

КВАЛІФІКАЦІЙНА РОБОТА

на здобуття освітнього ступеня

магістр

(назва освітнього ступеня)

на тему:

«Інформаційно-аналітична система прогнозування
портфеля криптовалютних фінансових інструментів»

Виконав: студент VI курсу, групи СПм-61
спеціальності

121 – Інженерія програмного забезпечення
(шифр і назва спеціальності)

(підпис) Лукасевич М.А.
(прізвище та ініціали)

Керівник (підпис) Пастух О.А.
(прізвище та ініціали)

Нормоконтроль (підпис) Стоянов Ю.М.
(прізвище та ініціали)

Завідувач кафедри (підпис) Петрик М.Р.
(прізвище та ініціали)

Рецензент (підпис) (прізвище та ініціали)

Міністерство освіти і науки України
Тернопільський національний технічний університет імені Івана Пулюя

Факультет комп'ютерно-інформаційних систем і програмної інженерії

(повна назва факультету)

Кафедра програмної інженерії

(повна назва кафедри)

ЗАТВЕРДЖУЮ

Завідувач кафедри

проф. Петрик М.Р.

(підпис)

(прізвище та ініціали)

« »

20__ р.

ЗАВДАННЯ
НА КВАЛІФІКАЦІЙНУ РОБОТУ

на здобуття освітнього ступеня

магістр

(назва освітнього ступеня)

за спеціальністю

121 Інженерія програмного забезпечення

(шифр і назва спеціальності)

студенту

Лукаsevичу Максиму Анатолійовичу

1. Тема роботи

Інформаційно-аналітична система прогнозування
портфеля криптовалютних фінансових інструментів

Керівник роботи

Пастух Олег Анатолійович, д.т.н., проф. каф. ПІ
(прізвище, ім'я, по батькові, науковий ступінь, вчене звання)

Затверджені наказом по університету від « 6 » грудня 2024 року № 4/7-1159

2. Термін подання студентом роботи 13.12.2024

3. Вихідні дані до роботи

наукові літературні джерела

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

1 Аналіз та класифікація методів прогнозування ціни активу. 2. Проектування рішення

3. Практична частина. 4. Охорона праці та безпека в надзвичайних ситуаціях

5. Перелік графічного матеріалу (з точним зазначенням обов'язкових креслень, слайдів)

1. Тема роботи. 2. Мета, задачі дослідження, об'єкт, предмет, наукова новизна

3. Способи прогнозування ціни активу. 4. Вимоги до системи.

5. Архітектура системи (діаграма компонентів UML). 6. Технологічний стек.

7. Діаграма класів UML для моделі даних системи. 8. Найбільш популярні криптовалюти.

9. Приклад зберігання даних часового ряду в датафреймі Pandas.

10. Розроблений алгоритм розрахунку та отримання значень індикаторів у комбінації
MACD+CCI

11. Діаграма діяльності UML для розробленого алгоритму прогнозування із застосуванням
комбінацій індикаторів.

12. Результати роботи невдалої та вдалої моделі SARIMA

13. Аналіз та порівняння результатів роботи різних рішень.

14. Фрагмент вихідного коду модуля прогнозування.

15. Основні результати проведеного дослідження.

6. Консультанти розділів роботи

Розділ	Прізвище, ініціали та посада консультанта	Підпис, дата	
		завдання видав	завдання прийняв
Охорона праці	Осухівська Г.М., к. т. н., доц., зав. каф. КС		
Безпека в надзвичайних ситуаціях			

7. Дата видачі завдання _____ 2024 р.

КАЛЕНДАРНИЙ ПЛАН

№ з/п	Назва етапів кваліфікаційної роботи	Термін виконання етапів роботи	Примітка
1.	Аналіз предметної галузі дослідження	19.09–27.09.24	Виконано
2.	Обґрунтування актуальності дослідження	30.09–10.10.24	Виконано
3.	Огляд існуючих механізмів та засобів	11.10–28.10.24	Виконано
4.	Проведення дослідження механізмів та засобів опрацювання даних	29.10–15.11.24	Виконано
5.	Оформлення розділу «Аналіз та класифікація методів прогнозування ціни активу»	16.11–28.11.24	Виконано
6.	Оформлення розділу «Проектування рішення»	29.11–05.12.24	Виконано
7.	Оформлення розділу «Практична частина»	06.12–10.12.24	Виконано
8.	Оформлення розділу «Охорона праці та безпека в надзвичайних ситуаціях»	05.12–12.12.24	Виконано
9.	Нормоконтроль	13.12–15.12.24	Виконано
10.	Перевірка на плагіат	13.12–17.12.24	Виконано
11.	Попередній захист роботи	20.12.24	Виконано
12.	Захист кваліфікаційної роботи	27.12.24	

Студент

(підпис)

Лукаsevич М.А.

(прізвище та ініціали)

Керівник роботи

(підпис)

Пастух О.А.

(прізвище та ініціали)

РЕФЕРАТ

Кваліфікаційна робота магістра містить: 72 сторінки, 17 рисунків, 36 джерел та 2 додатки.

ДАТАСЕТ, КРИПТОВАЛЮТА, МАШИННЕ НАВЧАННЯ ПРОГНОЗУВАННЯ КОТИРУВАНЬ, ТЕХНІЧНІ СТАТИСТИЧНІ ІНДИКАТОРИ

У кваліфікаційній роботі магістра проведено огляд торгових індикаторів та авторегресійних моделей для аналізу та передбачення руху тренду тимчасового ряду котирувань активу. Розроблено алгоритми для передбачення напрямку руху тренду котирувань. Виконана їх програмна реалізація, проведено аналіз їх роботи та здійснено візуалізацію отриманих даних як при застосуванні комплексу індикаторів, так і при застосуванні моделі із сезонним авторегресійним інтегрованим ковзним середнім.

Використана модель машинного навчання більш точно точніше і впевненіше прогнозує майбутню ціну та сам напрямок руху тренд криптовалюти, що відіграє важливу роль у торгівлі криптоактивами.

ABSTRACT

The qualifying thesis contains: 72 pages, 17 figures, 36 sources and 2 appends.

DATASET, CRYPTOCURRENCY, MACHINE LEARNING, QUOTE FORECAST, TECHNICAL STATISTICAL INDICATORS

In the master's thesis, a review of trading indicators and autoregression models was conducted to analyze and predict the movement of the trend of the time series of the asset's quotations. Algorithms have been developed for predicting the direction of the quote trend. Their software implementation was carried out, an analogue of their work was carried out and the visualization of the obtained data was carried out both when using a set of indicators and when using a model with a seasonal autoregressive integrated moving average.

The used machine learning model predicts the future price and the very direction of the cryptocurrency trend more accurately, more accurately and more confidently, which plays an important role in the trading of cryptoassets.

ПЕРЕЛІК СКОРОЧЕНЬ І ТЕРМІНІВ

ACD (Atomicity, Consistency, Isolation, Durability) - набір властивостей, що гарантують надійну роботу транзакцій бази даних: атомарність, узгодженість, ізолюваність, довговічність

ADX (Average directional index) – індекс середнього напрямку

API (Application Programming Interface) - прикладний програмний інтерфейс

AR (autoregressive) - авторегресія

ARIMA (autoregressive integrated moving average) - авторегресійне інтегроване ковзне середнє

BB (Bollinger Bands,) – Смуги Боллінджера

CCI (Commodity channel index) – індекс товарного каналу

EMA (Exponential Moving Average) – експоненційне згладжування

IDE (Integrated Development Environment) – інтегроване середовище розробки

HTML (HyperText Markup Language) – мова гіпертекстової розмітки

MA (moving average) - ковзне середнє

MACD (Moving Average Convergence/Divergence) – ковзна середня конвергенція/розбіжність

SAR (Stop and Revers) – індикатор трендів та розворотів

RSI (Relative strength index) – індекс відносної сили.

SARIMA (seasonal autoregressive integrated moving average) - сезонне авторегресійне інтегроване ковзне середнє

SARIMAX (Seasonal autoregressive–moving-average model with exogenous inputs) - сезонне авторегресійне інтегроване ковзне середнє з екзогенними регресорами

UTC (Coordinated Universal Time) – всесвітній координований час

БД – база даних

ПЕОМ - персональна електронно-обчислювальна машина

СУБД – система управління базами даних

ЗМІСТ

Вступ	9
1 Аналіз та класифікація методів прогнозування ціни активу	111
1.1 Прогнозування на основі моделей авторегресії	11
1.1.1 ARIMA	11
1.1.2 SARIMA	12
1.1.3 SARIMAX	13
1.2 Прогнозування на основі торгових індикаторів	133
1.2.1 RSI	14
1.2.2 CCI	15
1.2.3 MA	16
1.3 Групи індикаторів	20
1.4 Допустимі комбінації індикаторів для аналізу	23
1.5 Висновок до першого розділу	26
2 Проектування рішення	27
2.1 Вимоги до системи	27
2.2 Архітектура системи	29
2.3 Технологічний стек	32
2.4 Модель даних	34
2.5 Вибір торгової біржі та криптовалют для аналізу	36
2.5.1 Інформація про біржу та API	36
2.5.2 Вибір криптовалют для аналізу	37
2.6 Отримання історичних даних	39
2.7 Зберігання даних	43
2.8 Висновок до другого розділу	44
3 Практична частина	45
3.1 Бібліотеки	45
3.2 Визначення констант	46

3.3 Програмна реалізація алгоритму для передбачення поведінки ціни	47
3.3.1 Реалізація індикаторів на основі статистики	47
3.3.2 Реалізація авторегресійної моделі	50
3.4 Візуалізація та аналіз результатів прогнозування	52
3.4.1 Результат роботи RSI + CCI	52
3.4.2 Результат роботи EMA + CCI	52
3.4.3 Результат роботи MACD + CCI	54
3.4.4 Результат роботи MACD + Stochastic + 2 EMA	55
3.4.5 Результат роботи Double CCI + RSIOMA	56
3.4.6 Результат роботи BB + 3 EMA + MACD + RSI	57
3.4.7 Результат роботи авторегресійної моделі SARIMA	57
3.5 Порівняння найкращих рішень	59
3.6 Висновок до третього розділу	60
4 Охорона праці та безпека життєдіяльності	61
4.1 Охорона праці	61
4.2 Функціонування державної системи спостереження, збирання, обробки та аналізу інформації про стан довкілля під час надзвичайних ситуацій мирного та воєнного часу	64
4.3 Висновок до четвертого розділу	66
Висновки	67
Перелік джерел посилання	68
Додаток А. Тези конференції	
Додаток Б. Фрагмент вихідного коду модуля прогнозування	
Додаток В. Диск з роботою	

ВСТУП

Актуальність теми дослідження. В даний час все більше людей цікавиться трейдингом як додатковим або основним видом діяльності. Підприємства отримують ліцензії в Національному банку України і наймають десятки працівників для роботи у цій сфері. Глобалізація трейдингу, відповідно, дає негайні вигоди: швидше зростання у цій сфері, вищий рівень життя та нові економічні можливості як для населення, так і для країни.

Ринок криптовалют є досить молодим та перспективним. На просторах Інтернету можна знайти безліч рекомендацій, як заробити на крипторинку. Однак часто дані рекомендації несуть суперечливу інформацію, а часом навіть і шкідливу для криптотрейдера-початківця.

На криптовалютному ринку є велика конкуренція. Часто новачки втрачають свої заощадження, що дозволяє криптокитам постійно збільшувати свій капітал. Одним із найперспективніших напрямків заробітку на крипторинку є скальпінг. При скальпінгу необхідно проводити на день до 20 угод із невеликим сальдо. Для проведення угод необхідно знати точки входу в угоду та точки виходу, а дані точки безпосередньо залежать від багатьох факторів, що впливають. Актуальним завданням є розробка алгоритмів передбачення напрямку руху тренду котирувань та їх програмна реалізація.

Метою роботи є розробка інформаційно-аналітичної системи для прогнозування котирувань в галузі торгівлі криптовалютами активами.

Для досягнення мети потрібно вирішити наступні завдання:

- проведення літературного огляду в галузі статистичних індикаторів для технічного аналізу тимчасового ряду котирувань активу;
- проведення літературного огляду в галузі авторегресійних моделей для передбачення руху тренду тимчасового ряду котирувань активів ;
- розробка алгоритмів передбачення напряму руху тренду котирувань;
- програмна реалізація алгоритмів передбачення та аналіз отриманих

результатів.

Об'єкт дослідження: прогнозування курсу криптовалют.

Предмет дослідження: моделі для прогнозування курсу криптовалют.

Наукова новизна роботи: запропоновано використати модель авторегресії SARIMA для досягнення більшої точності прогнозування.

Практичне значення одержаних результатів. Розроблений програмний продукт дозволить криптотрейдерам значно покращити свої можливості при застосуванні торгових стратегій щодо криптовалют.

Апробація. Окремі результати дослідження доповідалися на XII науково-технічній конференції «Інформаційні моделі, системи та технології» як опубліковані тези.

1 АНАЛІЗ ТА КЛАСИФІКАЦІЯ МЕТОДІВ ПРОГНОЗУВАННЯ ЦІНИ АКТИВУ

В даний час для проблем прогнозування широко використовують штучний інтелект та нейронні мережі. Ці способи, безумовно, можна використовувати для прогнозування цінових рухів на ринку, де ціна активу водночас представляє тимчасовий ряд. Але, слід зазначити, що для торгівлі на малих таймфреймах, трейдерам потрібна швидкість прогнозування [1]. При скальпінгу час прийняття рішення при торгівлі досягає хвилинних інтервалів, отже, у зв'язку з тим, що штучний інтелект не хвалиться високою швидкістю тренування моделі та обчислення результатів, необхідно вдатися до інших рішень. У даному випадку слід розглянути можливість використання моделей авторегресії та класичних статистичних методів аналізу ринку. Під другим методом мається на увазі аналіз за допомогою торгових індикаторів на основі ковзної середовищем, її модифікацій та статистики. У комбінації, стратегія з індикаторами та моделлю авторегресії може досягати досить високої точності та швидкості, щоб реалізовувати успішну торгівлю активами.

1.1 Прогнозування на основі моделей авторегресії

Однією з найпопулярніших і найчастіше використовуваних моделей часових рядів є модель ARIMA. Вона має підкласи інших моделей, але для сезонного прогнозування часових рядів Бокса та Дженкінса (БД) пропонується досить вдалий варіант моделі ARIMA, а саме сезонний його варіант – SARIMA [2]. Кожна модель має свої переваги та недоліки, не існує єдиної моделі, яка перевершувала б інші моделі у всіх ситуаціях. У цій роботі будуть розглянуті основні моделі для роботи з прогнозуванням цінових тимчасових рядів, проте в результаті застосовується єдиний варіант у зв'язку з особливостями криптовалют.

1.1.1 ARIMA

Ця модель дає наступний крок у послідовності як лінійну функцію різних спостережень і залишкових помилок на попередніх тимчасових кроках [2].

Цей метод є поєднанням AR та MA, окрім того узгоджений з певними правилами етап попереднього опрацювання послідовності, з метою зробити її стаціонарною, можна застосувати термін «інтеграція» (I).

ARIMA окрім того прогнозує майбутні значення, використовуючи минулі значення. Однак цей метод прогнозування часових рядів недосконалий, оскільки не враховує сезонні тенденції. Він також припускає, що дані тимчасового ряду є стаціонарними, що означає, що їх статистичні властивості не зміняться з часом. Цей тип поведінки є ідеалізованим припущенням, яке, однак, не виконується на практиці, а це означає, що ARMA може давати спотворені результати.

Метод ARIMA підходить для одновимірних часових рядів із трендом та без присутності сезонних компонент [3].

1.1.2 SARIMA

В основі методу лежить моделювання наступного кроку у числовому ряді як лінійної функції різноманітних спостережень, помилок, різних сезонних спостережень та сезонних власне помилок на попередніх властиво тимчасових кроках. Є поєднанням SARIMA із перспективою виконувати ту ж авторегресію, диференціювання та моделювання ковзного середнього на сезонному рівні.

Метод добре зарекомендував себе для роботи із одновимірними часовими рядами із трендовими чи/або сезонними компонентами [2].

1.1.3 SARIMAX

Є розвитком попередньої моделі, котре містить на додачу симуляцію екзогенних величин. Інша назва таких величин - коваріати, і їх можливо вважати паралельними вхідними послідовностями, котрі мають спостереження на тих самих тимчасових кроках, що й вихідний ряд.

SARIMAX може бути використаний для симуляції включених моделей, які містять екзогенні змінні [3, 4]. Метод є застосовуваним для одновимірних часових рядів із наявністю трендових чи/альбо сезонних компонент та, властиво, екзогенних змінних.

1.2 Прогнозування на основі торгових індикаторів

У цьому випадку весь підхід до створення торгових стратегій можна грубо поділити на стратегії, що торгуються, ґрунтуючись на статистичних індикаторах, та стратегії, що ґрунтуються на поведінці учасників ринку. У першому випадку, наприклад, вибираються будь-які із запропонованих індикаторів, агрегуються, а потім на основі отриманих даних алгоритм заключає угоду. Слід зазначити, що технічні індикатори ґрунтуються на математичних розрахунках та статистиці [5]. Індикатори – це формули чи функції, побудовані на значеннях статистичних показників торгів. Ці формули засновані на минулих і реальних цінах з метою прогнозування очікуваної ціни. У другому випадку важливий критерій уміння торгівлі «руками», необхідно бачити та розуміти всі ситуації та причини руху тренду. Наприклад, аналіз ситуації у склянці на початку великих покупок (оголошується великий покупець), де поштовх ціни залежить від обсягу купівлі в моменті. Тут аналіз ситуації дозволяє ухвалити рішення для входу в угоду [6].

1.2.1 RSI

Цей індикатор є осцилятором, який коливається між двома крайніми значеннями від 0 до 100. Він вимірює імпульс і виділяється як один з найкращих інструментів для торгівлі на ринку тим, що він надає інформацію про перекупленість та перепроданість ринку.

Формула RSI є досить простою. Індикатор можна розбити на його основні компоненти: середній приріст, середній збиток і відносну силу. Для розрахунку використовується 14-денний період часу. Втрати повідомляються як позитивні значення, а не негативні [6, 7].

На рис. 1.1 показано приклад роботи індикатора RSI.

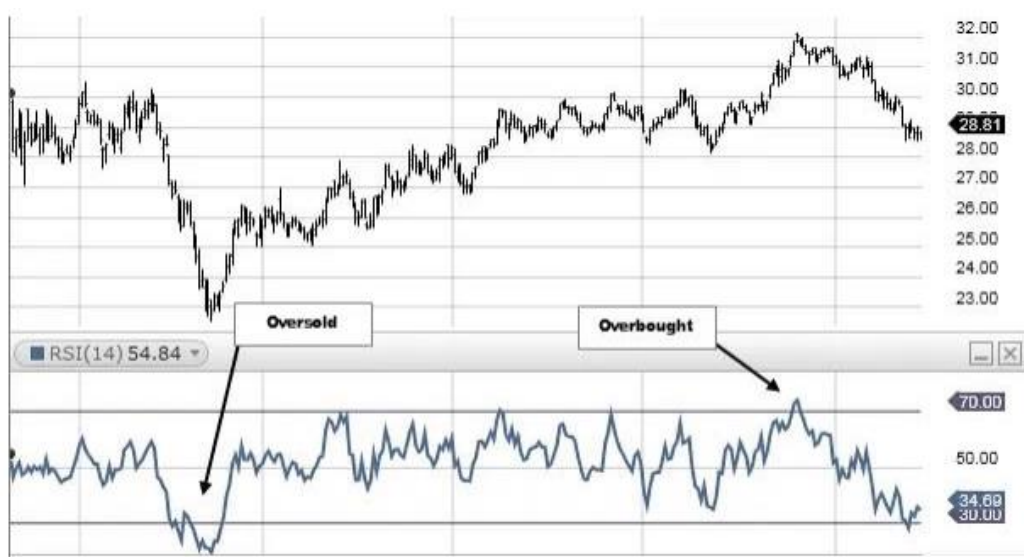


Рисунок 1.1 – Приклад роботи індикатора RSI

Щоразу, коли значення RSI вище 70, це інтерпретується як стан перекупленості ринку. У цьому випадку необхідно утримуватися від відкриття лонг-позицій, так як ймовірність розвороту ціни зазвичай висока. Так само щоразу, коли значення RSI нижче 30, це інтерпретується як стан перепроданості ринку. У цьому випадку ймовірність розвороту ціни і початку руху вгору зазвичай висока, оскільки базовий актив вважається дешевим.

1.2.2 CCI

Є важливим інструментом побудови графіків при аналізі, який вимірює поточну ціну активу щодо її середнього рівня ціни за певний період. Покази індикатора мають тенденцію бути відносно високими щоразу, коли переважна ціна активу набагато вища за середню [6, 7]. Так само показання індикатора мають тенденцію бути низькими кожного разу, коли ціна набагато нижча за середню.

На рис. 1.2 наведено приклад роботи індикатора CCI.



Рисунок 1.2 – Приклад роботи індикатора CCI

Як і індикатор RSI, індикатор CCI можна використовувати для визначення ринкових умов перекупленості і перепроданості досліджуваного активу. Щоразу, коли показання індикатора CCI вище +100, це означає, що починається сильний висхідний тренд, що сигналізує про стан перекупленості. У цьому випадку слід з обережністю відкривати лонг-позицію, оскільки ціна має тенденцію до розвороту. Так само щоразу, коли значення CCI нижче -100, це означає сильний низхідний тренд. У цьому випадку відкриваються лонг-позиції, щоб скористатися

потенційним розворотом ціни, оскільки актив вважається перепроданим.

1.2.3 МА

Це технічний індикатор, який показує середню ціну активу за визначений часовий проміжок. Він обчислюється як сума цін закриття активу за визначену кількість таких проміжків, яка ділиться на їх кількість. В результаті виходить згладжена лінія, яка слідує за рухом ціни активу [8, 9].

Існує три основні типи ковзних середніх: прості, експонентні та зважені. В основному при торгах розглядається експоненціальна ковзна середня, оскільки вона є найбільш популярною.

ЕМА. Є однією з основних провідних індикаторів від МА, описаної вище. Цей індикатор надає більшої ваги недавнім цінам і, отже, він чутливіший до не давнього цінового руху. ЕМА розраховується за формулою, яка приносить більшу вагу найостаннішим цінам.

Хоча існує безліч можливих варіантів коефіцієнта згладжування, найбільш поширеним варіантом вважається згладжування $= 2$. Це надає останньому спостереженню більшу вагу. Якщо коефіцієнт згладжування збільшується, пізніші спостереження надають більший вплив на ЕМА. Також, слід зазначити, що ЕМА для перших позицій у розрахунковому періоді вважається за допомогою SMA [3].

Williams Alligator (3 МА - з періодами 5, 8 та 13). Складається з 3-х ліній, тому він надає сильніші торгові сигнали, засновані на більшій кількості даних, ніж багато інших однолінійних індикаторів.

Для побудови торгових сигналів індикатор Alligator використовує коефіцієнти конвергенції та дивергенції. На практиці, на графіку зеленою лінією позначають «губи», червоною лінією – «зуби», синьою лінією – «щелепа». Кожна з ліній є згладженою ковзною середньою (SMMA) з періодами 5, 8, 13 і зсувами 2, 5, 8 відповідно.

При торгівлі з використанням індикатора Alligator необхідно звертати увагу на три ключові стани:

- «сон» - це стан, коли зелена, червона і синя лінії знаходяться дуже близько одна до одної або переплітаються, можна зробити висновок про відсутність тренду на ринку;
- «пробудження» - коли червона і синя ковзні середні рухаються в одному напрямку, зелена ковзна середня проходить через них, і всі три зрештою рухаються в одному напрямку, це знак того, що тренд ось-ось сформується;
- «голод» - коли зелена лінія знаходиться вище червоної лінії, а червона лінія вище синій лінії, спостерігається висхідний тренд, а коли ряд ліній змінюється на протилежний (синя на вершині, червона всередині, зелена в основі), ринок знаходиться в низхідному тренді. Розширення ліній у разі є підтвердженням тренду.

MACD. Індикатор MACD є так званим трендовим осцилятором, що поєднує в собі властивості і трендового індикатора, і осцилятора. Розраховується MACD на підставі ковзних середніх. Мета індикатора спростити візуальне сприйняття сигналів, що подаються ковзними середніми знизити запізнення і прибрати ряд недоліків, властивих звичайним трендовим індикаторам. Існує два способи побудови: лінійний MACD та гістограма MACD.

Parabolic SAR. Індикатор використовує метод трейлінг-стопу та розвороту, чи метод зупинки та розвороту, для визначення відповідних точок входу та виходу.

Цей ідикатор відображається на діаграмі як набір точок, котрі розміщені вище чи нижче ціни активу, залежно від напрямку руху ціни. Точка розміщується нижче за ціну, коли вона рухається вгору, і вище за ціну, коли вона рухається вниз. Параболічний індикатор генерує сигнали на купівлю або продаж, коли положення точок переміщається від однієї сторони ціни активу до іншої. Наприклад, сигнал на купівлю виникає, коли точки починають рух зверху-вниз за ціною, а сигнал на продаж виникає, коли точки починають рух знизу-вгору за ціною.

На графіку часового ряду Parabolic SAR рухається незалежно від того, чи

рухається ціна. Це означає, що якщо ціна спочатку зростає, а потім рухається горизонтально, PSAR продовжуватиме зростати, незважаючи на бічне рух ціни. З цієї причини сигнал розвороту на індикаторі не обов'язково означає розворот ціни.

Цей індикатор дає великі збитки у бічних трендах, тому він використовується в основному лише для виходу з позиції.

Stochastic. І стохастичний індикатор, і RSI є осциляторами імпульсу. Це робить їх дуже схожими за своєю метою, оскільки обидва вони використовуються для виявлення та прогнозування ринкових тенденцій. Незважаючи на те, що вони мають однакову мету, кожен досягає її зовсім по-різному. Стохастичний індикатор працює на основі припущення, що ціни закриття повинні рухатися в тому ж напрямку, що і основний тренд. RSI використовує швидкість руху ціни, щоб визначити, коли ціна перекуплена чи перепродана. Обидва вони корисні, хоча стохастичний індикатор, як правило, краще працює на бічних і мінливих ринках, тоді як RSI краще для трендових ринків.

Стохастичний осцилятор має 4 підвиди: повільний, швидкий, повний та Stochastic RSI. У бібліотеці TaLib, що використовується надалі в розробці: STOCH - повільний, STOCHF - швидкий, STOCHRSI - STOCHF, розрахований за даними RSI. Триперіодна ковзна середня %K, що називається %D, зазвичай включається як сигнальна лінія. Сигнали до угоди зазвичай подаються, коли %K перетинає %D.

ADX. Є індикатором технічного аналізу, котрий застосовується з метою з'ясування сили тренду. ADX не вказує напрямок або імпульс тренду. Це індикатор, що запізнюється, тобто тренд повинен встановитися до того, як ADX згенерує сигнал про початок тренду. Розрахункові значення ADX знаходяться в діапазоні від 0 до 100, але індикатор рідко заходить вище 60 пунктів.

Рівні відновлення Фібоначчі. Є технічним індикатором, який використовується для визначення потенційних рівнів підтримки та опору. Він заснований на ідеї, що ціни часто відновлюють передбачувану частину руху, після чого продовжують рух у початковому напрямку.

Щоб розрахувати рівні відновлення Фібоначчі, проводиться лінія тренду між двома крайніми точками на графіку, а потім ділиться відстань по вертикалі на

ключові коефіцієнти Фібоначчі: 23,6%, 38,2%, 50%, 61,8% та 100%. Потім ці рівні можна використовувати для визначення потенційних областей, де ціна може відчувати підтримку або опір.

Щоб інтерпретувати відновлення Фібоначчі, проводиться аналіз цінової поведінки до досягнення ключових рівнів, а потім спостерігається реакція. Якщо ціна відскакує від ключового рівня, це може бути ознакою підтримки чи опору. Якщо ціна пробиває ключовий рівень, то це може бути ознакою розвороту тренду.

Переваги використання відновлення Фібоначчі включають можливість визначати потенційні галузі підтримки та опору та підтверджувати угоди.

Pivot Points. Один із найефективніших інструментів підтримки та опору серед усіх технічних індикаторів, про які вже йшлося раніше. Рекомендується використовувати його як точку установки для ваших ордерів стоп-лосс і тейк-профіт. Точки розвороту розраховують середнє значення мінімуму, максимуму, ціни відкриття та закриття кожної з останніх свічок. Точки розвороту краще працюють у короткостроковій перспективі (внутрішні угоди та скальпінг). Він вважається дуже об'єктивним інструментом, подібним до Фібоначчі, що допомагає уникнути суб'єктивних інтерпретацій.

BB. Складаються із трьох ліній. Один з найбільш поширених розрахунків використовує 20-денну просту ковзну середню (SMA) для середньої смуги. Верхня смуга розраховується шляхом взяття середньої смуги та додавання до цієї суми подвоєного денного стандартного відхилення. Нижня смуга розраховується шляхом віднімання середньої смуги з відрахуванням дворазового денного стандартного відхилення.

У BB оптимальне значення, значення за замовчуванням, є MA20 і стандартне відхилення рівне 2. У такому разі 95.5% цін потраплять у цей діапазон і даватиме оптимальну кількість сигналів. Якщо відхилення дорівнюватиме 1 - то всього 68.3% цін будуть знаходитися в цьому діапазоні, і ми побачимо занадто багато сигналів. Якщо відхилення буде дорівнює 3 - то в діапазоні буде 99,7% цін, але тоді сигналів буде занадто мало, і вони будуть рідкісні, так як практично 100% цін будуть в діапазоні.

Стандартний підхід при використанні ВВ полягає у визначенні умов перекупленості або перепроданості ринку. Коли ціна активу проривається нижче смуги нижньої ВВ, ціни, можливо, впали занадто сильно і повинні відскочити. З іншого боку, коли ціна пробиває верхню смугу, ринок, можливо, перекуплений і чекає на відкат [7].

1.3 Групи індикаторів

Є десятки індикаторів, які можна використовувати, але вони з іншого боку також поділяються на групи за типом інформації, яку вони надають. Деякі індикатори можуть надавати кілька наборів інформації і, отже, можуть належати до однієї групи [7, 9].

Відправною точкою є поділ різних типів індикаторів на основі їхньої здатності передбачати ціну або визначати сигнал після руху ціни. Індикатори поділяються на великі групи.

Випереджаючі індикатори - це тип індикатора, здатний передбачити майбутній рух ціни. Через їхню передбачувану силу ці індикатори використовуються для визначення потенційних розворотів ціни до події.

Негативна сторона використання виключно випереджальних індикаторів полягає в тому, що вони можуть давати кілька помилкових сигналів, оскільки заздалегідь передбачають напрямок та рівень ціни. Можна підтвердити ці сигнали за допомогою індикаторів, що запізнюються, але це буде означати втрату частини прибутку при вході в позицію після підтвердження. Прикладами випереджуючих індикаторів є RSI, Stochastic, CCI, Parabolic SAR, балансовий об'єм (OBV) та осцилятор Чайкіна.

Індикатори, що запізнюються - це індикатори, які йдуть за ціною і зазвичай подають сигнали після того, як відбулися зміни ціни. Індикатори можна використовувати для підтвердження попереджень та сигналів від інших

індикаторів. Типовими прикладами індикаторів, що запізнюються, є ковзні середні, MACD, смуги Боллінджера, ATR, стандартне відхилення і швидкість зміни обсягу.

З погляду аспекту, аналізованого індикаторами, їх можна розділити на чотири основні категорії. Дані категорії та можливі їх комбінації відображені в табл. 1.1.

Таблиця 1.1 – Основні категорії індикаторів

Імпульсні індикатори	Трендові індикатори	Волатильні індикатори	Індикатори об'єму
Стохастичний осцилятор	ADX	Смуги Боллінджера	OBV
RSI	Parabolic SAR	ATR	Індикатор накопичення/розподілу
MACD	MA	Standard Deviation	
CCI	MACD	Канал Дончяна	
MA	Індикатор Ішімоку		

Трендові індикатори. У трейдингу під трендом розуміється напрямок руху ціни протягом тривалого періоду часу. Наприклад, коли ціна постійно зростає, це висхідний тренд, а коли вона знижується, це спадний тренд. Індикатори тренду можуть допомогти визначити напрямок руху ринку.

Деякі поширені індикатори тренду: MA (SMA та EMA), Parabolic SAR, MACD, індикатор Ішімоку, рівні відновлення Фібоначі та точки розвороту. Хоча останні два індикатори можна віднести до такої окремої підкатегорії, як індикатори рівня підтримки/опору. Вони особливо корисні при спробі виявити прориви в тренді, щоб встановити точки входу та виходу.

Індикатори обсягу. У трейдингу під обсягом розуміється кількість угод, скоєних протягом певного періоду часу. Це пряма вказівка на попит та пропозицію активу. Вимірюючи обсяг торгів, індикатори обсягу показують, чи тенденція

триватиме. Наприклад, великий обсяг на висхідному тренді вказує на високий попит і, отже, подальше підвищення ціни. Так само це працює і у зворотному сценарії: великий обсяг на низхідному тренді вказує на високу пропозицію і ймовірність подальшого падіння ціни.

Деякі поширені індикатори обсягу: балансовий обсяг (OBV), індикатор накопичення/розподілу (ADL) та індекс грошового струму (MFI).

Індикатори імпульсу. Ці індикатори попереджають трейдерів про початок тренду. Вони виступають у ролі інформаторів – повідомляють, коли з'являється тренд. Прикладами індикаторів імпульсу є MA, MACD, RSI, стохастичний осцилятор .

Можна відзначити деякі особливості індикаторів імпульсу. До плюсів належить те, що з ними безпечніше торгувати. Вони дають вищі результати, якщо правильно їх використовувати. До мінусів можна віднести те, що вони іноді «промахуються», даючи результат надто пізно, пропускаючи важливі зміни.

Індикатори волатильності. Волатильність у трейдингу визначає ступінь зміни ціни у часі . Висока волатильність вказує на швидкі та непередбачувані зміни ціни. Індикатори волатильності вимірюють діапазон цін активу і допомагають вловити моменти високої волатильності. Найчастіше віддається перевага активам і ринкам з високою волатильністю , тому що вони надають численні можливості для торгівлі, а також відносно швидкий і високий прибуток [10].

Найбільш поширеними індикаторами волатильності є: BB, канал Дончяна, ATR (середній дійсний діапазон), стандартний розподіл.

Для короткострокової торгівлі найбільш відповідними індикаторами є: SMA, EMA, MACD, Parabolic SAR, Stochastic Oscillator, BB. Як можна помітити, деякі з перерахованих індикаторів не рідко застосовують різні напрямки.

1.4 Допустимі комбінації індикаторів для аналізу

Як можна помітити, існує велика різноманітність технічних індикаторів і не обов'язково використовувати їх усі. Трейдери не повинні використовувати занадто багато інструментів. Робота з більш ніж 3 інструментами сповільнює роботу і призведе до помилок. Як і в будь-якій іншій галузі життя, на графіку прогресу є точка, після порушення якої ефективність починає падати. Ідея полягає в тому, щоб вибрати 2-3 потужні, ефективні інструменти для комфортної роботи [7, 8].

Однак, варто зазначити, що найчастіше комбінуються дві однакові категорії індикаторів разом. Наприклад, якщо об'єднати два індикатори імпульсу, такі як RSI і ковзне середнє, вони дадуть повторювані результати. Далі будуть розглянуті допустимі комбінації індикаторів, які вже були протестовані співтовариством трейдерів і мають гарну репутацію.

RSI + CCI. Щоб використовувати два індикатори, по-перше, необхідно налаштувати параметри RSI на 30-періодний RSI замість 14-періодного за замовчуванням. Зміна налаштувань допомагає захиститися від короткострокових коливань цін, які можуть ввести в оману. Аналогічним чином додається лінія 50 на індикаторі RSI на додаток до ліній тренду за умовчанням 70 та 30, які вказують на умови перекупленості та перепроданості.

Приклад спільної роботи обох індикаторів відображено на рис. 1.3.



Рисунок 1.3 – Приклад роботи комбінації статистичних індикаторів RSI та CCI

Щодо індикатора CCI, було б розумно дотримуватися 20-першого одного CCI, оскільки він має тенденцію бути більш стійким до короткострокових змін цін, які можуть ввести в оману. Але на коротких таймфреймах можна встановлювати і 6-періодний варіант. Обмеження перекупленості та перепроданості на рівні +100 та -100 повинні залишитися незмінними.

CCI + ЕМА. ЕМА використовується визначення тренда. Осцилятор CCI подає сигнали для входу в угоду. Необхідно спостерігати за ЕМА, щоб визначити тренд, а потім чекати сигналу від CCI, який надходить, коли лінія 100 (або -100) перетинається.

Зміна періодів також значно впливає на результат передбачення. Період ЕМА збільшується при використанні коротших таймф-реймів. Для CCI слід використовувати період 14, однак інші періоди не виключаються. Важливою умовою є використання періоду до 24.

CCI + MACD. Для роботи з даною стратегією найчастіше вибирається таймфрейм 4 -годинний або денний, але також дана комбінація застосовується і на малих скальпінгових таймфреймах.

Принцип стратегії полягає в тому, що перетин індикатором CCI своїх рівнів +100 і -100 є основним сигналом для відкриття позиції, але фільтруються вони за допомогою індикатора MACD. MACD немає переваг у визначенні часу входу через його щодо очевидного відставання.

RSI + Alligator. Комбінація діє на будь-яких торгових парах і малих таймфреймах. Алгоритм аналізу та купівлі активу виконується в наступній послідовності:

- «губа» Алігатора (зелена) перетинає «щелепу» (синя) знизу-вгору;
- "зуби" Алігатора (червона) перетинають "щелепу" (синя) знизу-вгору;
- RSI 8 долає позначку зі значенням 50 знизу-вгору;
- провадиться вхід в угоду на купівлю при відкритті наступної свічки;
- встановлюється стоп-лосс нижче останнього мінімуму коливання ціни;

- встановлюється тейк-профіт на таку кількість піпсов (процентних пунктів) від ціни входу, як і стоп-лосс, тобто. співвідношення ризик/прибуток 1:1 або більше.

RSI + MACD. Така комбінація відмінно підходить для торгівлі на малих таймфреймах, так як вона не запізнюється за ціною і вміло відфільтровує ринковий шум. Вона може замінити собою будь-який осцилятор або трендовий інструмент. Можна сміливо сказати, що це універсальна комбінація, яку можна адаптувати під будь-яку стратегію.

Stochastic + MA. Запізнювальна продуктивність ковзних середніх, доповнена чуйністю та підвищеною якістю сигналів, що демонструються стоха стичними осциляторами, робить цю комбінацію надзвичайно універсальним інструментом для всіх випадків торгівлі на ринку. Стохастичний осцилятор шукає зони перепроданості та перекупленості, використовуючи максимуми та мінімуми, використовуючи формулу нормалізації.

ADX + RSI + BB. Комбінація даних статистичних індикаторів добре підходить для скальпінгу, де ADX та RSI присвоюються періоди 14 та 7, відповідно. BB присвоюються період 20 до середньої лінії (SMA) і стандартне відхилення 2.

MA + MACD + рівні Фібоначчі. Рівні Фібоначчі часто використовуються для підтвердження точок входу або встановлення стоп-лоссів та тейк-профітів.

По-перше, необхідно реалізувати трендовий індикатор та осциль тор. Період MA дорівнює 10, а MACD встановлений за умовчанням.

Використовується наступний торговий алгоритм (в даному випадку для ордеру на продаж):

- ціна перетинає ковзну середню знизу-вгору;
- гістограма MACD перетинає лінію у тому напрямі;
- сітка Фібоначчі витягнута вздовж останньої видимої хвилі тренду.

Якщо ціна перетинає ковзну середню одночасно з відскоком або пробиттям рівня в бажаному напрямку, відкривається ордер на продаж. Тейк-профіт можна

встановити на одному з наступних рівнів Фібоначчі. Стоп-лос може бути встановлений на попередньому рівні.

1.5 Висновок до першого розділу

В цьому розділі проаналізовано предметну область дослідження та наведено класифікацію методів прогнозування ціни активу.

Описано прогнозування на основі моделей авторегресії (ARIMA, SARIMA, SARIMAX) та прогнозування на основі торгових індикаторів (RCI, CCI, MA).

Досліджено групи індикаторів та їх комбінації для проведення аналізу.

2 ПРОЕКТУВАННЯ РІШЕННЯ

2.1 Вимоги до системи

Розроблювана інформаційно-аналітична система повинна задовольняти такі вимоги:

- інтеграція із зовнішніми джерелами даних. Система повинна мати можливість отримувати актуальні історичні дані про котирування криптовалютних активів із зовнішніх джерел, таких як API криптовалютних бірж або сервіси, що надають фінансові дані;
- зберігання та управління даними. Система має мати БД для зберігання та управління історичними та поточними даними про котирування криптовалютних активів, а також інформацією про користувачів та їх торговельні операції;
- аналіз та обробка даних. Система має обробляти отримані дані та застосовувати статистичні моделі, алгоритми роботи торгових індикаторів та інші методи аналізу для прогнозування майбутніх котирувань криптовалют;
- візуалізація результатів. Розробка повинна давати юзеру наочну візуалізацію прогнозів та результатів аналізу даних, такі як графіки та сигнали, які допоможуть трейдерам приймати поінформовані рішення;
- інтерфейс користувача. Розробка має володіти зручним та нативним користувацьким інтерфейсом, який дозволить трейдерам взаємодіяти з системою;
- аутентифікація та авторизація. В системі повинна забезпечуватися безпека та захист даних користувачів, включаючи механізми автентифікації та авторизації, щоб тільки авторизовані користувачі могли одержувати доступ до системи та їх власних даних;
- інтеграція з торговими платформами. Система може надавати можливість інтеграції з торговельними платформами, щоб користувачі могли

автоматизувати свої торгові стратегії та виконувати операції на основі прогнозів системи;

- масштабованість і продуктивність. Система повинна бути здатна обробляти великі обсяги даних і забезпечувати високу продуктивність у режимі реального часу, щоб трейдери могли отримувати актуальні та точні прогнози.

На основі цих вимог були поставлені такі завдання, після виконання яких інформаційно-аналітична система для прогнозування котирувань вважатиметься мінімально життєздатним продуктом.

Завдання інтеграції даних:

- реалізувати модуль отримання даних про котирування криптових лютних активів із зовнішніх джерел;
- створити механізм збереження та оновлення отриманих даних у сховищі.

Завдання аналізу даних:

- розробити модуль для попереднього опрацювання даних, включно з очищенням, перетворенням та фільтрацією;
- реалізувати алгоритми статистичного аналізу та прогнозування на основі доступних даних.

Завдання візуалізації даних:

- створити інтерфейс для відображення графіків і сигналів, що представляють історичні дані та прогнози.

Завдання інтерфейсу користувача:

- розробити інтерфейс для аутентифікації та авторизації користувачів.

Завдання безпеки та захисту даних:

- впровадити механізми автентифікації та авторизації користувачів.

Завдання інтеграції з торговими платформами:

- Дослідити та інтегрувати API торгових платформ для автоматизації торговельних стратегій та виконання операцій на основі прогнозів системи.

Завдання продуктивності та масштабованості:

- оптимізувати процеси отримання даних, обробки та аналізу для забезпечення високої продуктивності системи при великих обсягах даних;
- передбачити масштабованість системи, щоб вона могла обробляти зростаючу кількість користувачів та даних.

2.2 Архітектура системи

На основі зібраної інформації, інформаційно-аналітична система для прогнозування котирувань у галузі торгівлі криптовалютами матиме багаторівневу архітектуру. Ця архітектура включає різні компоненти, такі як інтерфейс користувача, модуль прогнозування, модуль аналізу даних, БД і зовнішні джерела даних.

Інтерфейс користувача забезпечує взаємодію користувачів з системою, дозволяючи їм вводити параметри прогнозування, проглядати графіки та отримувати результати прогнозування.

Модуль прогнозування відповідає за розробку та застосування моделей прогнозування, таких як SARIMA, та використання індикаторів на основі статистики для створення прогнозів котирувань.

Модуль аналізу даних відповідає за збір, обробку та попередній аналіз історичних даних про котирування криптовалют. Він може включати операції передобробки даних, обчислення статистичних показників, створення графіків та інші аналітичні операції.

БД, така як PostgreSQL, використовується для зберігання та управління історичними даними про котирування, а також для зберігання інформації про користувачів-трейдерів.

Зовнішні джерела даних надають актуальні дані про котирування

криптовалют та іншу пов'язану інформацію, яка може бути використана для оновлення та розширення системи.

Таким чином, система має багаторівневу архітектуру, яка забезпечує поділ функціональності на різні компоненти та дозволяє ефективно виконувати завдання прогнозування та аналізу котирувань криптовалют.

Архітектура системи, що розробляється, зображена на рис. 2.1 як діаграма компонентів UML.

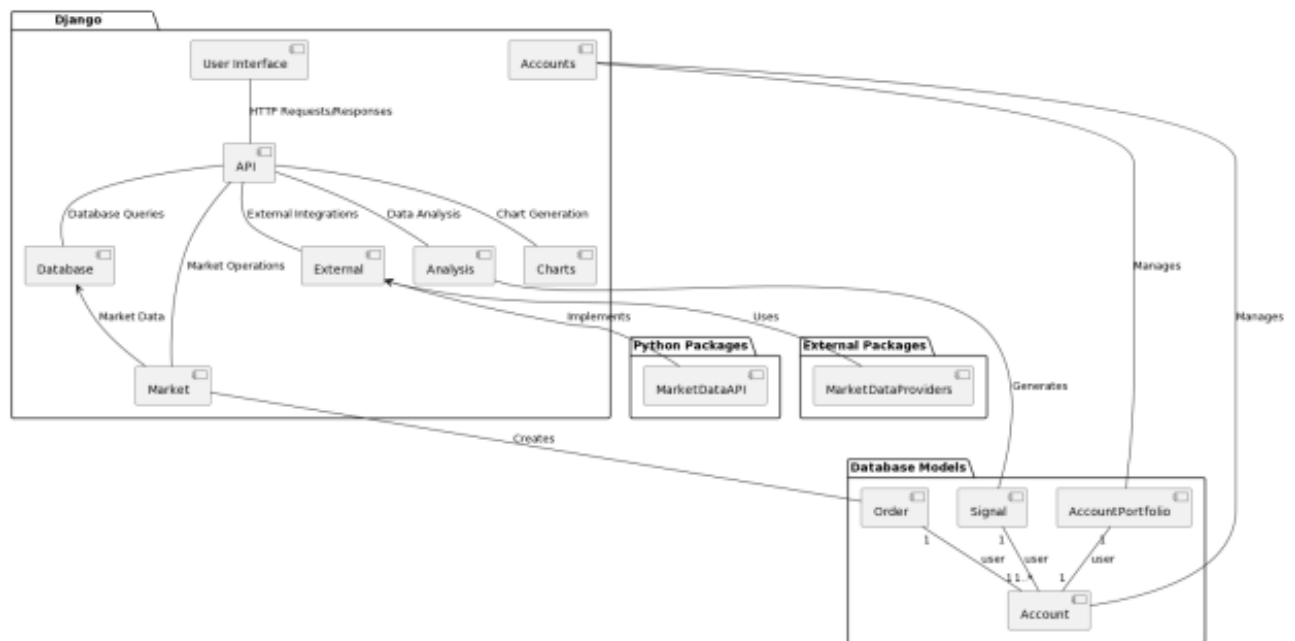


Рисунок 2.1 – Архітектура системи, що розробляється

Варто розглянути кожен компонент архітектури докладніше.

Інтерфейс користувача (UI). Відповідає за подання та взаємодію з користувачем. Він обробляє HTTP-запити і відображає веб-сторінки, які дозволяють користувачам взаємодіяти з системою.

API. Цей компонент надає інтерфейс для взаємодії з системою через HTTP - запити та відповіді. Він обробляє запити від інтерфейсу користувача і виконує відповідні операції в системі, включаючи операції з БД, операції на ринку та інтеграції із зовнішніми сервісами.

БД (DB). Відповідає за зберігання та керування даними системи. Він містить

таблиці та моделі даних, такі як акаунти трейдерів, портфелі, сигнали та ордери. БД використовується для збереження та вилучення даних під час операцій та аналізу.

Ринок (Market). Надає функціональність, пов'язану з торгівлею криптовалютами активами. Він виконує операції на ринку, такі як створення ордерів на купівлю та продаж активів, отримання даних про ринкові котирування та виконання торгових стратегій.

Зовнішні сервіси (External). Цей компонент інтегрується із зовнішніми провайдерами даних про ринкові котирування. Він використовує зовнішні API, такі як MarketDataProviders, для отримання актуальних даних про котирування та їх оновлення. Джерелами ринкових даних можуть бути API бірж, послуги постачальників даних про котирування та інші зовнішні ресурси. Пакет «External» забезпечує зв'язок між системою та цими зовнішніми джерелами даних, дозволяючи отримувати актуальні ринкові дані для аналізу та прогнозування. Цей елемент є важливою частиною архітектури системи, оскільки забезпечує доступ до актуальних зовнішніх даних, які можуть бути використані всередині системи для аналізу, прогнозування та прийняття торгових рішень. При цьому, зазначимо, що «MarketDataAPI» визначає загальний інтерфейс для доступу до ринкових даних, у той час як «MarketDataProviders» надає конкретні реалізації та провайдерів даних, які можуть бути використані для доступу до конкретних джерел даних.

Аналіз (Analysis). Відповідає за аналіз даних про ринкові котирування та генерацію прогнозів. Він використовує моделі прогнозування, такі як SARIMA, та статистичні індикатори, щоб визначити тренди та патерни на ринку та передбачити майбутні котирування.

Акаунти (Accounts). Цей компонент керує обліковими записами трейдерів, їх портфелями та операціями з ними. Він дозволяє трейдерам створювати та керувати своїми акаунтами, відкривати та закривати позиції, контролювати свої інвестиції та переглядати звіти про свою діяльність.

Генерація графіків (Charts). Відповідає за створення графіків та візуалізацію даних про ринкові котирування. Він використовує дані з БД та результати аналізу,

щоб створювати графічні представлення, які допомагають трейдерам візуалізувати та аналізувати ринкові дані.

2.3 Технологічний стек

Опишемо технологічний стек та інструменти, які будуть використані при розробці інформаційно-аналітичної системи для прогнозування котирувань у галузі торгівлі криптовалютними активами.

Враховуючи переваги та застосування технологій у контексті даної системи, виділимо деякі моменти.

При розробці системи буде використано:

- Python - є універсальною мовою програмування з багатою екосистемою та простим синтаксисом;
- Django - забезпечує зручну розробку веб-додатків з адміністративною панеллю та обробкою HTTP -запитів;
- PostgreSQL - забезпечує надійне зберігання даних з підтримкою ACID - транзакцій;
- matplotlib - надає потужні можливості візуалізації даних;
- Pandas – полегшує обробку та аналіз табличних даних.

Також у розробці системи використовується система контролю версій Git, яка полегшує керування та спільну розробку коду.

Використовувані технології при розробці системи зображені у круговій діаграмі на рис. 2.2.

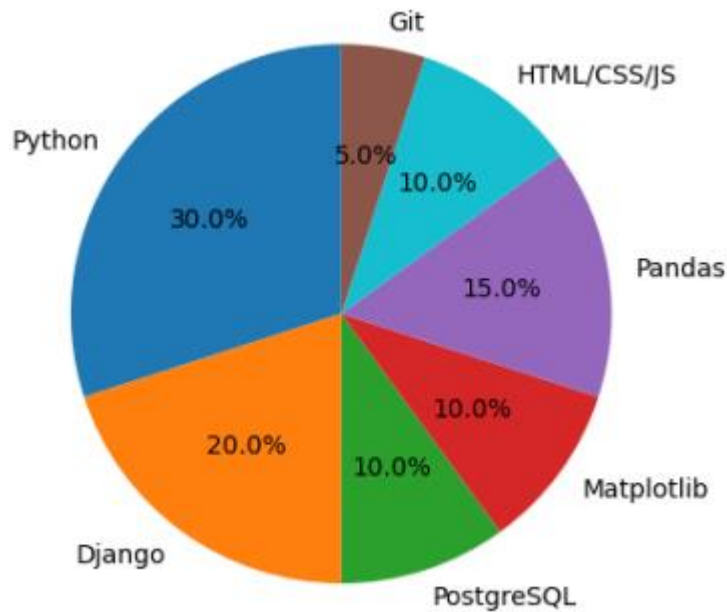


Рисунок 2.2 – Технології, що використовуються в системі

Python і R є популярними мовами програмування для аналізу даних та машинного навчання. У контексті даної системи, де потрібне прогнозування котирувань криптовалютних активів, Python є кращим вибором. Python має широкий спектр бібліотек для аналізу даних, машинного навчання і візуалізації, таких як Pandas, NumPy, scikit-learn і matplotlib. Python також володіє простим синтаксисом, який спрощує читання та підтримку коду.

Django має багаті можливості для розробки та управління веб-додатками, що робить його зручним вибором для побудови інформаційно-аналітичної системи з адміністративною панеллю та обробкою HTTP - запитів. У порівнянні з Flask, Django надає повноцінний фреймворк з безліччю вбудованих функцій і модулів, що спрощує та прискорює розробку. Flask, з іншого боку, є легковаговим і гнучким фреймворком, який надає базовий функціонал, але потребує більшого налаштування. У контексті даного програмного рішення, де потрібна розробка складної інформаційно-аналітичної системи, Django є більш відповідним вибором, оскільки надає інтеграцію з БД, адміністративну панель і безліч розширень для швидкої розробки.

PostgreSQL і MySQL є двома популярними реляційними БД з відмінною

продуктивністю та надійністю. У контексті даної системи, де потрібне зберігання та управління даними про торгові операції та аналітичні дані, PostgreSQL буде кращим. PostgreSQL пропонує широкі можливості по роботі з даними, включаючи підтримку ACID -транзакцій, гнучку схему даних та потужні засоби аналітики.

Matplotlib надає широкі можливості для візуалізації даних, що дозволяє зручно відображати результати прогнозування та аналізу даних у системі.

Pandas полегшує обробку та аналіз табличних даних, що особливо корисно при роботі з фінансовими часовими рядами та статистичними даними в системі.

Ці технології в сукупності забезпечують надійність, ефективність та зручність розробки інформаційно-аналітичної системи для прогнозування котирувань.

2.4 Модель даних

Модель даних є комплексною структурою, яка охоплює різні аспекти інформаційно-аналітичної системи для прогнозування котирувань криптовалютних активів. Вона включає моделі, такі як Account, AccountPortfolio, Order, Signal та Prediction, які взаємодіють один з одним для забезпечення функціональності системи.

Модель даних містить інформацію про трейдерів та їх акаунти (Account), їх портфелі (AccountPortfolio) з балансами і прибутком/збитком, ордери на купівлю та продаж (Order), сигнали (Signal), які рекомендують певні дії з торгівлі, та передбачення (Prediction), які містять інформацію про прогнозовані значення і пов'язані з певними сигналами.

Ця модель даних забезпечує систему засобами зберігання та організації даних, а також зв'язки між різними компонентами, дозволяючи системі ефективно обробляти та використовувати інформацію для прийняття рішень та надання користувачеві актуальних та точних результатів аналізу та прогнозування [11].

Кожен користувач (Account) має свій акаунтний портфель (AccountPortfolio), в якому зберігається інформація про баланс та прибуток. Користувачі можуть створювати замовлення, щоб купити чи продати криптовалютні активи. Кожне замовлення містить інформацію про символ активу, кількість та напрямок операції.

Діаграма класів UML для моделі даних у системі зображено на рис. 2.3.

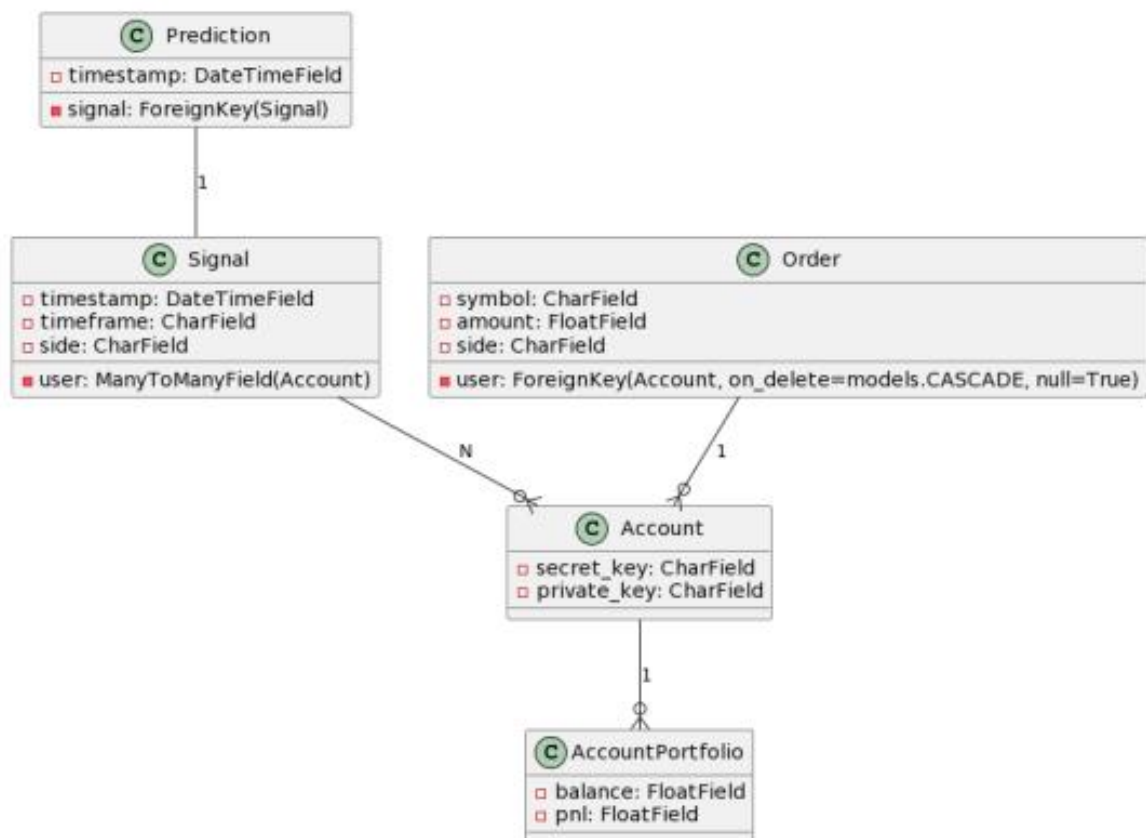


Рисунок 2.3 – Діаграма класів UML для моделі даних у системі

Дамо деякий опис представленої діаграми.

- відношення між класами Account та AccountPortfolio: 1 свідчить про зв'язок "один до одного" (One-to-One) між Account та AccountPortfolio. Один обліковий запис може перебувати у зв'язку лише з одним портфелем облікового запису;
- відношення між класами Signal та Account: N каже про зв'язок "багато до багатьох" (Many-to-Many) між Signal і Account. Один сигнал може бути пов'язаний з декількома обліковими записами, і один обліковий запис може бути

пов'язаний з кількома сигналами;

- відношення між класами Order та Account: тут 1 свідчить про зв'язок "один до багатьох" (One-to-Many) між Order та Account. Одне замовлення може бути пов'язане тільки із одним обліковим записом, але один обліковий запис може мати кілька замовлень.

- відношення між класами Prediction та Signal: цифра 1 вказує на зв'язок "один до одного" (One-to-One) між Prediction та Signal. Один прогноз може перебувати у зв'язку тільки з одним сигналом.

Таким чином, модель даних підтримує зберігання інформації про користувачів, портфелі, сигнали, прогнози та замовлення, що дозволяє системі ефективно прогнозувати і управляти торговими операціями з криптовалютами.

2.5 Вибір торгової біржі та криптовалют для аналізу

2.5.1 Інформація про біржу та API

За останні кілька років все більше людей почали торгувати крипто валютами, і по всьому світу існує безліч бірж, які підтримують різні криптовалюти. У цій роботі ми зосередимося на біржі Binance, яка сьогодні є найпопулярнішою біржею у світі [12, 13]. Вона підтримує практично всі криптовалюти та доступна у багатьох країнах.

Binance надає два види доступу до API [12, 14]:

- actual Binance API;
- test Binance API.

Actual API забезпечує прямий доступ до фактичного облікового запису, і будь-які угоди, вчинені за допомогою цього API, будуть відображатися в реальному обліковому записі.

Test Binance API дає ті ж функції, як і з реальним API. На баланс користувача

нараховуються стартові грошові одиниці для тестування. Тут можна переконатися, що розроблений додаток працює нормально.

Binance не надає бібліотеку Python для взаємодії з API, але є відома стороння бібліотека python-binance, яку можна використовувати для взаємодії з API [13].

Також необхідно зазначити, що ліміти курсу Binance становлять: 1200 запитів на хвилину, 10 ордерів на секунду та 100 000 ордерів на добу [14].

Деякі запити мають більшу вагу, ніж інші, особливо якщо запит повертає інформацію про всі активи.

2.5.2 Вибір криптовалют для аналізу

У скальпінгу необхідно отримувати прибуток від невеликих рухів цін, а на ринку криптовалют, де гарантована чимала волатильність, це дає чудову можливість для торгівлі. У зв'язку з цим, для скальпінгу переважно вибираються найбільш волатильні активи [15].

Для аналізу, безумовно, можна вибрати наступні 13 найбільш популярних криптовалют, котрі наведені в табл. 2.1.

Таблиця 2.1 – Найбільш популярні криптовалюти

Рік появи	Валюта	Символ
2009	Bitcoin	BTC
2011	Litecoin	LTC
2013	Dogecoin	DOGE
2013	Ripple	XRP
2014	Cosmos	ATOM
2015	Ethereum	ETH

Продовження таблиці 2.1

Рік появи	Валюта	Символ
2017	Cardano	ADA
2017	Binance Coin	BNB
2017	TRON	TRX
2017	Polygon	MATIC
2020	SHIBA INU	SHIB
2020	Polkadot	DOT
2020	Solana	SOL

Але для тестування програмно-розробленого алгоритму на даному етапі дослідження достатньо одного-двох інструментів. Всі перелічені криптовалюти користуються популярністю у співтоваристві трейдерів та інвесторів, тому їх ліквідність має високі показники. Висока ліквідність робить актив більш стійким до різких стрибків ціни на ринку, але зі збільшенням ліквідності знижується волатильність, яка також є важливою для отримання хорошого прибутку [16, 17]. Також мають місце деякі особливості поведінки кожного активу на ринку. Так, наприклад, біткойн - найбільша та найбільш ліквідна монета на ринку криптовалют, що робить її ідеальним варіантом для денної торгівлі та стратегій скальпінгу [18].

У криптовалюти Ethereum також непогана ліквідність, але є імпульсні ділянки часового ряду, коли ціна швидко набирає або знижує обороти.

Криптовалюта XRP завдяки своїм унікальним функціям, які надає цифровому світу, щодня демонструє значну волатильність та ліквідність, що робить її відповідним варіантом для короткострокових торгових стратегій.

BNB також через свою популярність має особливо високі обсяги торгів, волатильність і ліквідність, що робить його одним із найкращих варіантів для скальпінгу.

Litecoin - досить популярна криптовалюта, якою торгують при скальпінгу через її великий обсяг торгів, ліквідності та значні коливання цін, які LTC може

випробувати за один торговий день [19].

Для тестування розробленого алгоритму передбачення напрямки руху тренда часового ряду у цій роботі використовується LTC.

2.6 Отримання історичних даних

За проведеними спостереженнями, тестовий API не надає реальних історичних даних; натомість він надає фіктивні дані. Таким чином, для доступу до реальних даних у роботі використовується реальний API та фактичні ключі API. Можливе використання Binance WebSocket для потокової передачі даних в реальному часі. Для завантаження історичних ринкових даних через API у табл. 2.2 визначено типи даних та відповідні їм ендпоінти.

Таблиця 2.2 – Типи даних Binance API та ендпоінти

Тип даних	Спотовий ендпоінт	Ф'ючерсний ендпоінт UM	Ф'ючерсний ендпоінт CM
K-line	/api/v3/klines	/fapi/v1/klines	/dapi/v1/klines
Agg Trades	/api/v3/aggTrades	/fapi/v1/aggTrades	/dapi/v1/aggTrades
Trades	/api/v3/historicalTrades	/fapi/v1/trades	/dapi/v1/trades
Index Price K-line	N/A	/fapi/v1/indexPriceKlines	/dapi/v1/indexPriceKlines
Mark Price K-line	N/A	/fapi/v1/markPriceKlines	/dapi/v1/markPriceKlines
Premium Index K-line	N/A	/fapi/v1/premiumIndex	/dapi/v1/premiumIndex

Для скальпінгу досить рідко використовується спотова торгівля у зв'язку з

тим, що вона не призначена для подібних цілей, а також має обмежений функціонал і великі комісії за операції. Враховуючи особливості спотової торгівлі, для скальпінгу вибирається ф'ючерсна торгівля, яка за вищеописаними властивостями перевищує спотову торгівлю [20].

Для торгівлі на ф'ючерсах у реальному часі використовується модуль `um_futures` від `python-binance`. Клієнт та методи даного модуля надають доступ до свічкових історичних даних в обмеженому форматі. Обмеження за максимальною кількістю свічок, що запитуються, відповідає 1500 свічкам. Тимчасового ряду на 1500 позицій більш ніж достатньо для технічного аналізу за допомогою аналізованих статистичних методів [20].

Кількість свічок, що запитуються, і відповідна вага запиту відображені в табл. 2.3.

Таблиця 2.3 – Кількість свічок, що запитуються, і відповідна вага запиту

LIMIT	weight
[1,100)	1
[100, 500)	2
[500, 1000]	5
> 1000	10

Відповіддю від сервера є добові дані зі свічок у форматі, котрий описаний в табл. 2.4.

Для технічного аналізу, в основному, використовується тимчасовий ряд з параметрами Open time і Close [17, 20]. Крім того, для розрахунку деяких статистичних показників необхідно використовувати такі параметри, як High та Low. Дані параметри є екстремумами зміни ціни відкритої свічки, які, як мінімум, застосовуються у формулах розрахунку індикатора CCI.

Виходячи з наведеного вище, немає необхідності отримувати інші параметри. Стандартний набір параметрів для аналізу включає Open time, Open, High, Low і Close.

Таблиця 2.4 – Формат відповіді від сервера за свічковими даними

Параметр свічки	Опис	Приклад
Open time	Час відкриття свічки у форматі часу unix	1676320620000
Open	Ціна відкриття	89.89
High	Максимум ціни	89.94
Low	Мінімум ціни	89.86
Close	Ціна закриття	89.87
Volume	Обсяг торгів	229.570
Close time	Час закриття свічки у форматі часу unix	1676320679999
Quote asset volume	Обсяг активу, що котирується	20637.78595
Number of trades	Кількість угод	8
Taker buy base asset volume	Тейкер купує обсяг базового активу протягом цього періоду	148.119
Taker buy quote asset volume	Обсяг активу котирування тейкера за даний період	13316.37132
Can be ignored	Ігнорується	0

2.7 Зберігання даних

Отримані дані свічок при зборі допускається зберігати в змінних програмах. Для цього в роботі використовується бібліотека Pandas, яка дозволяє легко структурувати дані для аналізу, особливо для часових рядів. Pandas надає особливі структури даних, звані дата-фреймом і серією, до яких можна застосовувати також методи, які надаються Pandas для статистичного аналізу.

На рис. 2.4 відображено приклад зберігання даних часового ряду в

	Open	High	Low	Close
Open time				
2023-02-05 00:00:00	98.48	98.59	98.37	98.42
2023-02-05 00:01:00	98.36	98.36	98.24	98.28
2023-02-05 00:02:00	98.25	98.39	98.25	98.29
2023-02-05 00:03:00	98.30	98.36	98.22	98.24
2023-02-05 00:04:00	98.24	98.28	98.16	98.23
...	...	...	...	...
2023-02-05 23:56:00	96.51	96.55	96.35	96.55
2023-02-05 23:57:00	96.53	96.57	96.49	96.57
2023-02-05 23:58:00	96.55	96.63	96.49	96.57
2023-02-05 23:59:00	96.52	96.68	96.51	96.60
2023-02-06 00:00:00	96.64	96.67	96.56	96.67

Рисунок 2.4 – Зберігання даних тимчасового ряду в датафреймі Pandas

Проте, необхідно зберігати як дані параметрів свічок, а й розрахунки статистичних показників торгів. Статистичні розрахунки застосовуються для подальшого аналізу результатів та покращення алгоритмів передбачення. Оскільки розроблені алгоритми необхідно тестувати на різних криптовалютних активах, то відповідно постає питання про покращення структури зберігання даних [21, 22]. Створювати безліч різних змінних стає не доцільно, оскільки під час розробки відбуваються постійні перезапуски скрипта, у ході чого дані потрібно збирати знову.

Враховуючи сказане вище та масштабованість проекту, у роботі слід розглянути застосування СУБД. Як раніше було сказано, у проекті в комбінації з Django буде використана PostgreSQL, яка є однією з відкритих і найбільш розвинених у світі об'єктно-реляційна СУБД. З її допомогою також легко створюються, модифікуються та видаляються записи в БД.

2.8 Висновок до другого розділу

Було описано вимоги до системи, що розробляється, та завдання інтеграції даних. Наведено її архітектуру. Досліджено стек технологій, який використовується для розробки. Побудована модель даних, наведено її представлення в нотаціях UML.

Здійснено аргументований вибір торгової площадки та криптовалют для аналізу. Наведено опис механізму отримання історичних даних та їх збереження.

3 ПРАКТИЧНА ЧАСТИНА

Для впровадження функціоналу системи прогнозування використана мова Python і на її основі реалізовані окремі модулі. Python, фактично є стандартом для використання у різних контекстах науки про дані (data science). Ця мова містить одночасно і потужність, і простоту застосування. Як раніше було описано, Python володіє бібліотечними файлами для завантаження даних та їх візуалізації, також для проведення статистичних обчислень, і т.д. Такий значний обсяг засобів пропонує спеціалістам по роботі із даними (data scientists) власне потужний набір інструментів із різним призначенням.

У зв'язку з наведеними вище фактами, початкова розробка для зручності проводиться в інтерактивному веб-середовищі Jupyter Lab. У міру масштабування проекту, до розробки застосовуються IDE та багаторівневий підхід.

3.1 Бібліотеки

Безумовно, основною бібліотекою у проекті є python-binance. З її допомогою здійснюється підключення до сервера, отримання історичних даних та управління операціями, рахунком та загалом акаунтом на біржі.

Далі, важливу роль у роботі програми грають бібліотеки Pandas, Numpy, statsmodels і TA-Lib.

Як було зазначено раніше, Pandas у роботі застосовується в основному для зберігання отриманих історичних даних зі свічок торгових пар. У дата-фреймі як індекс позначається параметр «Open Time» (час відкриття свічки). Pandas зручний тим, що дозволяє легко перевести дату з формату тимчасової мітки Binance у формат дати і часу UTC, або місцевого часового поясу.

Numpy застосовується для отримання масивів даних із датафреймів Pandas.

Потім отримані масиви даних використовуються для розрахунку статистичних показників. Наприклад, під час розрахунку індикатора CCI.

Бібліотека TA-Lib є найпопулярнішою для розрахунку технічних індикаторів. Популярність отримана за рахунок того, що вона швидше за альтернативні їй бібліотеки. Швидкість, своєю чергою, пов'язана з тим, що ця бібліотека написана мовою програмування C, що робить її досить ефективною для обчислень.

TA-Lib надає готові формули майже всім індикаторам, крім графічних: рівні Фібоначі, точки розвороту та ін. Але дані індикатори також реалізуються на Python за допомогою статистичних даних по тимчасовому ряду.

Важливо відзначити, що у роботі також використовуються інші бібліотеки. Такі, як matplotlib, time, pytz, calendar, datetime та інші (у міру розширення системи). Дані бібліотеки використовуються для графічної візуалізації, для визначення тимчасових міток у тимчасовому ряді з урахуванням локальних тимчасових зон та для інших операцій.

3.2 Визначення констант

Насамперед, необхідно позначити основні константи, які використовуються в програмі. Такими є торгові пари та таймфрейм для торгівлі.

Для аналізу, у всіх торгових парах другою криптовалютою як котирування у цій роботі приймається USDT. Клієнт приймає торгову пару у подібному форматі «LTCUSDT».

Таймфрейм - це період часу або торговий період, під час якого групуються котирування та вибудовуються елементи цінового графіка (свічки, бари, лінії та ін.). Для торгівлі використовуються різні таймф-рейми [22]. Основні з них:

- M1 (одногохвилинний);

- M5 (п'ятихвилинний);
- M15 (п'ятнадцятихвилинний);
- H1 (годинний);
- D1 (денний);
- W1 (тижневий);
- MN (місцевий) і т.д.

Торгівля скальпінгом відповідає таймфреймам M1-M15, але в окремих випадках розглядається і M30 [23, 24]. Враховуючи дані умови, для алгоритму скальпінгової торгівлі як константу прийнято таймфрейм M1. Клієнт binance приймає таймфрейм у форматі "1 m".

Клієнт binance також запитує ключі API, які можна вказати при підключенні до клієнта, але вони необхідні тільки для фактичної відправки торгових ордерів на біржу binance в режимі реального часу і не потрібні для читання даних. Дані ключі будуть особливо необхідні, коли розробка торкнеться торгівлі на прогнозах в реальному час [25].

3.3 Програмна реалізація алгоритму для передбачення поведінки ціни

3.3.1 Реалізація індикаторів на основі статистики

Оскільки для передбачення напрямку тренду ціни використовуються різні комбінації статистичних методів, то внаслідок цього програмна реалізація алгоритму ґрунтується на поділі кожної комбінації індикаторів на окремі функції. Окремо кожен індикатор не вважається і не фіксується в окрему функцію в силу того, що періоди часового ряду кожної комбінації індикаторів змінюються, як і інші параметри. Наприклад, індикатор CCI має період 6 у комбінації RSI+CCI, але в комбінації MACD+CCI той самий індикатор має період 14.

Зобразимо на рис. 3.1 діаграму для розробленого алгоритму розрахунку та отримання датафрейму значень індикаторів у комбінації MACD+CCI.

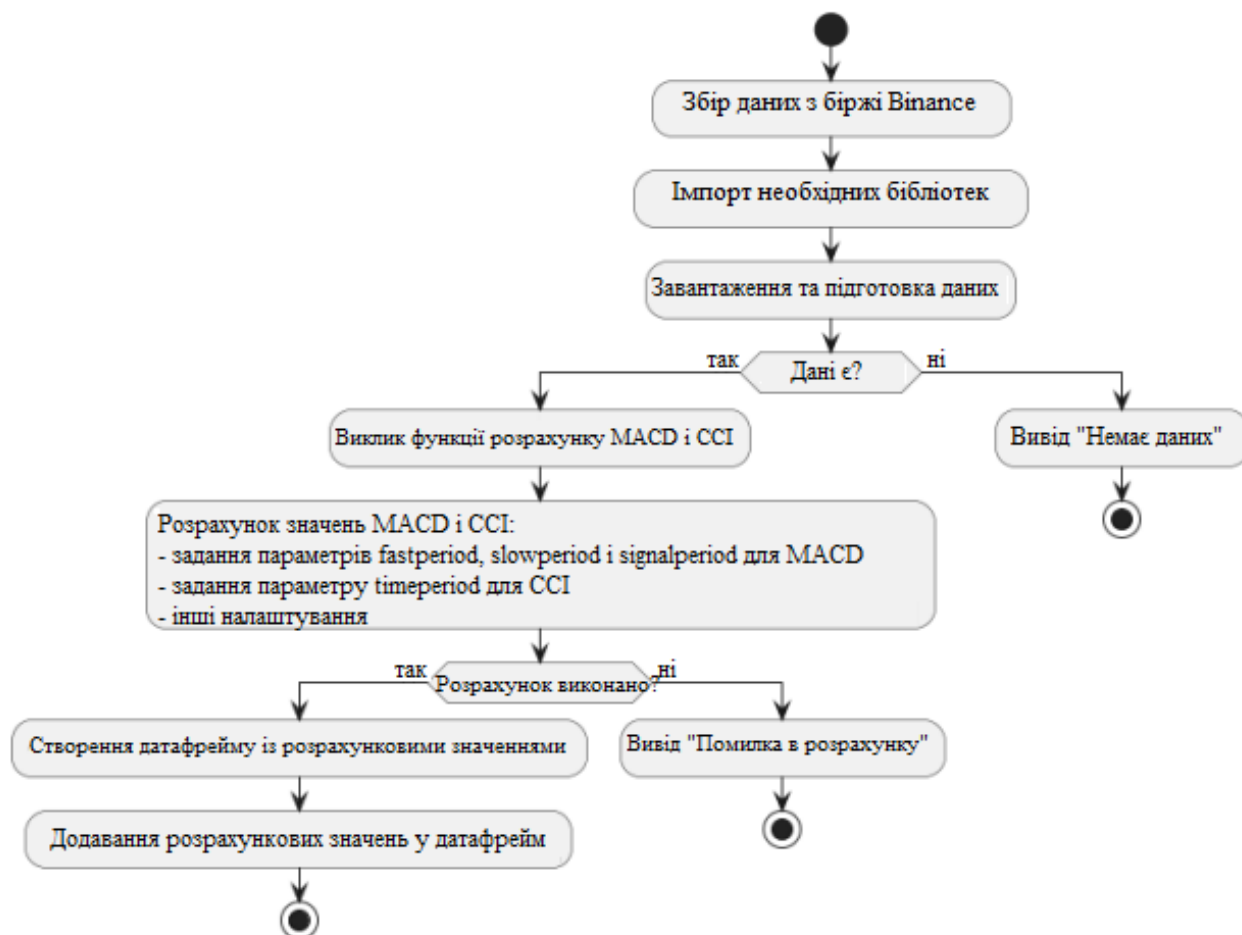


Рисунок 3.1 – Розроблений алгоритм розрахунку та отримання значень індикаторів у комбінації MACD+CCI

Функція розрахунку значень MACD і CCI приймає два аргументи, де перший аргумент відповідає датафрейму з історичними даними за певний період, а другий аргумент відповідає кількості останніх позицій, які необхідно проаналізувати. Функція повертає дата-фрейм з отриманими розрахунками.

В даному випадку використаний метод MACD() для розрахунку сходження-розбіжності ковзних середніх за ціною закриття свічки. Для проведення обчислень застосовуються три експоненційні ковзні середні із відмінними періодами fastperiod, slowperiod і signalperiod [26]. Розрахунки в даному випадку базуються на наступних основних формулах:

- від швидкої ковзної середньої з меншим періодом (EMA_S) віднімається

повільна ковзна середня з великим періодом (EMA_t), потім за отриманими значеннями будується лінія MACD:

$$MACD = EMA_s(P) - EMA_t(P); \quad (3.1)$$

- отримана лінія згладжується третьою експоненційною ковзною середньою (EMA_a), як правило, з періодом 9 і виходить, так звана, сигнальна лінія MACD (Signal):

$$Signal = EMA_a(EMA_s(P) - EMA_t(P)). \quad (3.2)$$

Метод `talib.MACD()`, у свою чергу розраховує та виводить 3 масиви:

- масив значень з різницями експоненційних ковзаючих середніх $EMA_s - EMA_t$, параметри `fastperiod` і `slowperiod` - відповідно, періоди;
- масив значень розрахованої експоненціальної ковзної середньої EMA_a за значеннями різниці $EMA_s - EMA_t$, де `signalperiod` - період EMA_a ;
- масив значень різниці $MACD - Signal$, де значення 0 означає перетин двох кривих.

Далі, за допомогою циклів і нескладних конструкцій умовних операторів, в датафреймі виконується пошук на відповідність умовам для відкриття long - позицій і short -позицій. Дані умови описані раніше в підзаголовку роботи, присвяченому допустимим комбінаціям індикаторів для аналізу. Отримані «сигнали» про рух цінового тренду записуються в раніше отриманий датафрейм і відповідають кожній позиції часового ряду.

На рис. 3.2 зображено повноцінну діаграму діяльності UML для розробленого алгоритму прогнозування із застосуванням комбінацій індикаторів.

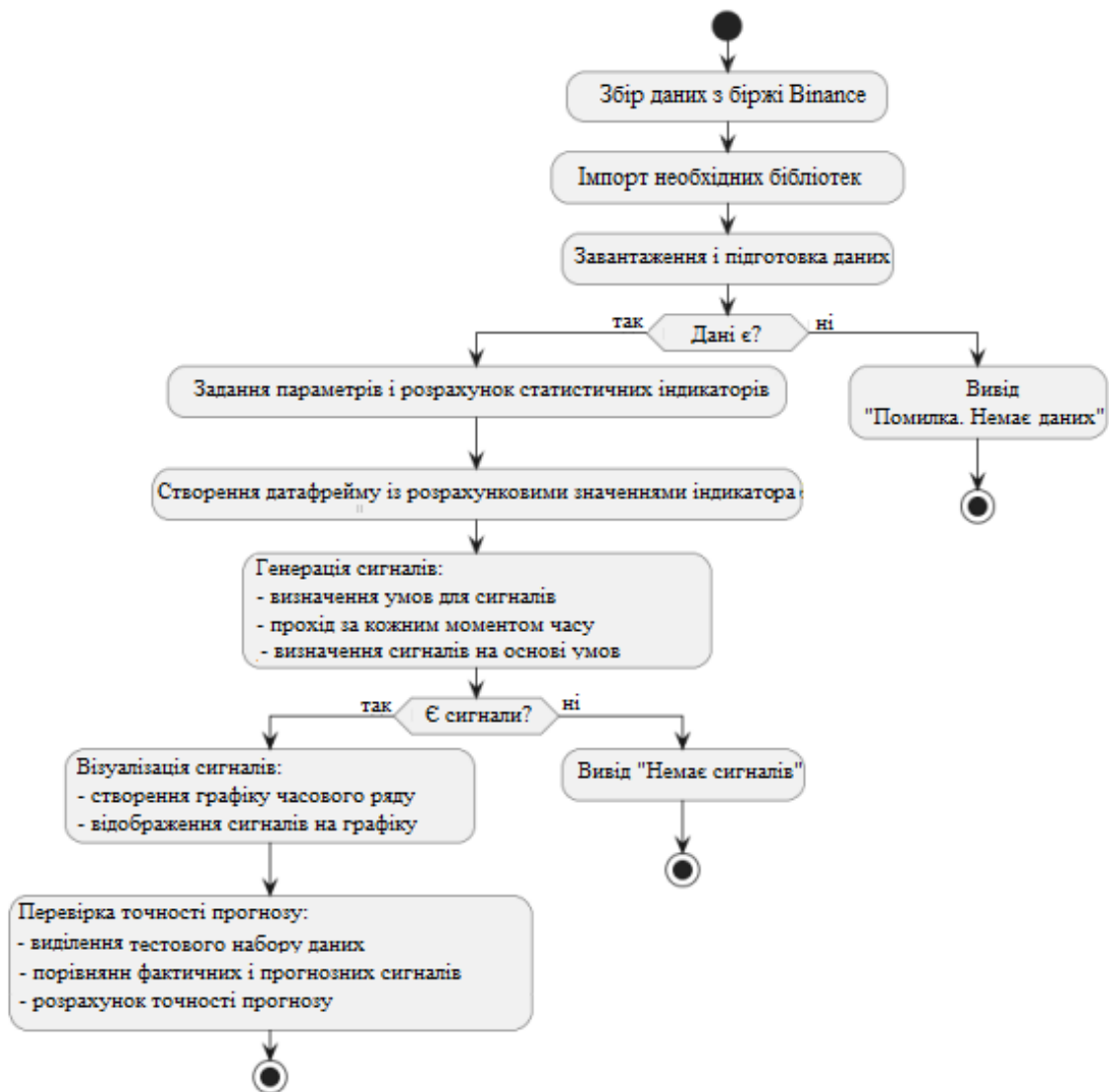


Рисунок 3.2 – Розроблений алгоритм прогнозування із застосуванням комбінацій індикаторів

У середині функцій для створення датафрейму з розрахунковими значеннями індикаторів у роботі застосовується бібліотека TA-Lib.

3.3.2 Реалізація авторегресійної моделі

SARIMA - це модель тимчасових рядів, яка є розширенням моделі ARIMA для врахування сезонності даних. SARIMA модель доступна в бібліотеці statsmodels в Python і надає зручні інструменти для аналізу та прогнозування

часових рядів.

SARIMA модель в statsmodels дозволяє враховувати як несезонні, так і сезонні компоненти часового ряду [27]. Вона визначається за допомогою пари метрів, що включають порядки авторегресії (p), різниці (d) і ковзного середнього (q) для несезонних компонентів, а також порядки авторегресії (P), різниці (D), ковзного середнього (Q) і періоду сезонності (s) для сезонних компонентів. Модель також може враховувати помилки (residuals) та зовнішні (екзогенні) змінні, якщо вони доступні.

У бібліотеці statsmodels SARIMA модель реалізована з використанням класу SARIMAX. Вона надає методи для навчання моделі на даних часового ряду, виконання прогнозування майбутніх значень та оцінки параметрів моделі.

Statsmodels також надає функції для аналізу та діагностики SARIMA моделі, такі як графіки автокореляційної функції (ACF) та часткової автокореляційної функції (PACF), тест Льюнга-Бокса для перевірки наявності автокореляції в залишках та інші.

Загалом бібліотека statsmodels надає потужні інструменти для роботи з SARIMA моделями, які можуть бути застосовані для аналізу, прогнозування та моделювання часових рядів з урахуванням сезонності [28].

Програмна реалізація алгоритму прогнозування ціни криптовалют з використанням моделі SARIMA з пакету statsmodels в Python включає наступні кроки:

- підготовка даних: імпортування необхідних бібліотек, завантаження та підготовка даних тимчасового ряду котирувань криптовалют. Зазвичай дані представлені у форматі тимчасової мітки та відповідного значення ціни;
- створення моделі SARIMA: імпортування класу SARIMAX із модуля statsmodels.tsa.statespace.sarimax. Потім необхідно задати параметри моделі, включаючи порядки (p , d , q) для авторегресійної, різницевої та ковзної складових, а також порядки (P , D , Q , s) для сезонних компонентів. Після цього, створення екземпляра моделі SARIMAX, використовуючи зазначені параметри;

- навчання моделі: використання методу `fit()` для навчання моделі на підготовлених даних часового ряду. Метод `fit()` оцінює параметри моделі з використанням методу максимальної правдоподібності (MLE) та виконує відповідну оптимізацію;
- прогнозування: використання методів `predict()` або `forecast()` для виконання прогнозу майбутніх значень ціни криптовалюти на основі навченої моделі SARIMA. Методи `predict()` і `forecast()` повертають прогностичні значення та відповідні довірчі інтервали.

3.4 Візуалізація та аналіз результатів прогнозування

Для тестування та аналізу результатів роботи алгоритму програмно реалізуються модель SARIMA та три комбінації статистичних індикаторів. Потім, за допомогою бібліотеки `matplotlib` і колекції функцій `pyplot`, здійснюється візуалізація отриманих результатів передбачень щодо руху ціни.

Візуалізація включає тенденцію зміни показників часового ряду і мітки, що означають майбутнє зростання_або падіння ціни. Зелена мітка «BUY» говорить користувачеві про те, що актив необхідно купувати, оскільки передбачається зростання ціни. У деяких випадках зростання може статися не відразу, а тільки після невеликого відхилення вниз по тренду. Червона мітка «SELL», відповідно, говорить користувачеві про те, що актив необхідно продавати. Сітка на графіку має часовий інтервал.

3.4.1 Результат роботи RSI + CCI

Як можна помітити, на тимчасовому ряді є деякі аномалії. Можливо, у цей проміжок часу на ринок заходять більші учасники. Інші аномалії, пов'язані з викидами, були усунені у процесі роботи шляхом виділення лише якісної ділянки часового ряду.

Методом проб і помилок для даної комбінації індикаторів були підібрані відповідні періоди для максимальної точності ймовірності.

На ділянці до 19:00 годин відзначається спадаючий тренд, де алгоритм розставляє правильні мітки. Потім у тренда з'являється міцний опір, де алгоритм помилково виявляє висхідний тренд і ставить відповідну мітку. Але, далі алгоритм безпомилково визначає падіння та зростання ціни активу. Тільки на ділянці після 21:00 оголошується помилкова мітка, що відповідає спадаючому тренду.

Результат передбачення за рішенням RSI+CCI зображено на рис .3.4.

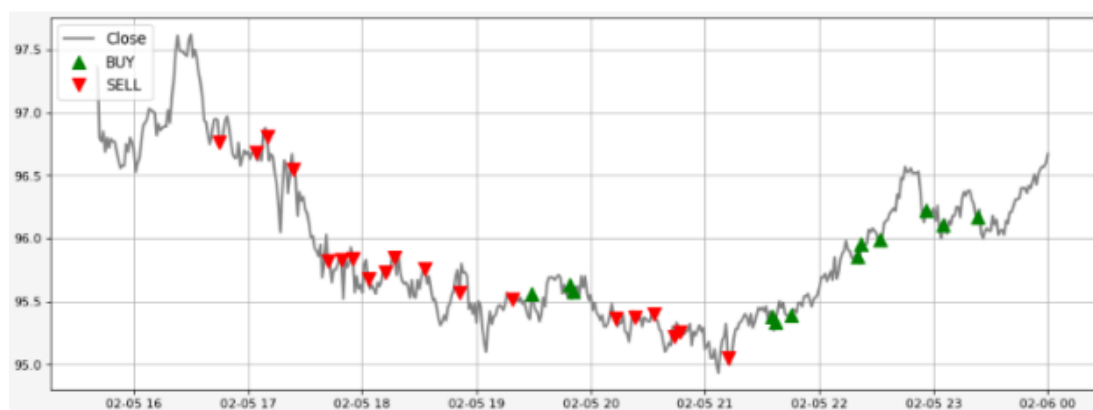


Рисунок 3.4 – Передбачення напрямку руху по RSI+CCI

3.4.2 Результат роботи EMA + CCI

Ця комбінація індикаторів має свої позитивні та негативні сторони у передбаченні руху ціни.

У першу чергу, необхідно відзначити те, що дана стратегія показала більшу кількість позначок. Однак, це також позначилося і на точності передбачення.

Помітно, що на позиції 16:00 алгоритм визначає додаткові зелені мітки, які підтверджуються висхідним трендом. Також слідом відзначається і червона мітка, що відповідає низхідному тренду. Далі, відбувається низка помилкових прогнозів, де після зелених міток суворо відбувається падіння ціни. Після бокового тренду, також є і помилкові передбачення на висхідному тренді.

Можливо, що ця комбінація індикаторів підходить для певних часових проміжків усередині дня, де тренд матиме більше аномалій. Наприклад, коли на ринок виходять «великі» користувачі і здійснюють угоди з великими сумами.

Результат передбачення напрямку руху ЕМА+CCI зображено на рис. 3.5.

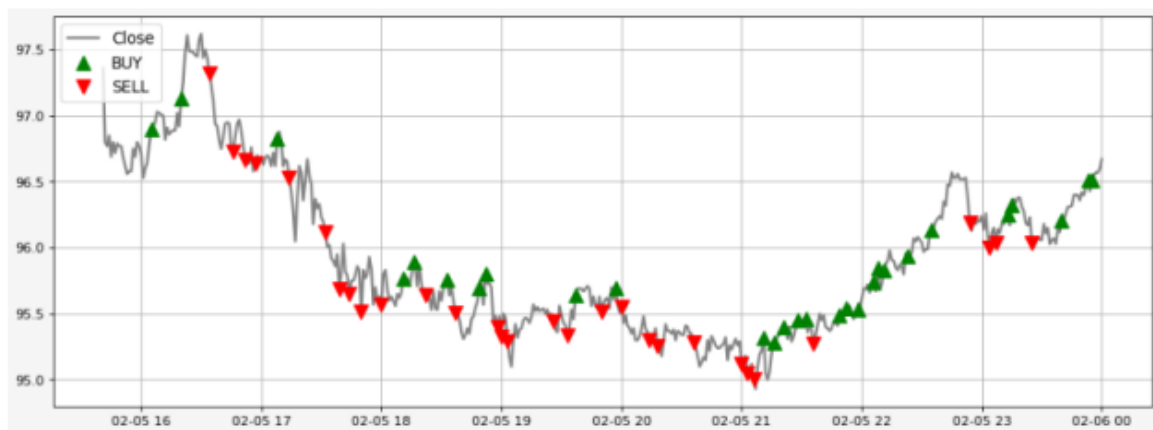


Рисунок 3.5 – Передбачення напрямку руху по ЕМА+CCI

3.4.3 Результат роботи MACD + CCI

Ця комбінація індикаторів має схожий зі стратегією ЕМА + CCI результат передбачень, оскільки розрахунок MACD заснований на експоненціальних ковзаючих середніх. Але, варто зазначити, що на окремих позиціях часового ряду при роботі цієї комбінації виявляються або опускаються деякі мітки.

Помітно, що на позиції близько 16.05 алгоритм не виділив висхідний тренд, який був відзначений за допомогою раніше згаданої стратегії. Також, не було передбачено зниження ціни на 17:15, що свідчить про зниження точності

прогнозування. Разом, важливо відзначити, що дана стратегія або потребує доопрацювання, або некоректно працює на даному таймфреймі часових рядів.

Результат передбачення напрямку руху по MACD+CCI наведено на рис. 3.6.

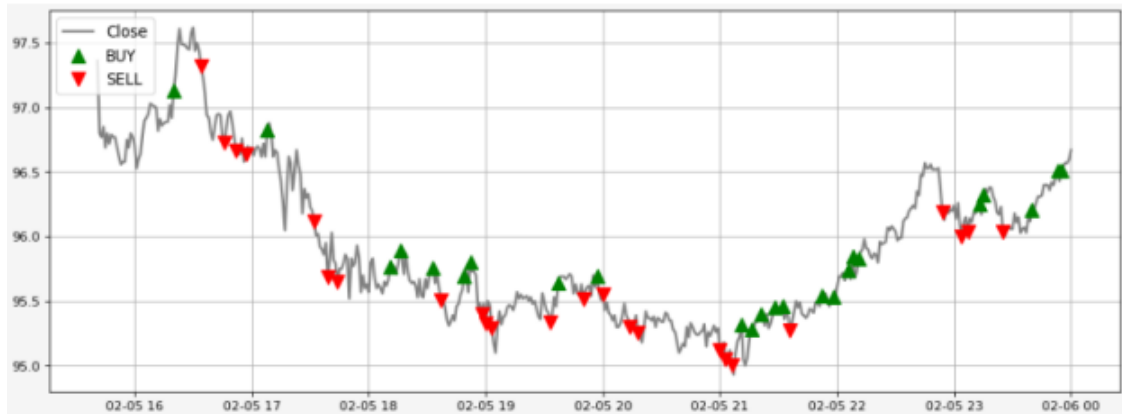


Рисунок 3.5 – Передбачення напрямку руху по MACD + CCI

3.4.4 Результат роботи MACD + Stochastic + 2 EMA

На рис. 3.6 зображено результат роботи розробленого алгоритму.



Рисунок 3.6 – Передбачення напрямку руху по MACD+Stochastic+2EMA

Необхідно відзначити, що кількість сигналів значно знижено у зв'язку з тим, що працюють умови для кількох складових стратегії. Зауважимо також, що комбінація в деяких ситуаціях помилково прогнозує падіння або зростання активу.

Крім того, точність прогнозування залежить від кількості сигналів, які

потрібно вивести користувачеві. Наведемо приклади тестів:

- точність 65% на 1-хвилинному таймфреймі на даних 5-6 лютого (500 позицій);
- точність 62% на 5-хвилинному таймфреймі на даних 5-9 лютого;
- на 15-хвилинному таймфреймі також видано більше 50% правильних сигналів.

3.4.5 Результат роботи Double CCI + RSIOMA

Ця комбінація індикаторів працює вже краще, ніж попередня і в деяких випадках досягає точності 80%. Стабільно точність не знижується нижче 50%, що дозволяє користувачеві не втратити свої заощадження.

Результат передбачення для Double CCI + RSIOMA зображений на рис. 3.7.



Рисунок 3.7 – Передбачення напрямку руху по Double CCI + RSIOMA

Також можна помітити, що алгоритм погано відпрацьовує на аномальних ділянках, де ціна сильно різко падає або зростає. У ці моменти здійснюють угоди великі учасники ринку і часто такі явища не вдається уникнути.

3.4.6 Результат роботи BB + 3 EMA + MACD + RSI

Результат роботи цієї стратегії з індикаторами показав найкращу точність серед описаних комбінацій. Але разом із точністю, звичайно, знижується кількість сигналів до операцій. Точність даної стратегії в певних ситуаціях досягає 85%, що є добрим показником. Постійна точність стратегії також не опускається нижче 50%.

Результат редагування напрямку руху за цим рішенням показано на рис. 3.8.



Рисунок 3.8 – Передбачення напрямку руху по BB + 3 EMA + MACD + RSI

3.4.7 Результат роботи авторегресійної моделі SARIMA

Оскільки стратегії на основі індикаторів не дозволяють отримати достатньо високу точність, було прийнято рішення розробити алгоритм прогнозування на основі моделі машинного навчання. Попередній аналіз та обробка даних дозволила привести ряд до стаціонарного вигляду. Необхідно відзначити, що при приведенні часового ряду до стаціонарного вигляду використовувався тест Діккі-Фуллера, який після доопрацювань показав задовільне негативне значення. А також, на доказ стаціонарності ряду додається Р -значення, що вийшло менше, ніж 0,05. Крім того, підбиралася модель з найменшим значенням інформаційного критерію Акаїке, що є найкращим показником. Ця авторегресійна модель на момент розробки показала найкращі результати порівняно з попередніми невдалими моделями. Також даний

результат до певної міри перевищує результат передбачення по індикаторах.

Виділимо спочатку невдалу модель, за якою можна відзначити меншу точність прогнозування (рис. 3.9). А потім виділимо вдалий варіант моделі, де параметри налаштування моделі відіграли важливу роль (рис. 3.10). Результат роботи моделі відобразимо на графіку. Тимчасовий ряд має інтервал 5 хвилин, сітка графіка розбита за інтервалами 10 хвилин.

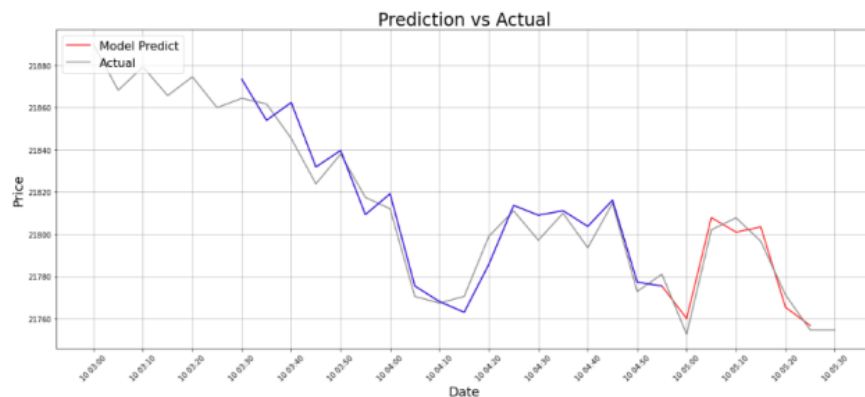


Рисунок 3.9 – Невдала модель SARIMA з тренувальним даними до 04:55 10 грудня

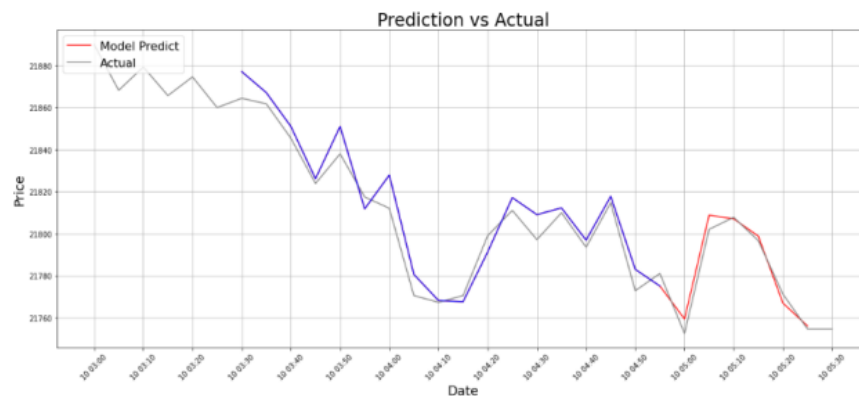


Рисунок 3.9 – Вдала модель SARIMA з тренувальним даними до 04:55 10 грудня

Як можна помітити з рис. 3.9, модель досить добре передбачає зростання і падіння ціни. Помітна похибка помічається на 04:00 і після 05:55, оскільки на практиці даним моделям важко передбачити більше 2-х часових інтервалів.

Тут модель SARIMA помилилася в 2 випадках з 19, що означає, що вона правильно передбачила 17 з 19 спостережень. Таким чином, модель SARIMA,

заснована на досліджуваних даних, має приблизно 89% точності в прогнозуванні часового ряду. Але варто помітити найголовніше в цьому те, що модель адекватно прогнозує майбутнє зростання та падіння ціни як мінімум на 2 позиції.

3.5 Порівняння найкращих рішень

Аналіз та порівняння результатів роботи різних рішень зазначені у табл. 3.1.

Таблиця 3.1 – Аналіз та порівняння результатів роботи різних рішень

Стратегія	Точність прогнозування	Швидкість прогнозування	Складність реалізації	Ресурсоємність	Коментар
MACD + Stochastic + 2 EMA	<75%	10 секунд	Середня	Низька	Швидші передбачення, але знижена точність
Double CCI + RSIOMA	<75%	10 секунд	Середня	Низька	Швидші передбачення, але знижена точність
BB + 3 EMA MACD + RS	<80%	10 секунд	Висока	Низька	Швидші передбачення, але знижена точність
SARIMA	>85%	2-3 хвилини	Висока	Висока	Вища точність, але повільні передбачення (початковий показ близько 15 хвилин). Однак, модель була оптимізована і тепер швидкості в 2-3 хвилини вистачає для того, щоб встигати приймати рішення

З таблиці видно, що стратегії, засновані на статистичних індикаторах, демонструють швидші передбачення порівняно з моделлю SARIMA. Проте,

точність прогнозування даних стратегій становить трохи більше 80%.

З іншого боку, модель SARIMA показує більш високу точність прогнозування (понад 85%), але вимагає значно більше часу на передбачення (2-3 хвилини).

Важливо, що вибір оптимальної стратегії залежить від конкретних вимог та обмежень. Якщо швидкість прийняття рішень є пріоритетом і деяка втрата точності допустима, то використання стратегій на основі статистичних індикаторів може бути кращим. З іншого боку, якщо точність прогнозування є головним фактором, модель SARIMA може бути кращою, незважаючи на більш тривалий час передбачення.

3.6 Висновок до третього розділу

У цьому розділі були описані Python-бібліотеки (Pandas, Numpy, statsmodels і TA-Lib), які використовуються для розробки. Визначені константи, які застосовуватимуться для прогнозування - це торгові пари та таймфрейм.

Наведена програмна реалізація алгоритму для передбачення поведінки ціни криптовалюти із застосуванням індикаторів на основі статистики та авторегресійної моделі. Із використанням засобів бібліотеки matplotlib і колекції функцій pyplot, виконано візуалізацію одержаних результатів прогнозу щодо руху ціни для різних комбінацій торгових індикаторів.

Виконано порівняння найкращих рішень, з'ясовано, що модель SARIMA найкраще підходить для прогнозування, навіть незважаючи на триваліші часові рамки її роботи.

4 ОХОРОНА ПРАЦІ ТА БЕЗПЕКА ЖИТТЄДІЯЛЬНОСТІ

4.1 Охорона праці

Метою кваліфікаційної роботи магістра є розробка інформаційно-аналітичної системи для прогнозування котирувань в галузі торгівлі криптовалютними активами. Оскільки, проведення робіт з розробки та використання такої системи передбачає застосування комп'ютерної техніки, зокрема ПК та периферійних пристроїв, то обов'язковим є дотримання вимог з охорони праці і техніки безпеки.

Для ефективної і безпечної роботи колективу працівників з розробки ПЗ комп'ютерних систем, в тому числі і фахівців з підвищення ефективності контролю доступу в приміщення, необхідно організувати безпечні умови праці. При цьому керівник організації несе безпосередню відповідальність за порушення нормативно-правових актів з охорони праці [35]. Окрім цього, на робочих місцях працівників необхідно забезпечити дотримання вимог, затверджених Наказом Мінсоцполітики від 14.02.2018 за № 207 «Про затвердження Вимог щодо безпеки та захисту здоров'я працівників під час роботи з екранними пристроями». Згідно Вимог приміщення, де розміщені робочі місця операторів, крім приміщень, у яких розміщені робочі місця операторів великих ЕОМ загального призначення (сервер), мають бути оснащені системою автоматичної пожежної сигналізації відповідно до цих вимог;

– переліку однотипних за призначенням об'єктів, які підлягають обладнанню автоматичними установками пожежогасіння та пожежної сигналізації, затвердженого наказом Міністерства України з питань надзвичайних ситуацій та у справах захисту населення від наслідків Чорнобильської катастрофи від 22.08.2005 N 161, зареєстрованого в Міністерстві юстиції України 05.09.2005 за N 990/11270 (НАПБ Б.06.004-2005);

– Державних будівельних норм "Інженерне обладнання будинків і споруд. Пожежна автоматика будинків і споруд", затверджених наказом Держбуду України

від 28.10.98 N 247 (далі - ДБН В.2.5-56:2014, з димовими пожежними сповіщувачами та переносними вуглекислотними вогнегасниками.

В інших приміщеннях допускається встановлювати теплові пожежні сповіщувачі. Приміщення, де розміщені робочі місця операторів, мають бути оснащені вогнегасниками, кількість яких визначається згідно з вимогами ДСТУ 4297:2004 «Пожежна техніка. Технічне обслуговування вогнегасників». Загальні технічні вимоги і з урахуванням граничнодопустимих концентрацій вогнегасної рідини відповідно до вимог НАПБ А.01.001-2014. Приміщення, в яких розміщуються робочі місця операторів сервера загального призначення, обладнуються системою автоматичної пожежної сигналізації та засобами пожежогасіння відповідно до вимог ДБН В.2.5-56:2014, НАПБ А.01.001-2014 і вимог нормативно-технічної та експлуатаційної документації виробника. Проходи до засобів пожежогасіння мають бути вільними.

Лінія електромережі для живлення комп'ютера та периферійних пристроїв повинні бути виконаними як окрема групова трипровідна мережа шляхом прокладання фазового, нульового робочого та нульового захисного провідників. Нульовий захисний провідник використовується для заземлення (занулення) електроприймачів. Не допускається використовувати нульовий робочий провідник як нульовий захисний провідник. Нульовий захисний провідник прокладається від стійки групового розподільного щита, розподільного пункту до розеток електроживлення. Не допускається підключати на щиті до одного контактного затискача нульовий робочий та нульовий захисний провідники.

Площа перерізу нульового робочого та нульового захисного провідника в груповій трипровідній мережі має бути не менше площі перерізу фазового провідника. Усі провідники мають відповідати номінальним параметрам мережі та навантаження, умовам навколишнього середовища, умовам розподілу провідників, температурному режиму та типам апаратури захисту, вимогам НПАОП 40.1-1.01-97.

У приміщенні, де одночасно експлуатуються понад п'ять комп'ютерів, на помітному, доступному місці встановлюється аварійний резервний вимикач, який

може повністю вимкнути електричне живлення приміщення, крім освітлення. Комп'ютери повинні підключатися до електромережі тільки за допомогою справних штепсельних з'єднань і електророзеток заводського виготовлення.

У штепсельних з'єднаннях та електророзетках, крім контактів фазового та нульового робочого провідників, мають бути спеціальні контакти для підключення нульового захисного провідника. Їхня конструкція має бути такою, щоб приєднання нульового захисного провідника відбувалося раніше, ніж приєднання фазового та нульового робочого провідників. Порядок роз'єднання при відключенні має бути зворотним. Не допускається підключати комп'ютери до звичайної двопровідної електромережі, в тому числі – з використанням перехідних пристроїв. Електромережі штепсельних з'єднань та електророзеток для живлення комп'ютерної техніки повинні бути виконаними за магістральною схемою, по 3-6 з'єднань або електророзеток в одному колі. Штепсельні з'єднання та електророзетки для напруги 12 В та 42 В за своєю конструкцією мають відрізнятися від штепсельних з'єднань для напруги 127 В та 220 В. Штепсельні з'єднання та електророзетки, розраховані на напругу 12 В та 42 В, мають візуально (за кольором) відрізнятися від кольору штепсельних з'єднань, розрахованих на напругу 127 В та 220 В.

При підвищенні ефективності контролю доступу в приміщення, де для забезпечення безпеки мешканців, співробітників і збереження майна використовуються ДС, важливим, з точки зору охорони праці, є забезпечення достатньої величини природного та штучного освітлення, які визначені у НПАОП 0.00-7.15-18. Організація робочого місця фахівця із дослідження методів та програмно-апаратних засобів оптимізаційних процесів на основі ГА повинна забезпечувати відповідність усіх елементів робочого місця та їх розташування ергономічним вимогам ДСТУ 8604:2015 «Дизайн і ергономіка. Робоче місце для виконання робіт у положенні сидячи. Загальні ергономічні вимоги». Відстань від екрана до ока фахівців, які працюють за комп'ютером визначається згідно з вимогами ДСанПіН 3.3.2.007-98.

Розміщення принтера або іншого пристрою введення-виведення інформації

на робочому місці має забезпечувати добру видимість екрана комп'ютера, зручність ручного керування пристроєм введення-виведення інформації в зоні досяжності моторного поля згідно з вимогами ДСанПіН 3.3.2.007-98.

Таким чином, у результаті аналізу вимог щодо охорони праці користувачів комп'ютерів, визначено особливості організації робочих місць, вимог з електробезпеки, природного та штучного освітлення для ефективної і безпечної роботи фахівців з розробки інформаційно-аналітичної системи для прогнозування котирувань в галузі торгівлі криптовалютами активами.

4.2. Функціонування державної системи спостереження, збирання, оброблення та аналізу інформації про стан довкілля під час надзвичайних ситуацій мирного та воєнного часу

Моніторинг довкілля – це система спостереження, збирання та аналізу інформації про ситуацію, що може скластись під час надзвичайних ситуацій мирного та воєнного часу [36]. Також це система спостереження за визначеними об'єктами, явищами та процесами з метою оперативного оцінювання їх стану, виявлення результатів впливу на них зовнішніх чинників та прийняття відповідних управлінських рішень (ДСТУ 3891:2013) (див. ДСТУ 7295:2013).

Моніторинг потенційно небезпечних об'єктів це спостереження, контролювання за зміною параметрів технологічних режимів з метою збирання, збереження, передавання та аналізування інформації щодо поточного стану потенційно небезпечних об'єктів, наявності та кількості порушень вимог безпеки, відпрацювання рекомендацій щодо проведення 98 робіт із запобігання та ліквідування техногенних надзвичайних ситуацій та їх наслідків (ДСТУ 7295:2013).

Моніторинг джерел надзвичайних ситуацій це система спостереження за об'єктами, які можуть бути джерелами надзвичайних ситуацій, що має на меті

виявлення небезпеки, збирання, узагальнення та аналізування оперативної інформації стосовно стану об'єктів моніторингу та розроблення науково-обґрунтованих рекомендацій щодо проведення заходів із запобігання та ліквідування надзвичайних ситуацій (ДСТУ 7295:2013).

Моніторинг довкілля – це систематичні спостереження і контролювання, які проводять регулярно, за єдиною програмою для оцінювання стану довкілля, аналізування процесів, які відбуваються в ньому і своєчасне виявлення тенденцій його змінювання (ДСТУ 7295:2013).

Моніторинг надзвичайних ситуацій (НС) – система спостереження за об'єктами, які можуть бути джерелами надзвичайних ситуацій, що має на меті виявлення небезпеки, збирання, узагальнення та аналізування оперативної інформації щодо об'єктів моніторингу та розроблення науково обґрунтованих рекомендацій щодо проведення заходів із запобігання та ліквідування НС.

Моніторинг небезпечних явищ та процесів це система спостереження та контролювання за розвитком небезпечних та стихійних природних явищ і процесів, чинниками, які спричиняють їх формування та розвиток, аналізування, збереження та передавання інформації щодо виявлення тенденцій їх змінювання, розроблення комплексу заходів щодо запобігання природним надзвичайним ситуаціям та ліквідування їх наслідків. Небезпечні природні явища і процеси підрозділяють на геофізичні, геологічні, гідрологічні, метеорологічні, медико-біологічні та пожежі в природних екосистемах (ДСТУ 7295:2013).

Моніторинг пожеж в екосистемах це спостереження, контролювання, збирання, аналізування, збереження та передавання інформації щодо 99 пожежної небезпеки в природних екосистемах (умов погоди, стану горючих матеріалів, інших пожежонебезпечних чинників), з метою своєчасного планування та здійснення заходів щодо запобігання виникненню і ліквідування пожеж та їх наслідків (ДСТУ 7295:2013).

Моніторинг радіаційної безпеки це спостереження і контролювання рівня радіоактивного забруднення місцевості, повітря, води, продовольства, об'єктів господарювання, дозових навантажень на населення з метою прийняття

оперативних рішень щодо запобігання виникненню та ліквідування надзвичайних ситуацій та їх наслідків (ДСТУ 7295:2013).

Моніторинг хімічної небезпеки це спостереження, контролювання, збирання, аналізування, збереження та передавання інформації щодо визначення ступеня і характеру хімічного забруднення довкілля, санітарногігієнічний нагляд за дотриманням установлених нормативів з метою виявлення джерела надходження небезпечних хімічних речовин, запобігання виникненню та ліквідування надзвичайних ситуацій та їх наслідків (ДСТУ 7295:2013).

Збір та аналіз інформації про стан довкілля під час мирного та воєнного стану дає можливість приймати оперативні рішення для адекватного реагування на ситуацію [36].

4.3. Висновки до розділу

В цьому розділі проаналізовано важливі питання охорони праці та безпеки в надзвичайних ситуаціях, висвітлено питання функціонування державної системи спостереження, збирання, оброблення та аналізу інформації про стан довкілля під час надзвичайних ситуацій мирного та воєнного час.

ВИСНОВКИ

Розроблені рішення успішно протестовані, система дозволяє проводити торгові операції у реальному часі за прогнозами цінових рухів часового ряду.

Дослідження різних інструментів, програмних бібліотек, методів і моделей для роботи з тимчасовими рядами є досить трудомістким завданням. За результатами проведених досліджень у даній науково- дослідній роботі можна зробити висновок, що алгоритмічний трейдинг дозволяє значно спростити процес обробки та прогнозування часового ряду, а також дозволяє зосереджуватися на розумінні проблеми досліджуваного часового ряду.

Візуалізація та аналіз результатів передбачень вказують на те, що розроблені та програмно-реалізовані алгоритми мають як ряд переваг, так ряд і недоліків. Зважаючи на те, що рішення на основі індикаторів мають межі точності, було вирішено застосувати модель машинного навчання SARIMA. ARIMA не була застосована, так як тимчасовий ряд котирувань криптовалют має сезонну складову.

Порівняльний аналіз показав, що рішення на основі статистичних індикаторів та на основі моделі машинного мають свої переваги та недоліки. Таким чином, модель машинного навчання більш точно і впевнено прогнозує майбутню ціну та сам напрямок руху тренду, що має важливу роль у торгівлі. У свою чергу рішення на основі індикаторів мають велику швидкість, в той час як модель тренується близько 2-3 хвилин. Однак максимальна точність рішень на основі індикаторів досягає 75-80 %, тут же модель машинного навчання SARIMA має точність понад 85%.

Крім того, варто зауважити, що результати роботи розрахункових алгоритмів вказують на прийнятну для торгівлі точність комбінації індикаторів, де комбінація BB+3EMA+MACD+RSI показала найкращі результати. Точність даної реалізації в тестах досягала 80%.

ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ

1. Hyndman R. J., Athanasopoulos G. Forecasting: Principles and Practice (3rd ed). Australia: OTexts, 2021. 442 p
2. Бідюк П. І., Романенко В. Д., Тимошук О. Л. Аналіз часових рядів. Київ: НТУУ «КПІ», 2010. 223 с..
3. Armstrong J.S. Illusions in Regression Analysis [Текст] / J.S. Armstrong. Pennsylvania: Penn Press, 2011. 147 p.
4. Adhikari R. An Introductory Study on Time Series Modeling and Forecasting [Текст] / Adhikari R. Riga: LAP Lambert Academic Publishing, 2013. 76 p.
5. Найкращі індикатори для криптотрейдингу та аналізу в 2024 році . URL: <https://plisio.net/uk/blog/best-indicators-for-crypto-trading-and-analysis-in-2024> (дата звертання 20.11.2024)
6. Як ефективно використовувати торгові індикатори? URL: <https://www.binance.com/uk-UA/blog/futuresяк-ефективно-використовувати-торгові-індикатори-7213961428302983386> (дата звертання 21.11.2024)
7. PRICE ACTION: ЩО ЦЕ TAKE? URL: <https://cryptology.school/blog/price-action-shho-ce-take> (дата звертання 20.11.2024)
8. Best Forex Indicators All Traders Should Know in 2024. URL: <https://www.xs.com/en/blog/best-forex-indicators> (дата звертання 21.11.2024)
9. 13 Найкращих Індикаторів для Криптотрейдингу: Посібник з Денних і Свінгових Торгових Індикаторів. URL: <https://goodcrypto.app/uk/cryptocurrency-analysis-the-best-indicators/> (дата звертання 21.11.2024)
10. Найкращі стратегії для торгівлі на волатильності у 2024 році. URL: <https://tradersunion.com/ua/interesting-articles/volatility-trading/> (дата звертання 22.11.2024)
11. Лукасевич М.А. Модель даних для інформаційно-аналітичної системи для прогнозування котирувань криптовалютних активів // Інформаційні моделі, системи та технології: Праці XII наук.-техн. конф. Тернопіль, 2024. С. 204.

12. Binance API Documentation. URL:<https://binance-docs.github.io/apidocs/> (дата звертання 23.11.2024)
13. Welcome to python-binance. URL: <https://pythonbinance.readthedocs.io/en/latest/index.html> (дата звертання 23.11.2024)
14. Поширені запитання про API <https://www.binance.com/uk-UA/support/faq/поширені-запитання-про-апі-360004492232> (дата звертання 24.11.2024)
15. Hansen J. D. Virtual currencies: international actions and regulations. Perkins Coie. URL: <https://www.perkinscoie.com/en/news-insights/virtualcurrenciesinternational-actions-and-regulations.html> (дата звернення: 25.11.2024).
16. Investors Underground: Introduction to Technical Analysis. URL: <https://www.investorsunderground.com/technical-analysis/> (дата звертання 27.11.24).
17. Bakker Joris. Technical Analysis in the Cryptocurrency Market / Joris Bakker. Erasmus School of Economics, 2017. 53 с.
18. Ord K., Fildes R. Kourentzes N. Principles of Business Forecasting (2nd ed). England: wessex, inc., 2017. 588 p.
19. Litecoin. URL: <https://litecoin.org/> (дата звертання 27.11.2024)
20. Bakker Joris. Technical Analysis in the Cryptocurrency Market / Joris Bakker. Erasmus School of Economics, 2017. 53 p.
21. Gerlein, E. A., McGinnity, M., Belatreche, A., & Coleman, S. (2016). Evaluating machine learning classification for financial trading: An empirical approach. Expert Systems with Applications, 54, 193–207.
22. Quantitative cryptocurrency trading: exploring the use of machine learning techniques. вебсайт. URL: <https://dl.acm.org/doi/10.1145/3336499.3338003> (дата звертання 10.12.24)
23. Patel, J., Shah, S., Thakkar, P., & Kotecha, K. (2015). Predicting stock and stock price index movement using Trend Deterministic Data Preparation and machine learning techniques. Expert Systems with Applications, 42(1), 259–268.
24. Sarkis Agaian. Financial Sentiment Analysis Using Machine Learning

Techniques / Agaian Sarkis, Kolm Petter. New York: International Journal of Investment Management and Financial Innovations, 2017. 9 p .

25. Chatzis, S.P., Siakoulis, V., Petropoulos, A., Stavroulakis, E., Vlachogiannakis, N. Forecasting stock market crisis events using deep and statistical machine learning techniques. *Expert Systems with Appl.* 112, pp. 353– 371 (2018)

26. Machine learning applications in time series hierarchical forecastin. вебсайт. URL: <https://arxiv.org/pdf/1912.00370.pdf>? (дата звертання 12.12.24)

27. SARIMA models for time series forecasting. URL: <https://people.duke.edu/~rnau/411arim.html> (дата звертання 12.12.24).

28. Cryptocurrency Predictions with SARIMA. URL: <https://www.kaggle.com/taniaj/cryptocurrency-predictions-with-sarima> (дата звернення: 16.12.2023).

29. Stefanyshyn, V. , Stefanyshyn, I. , Pastukh, O. , Yatsyshyn, V. , Yakymenko. Accuracy of software and hardware of computer systems for human-machine interaction, I. *CEUR Workshop Proceedings*, 2024, 3842, pp. 178–183.

30. Stefanyshyn, I. , Pastukh, O., Stefanyshyn, V. , Baran, I. , Boyko, I. Robustness of AI algorithms for neurocomputer interfaces based on software and hardware technologies *CEUR Workshop Proceedings*, 2024, 3742, pp. 137–149.

31. Petryk, M. , Bachynskyi, M. , Brevus, V. , Mudryk, I. , Mykhalyk, D. Analysis technology of neurological movements considering cognitive feedback influences of cerebral cortex signals *CEUR Workshop Proceedings*, 2022, 3309, pp. 45–54.

32. Boyko, I. , Petryk, M. , Mudryk, I. , Stoianov, Y., Tsupryk, H. Mathematical Model of the Capacitor Based on Zeolite Material. *Proceedings - International Conference on Advanced Computer Information Technologies, ACIT*, 2021, pp. 45–48.

33. Boyko, I. , Petryk, M. , Mykhailyshyn, R . Excitons in resonant tunnelling structures based on AlN/GaN/AlN/AlGaIn/AlN nitride: spectral dependences and intensities of interband optical transitions. *Ukrainian Journal of Physical Optics*, 2022, 23(3), pp. 180–191.

34. M. R. Petryk, A. Khimich, M. M. Petryk, and J. Fraissard, “Experimental and computer simulation studies of dehydration on microporous adsorbent of natural gas used

as motor fuel,” Fuel, Vol. 239, 1324–1330 (2019).
<https://doi.org/https://doi.org/10.1016/j.fuel.2018.10.134>

35. Заїкіна Д., Глива В. Основи охорони праці та безпека життєдіяльності. 2019. URL: <https://doi.org/10.31435/rsglobal/001> (дата звернення: 14.12.2024).

36. Безпека в надзвичайних ситуаціях. Методичний посібник для здобувачів освітнього ступеня «магістр» всіх спеціальностей денної та заочної (дистанційної) форм навчання / укл.: Стручок В. С. Тернопіль: ФОП Паляниця В. А., 2022. 156 с.

ДОДАТКИ

ДОДАТОК А

Тези конференції

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
ТЕРНОПІЛЬСЬКИЙ НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ
імені ІВАНА ПУЛЮЯ
НАУКОВЕ ТОВАРИСТВО ім. ШЕВЧЕНКА

МАТЕРІАЛИ

ХІІ НАУКОВО-ТЕХНІЧНОЇ КОНФЕРЕНЦІЇ

**«ІНФОРМАЦІЙНІ МОДЕЛІ,
СИСТЕМИ ТА ТЕХНОЛОГІЇ»**



18-19 грудня 2024 року

ТЕРНОПІЛЬ
2024

К. Кулішова, С. Дячук РОЗРОБКА РІШЕНЬ НА ОСНОВІ ШТУЧНОГО ІНТЕЛЕКТУ ДЛЯ АВТОМАТИЗАЦІЇ КОНТЕНТУ В СУЧАСНИХ CRM-СИСТЕМАХ K. Kulishova, S. Dyachuk DEVELOPMENT OF ARTIFICIAL INTELLIGENCE-BASED SOLUTIONS FOR CONTENT AUTOMATION IN MODERN CRM SYSTEMS	171
А. О. Круглюк ВІЗУАЛІЗАЦІЯ ЯК ВИБІР ДИЗАЙНУ A. O. Kruhlyk VISUALIZATION AS A DESIGN CHOICE	172
Т. Лавренчук МЕТОДИ РЕФАКТОРИНГУ В ГНУЧКІЙ РОЗРОБЦІ T. Lavrenchuk REFACTORING METHODS IN AGILE DEVELOPMENT	173
Т. Лавренчук ВПЛИВ TEST-DRIVEN DEVELOPMENT НА ЯКІСТЬ РОЗРОБКИ T. Lavrenchuk INFLUENCE OF TEST-DRIVEN DEVELOPMENT ON DEVELOPMENT QUALITY	174
А. С. Луговий, І. В. Тих МЕТОДИ ОПТИМІЗАЦІЇ ПРОДУКТИВНОСТІ ETL-СИСТЕМ У БАГАТОРІВНЕВИХ АНАЛІТИЧНИХ ПЛАТФОРМАХ A. S. Lugovyi, I. V. Tykh METHODS FOR OPTIMIZING ETL SYSTEM PERFORMANCE IN MULTI-TIER ANALYTICAL PLATFORMS	175
М. А. Лукасевич МОДЕЛЬ ДАНИХ ДЛЯ ІНФОРМАЦІЙНО-АНАЛІТИЧНОЇ СИСТЕМИ ДЛЯ ПРОГНОЗУВАННЯ КОТИРУВАНЬ КРИПТОВАЛЮТНИХ АКТИВІВ M. Lukasevych DATA MODEL FOR INFORMATION AND ANALYTIC SYSTEM FOR FORECASTING QUOTATIONS OF CRYPTOCURRENCY ASSETS	176
Н. Мельник, А. М. Лутіків РОЗГОРТАННЯ ВЛАСНИХ АРТЕФАКТІВ ЕКОСИСТЕМИ HADOOP N. Melnyk, A. Lutskiv DEPLOYMENT OF OWN ARTIFACTS OF THE HADOOP ECOSYSTEM	177
Б. Б. Млунко, І. М. Клушко ЗАСТОСУВАННЯ МЕТОДІВ ШТУЧНОГО ІНТЕЛЕКТУ ДЛЯ ОБРОБКИ ПРИРОДНОЇ МОВИ ЗАСОБАМИ ГЛИБИННОГО НАВЧАННЯ B. B. Mlynko, I. M. Klyushko APPLICATION OF ARTIFICIAL INTELLIGENCE METHODS FOR NATURAL LANGUAGE PROCESSING USING DEEP LEARNING TECHNIQUES	179
Є. В. Огіньський, Д. С. Антошук ВИКОРИСТАННЯ ЙМОВІРНІСНОЇ ПРИРОДИ ВЕЛИКИХ МОВНИХ МОДЕЛЕЙ В СФЕРІ ПЕРСОНАЛЬНИХ ФІНАНСІВ Y. V. Ohinskiy, D. S. Antosniuk UTILIZING THE PROBABILISTIC NATURE OF LARGE LANGUAGE MODELS IN THE FIELD OF PERSONAL FINANCE	180
П. С. Панчухін, М. І. Паламар ПІДВИЩЕННЯ ЯКОСТІ РОЗПИЗНАВАННЯ ОБ'ЄКТІВ У ВІДЕОПОТОЦІ В РЕАЛЬНОМУ ЧАСІ P. Panchukhin, M. I. Palamar IMPROVING THE QUALITY OF OBJECT RECOGNITION IN A REAL-TIME VIDEO STREAM	181

МОДЕЛЬ ДАНИХ ДЛЯ ІНФОРМАЦІЙНО-АНАЛІТИЧНОЇ СИСТЕМИ ДЛЯ ПРОГНОЗУВАННЯ КОТИРУВАНЬ КРИПТОВАЛЮТНИХ АКТИВІВ

M. Lukasevich

DATA MODEL FOR INFORMATION AND ANALYTIC SYSTEM FOR FORECASTING QUOTATIONS OF CRYPTOCURRENCY ASSETS

Модель даних (МД) є комплексною структурою, яка охоплює різні аспекти системи. Вона включає моделі, такі як Account, AccountPortfolio, Order, Signal та Prediction, які взаємодіють один з одним для забезпечення функціональності системи.

МД містить інформацію про трейдерів та їх акаунти (Account), їх портфелі (AccountPortfolio) з балансами і прибутком/збитком, ордери на купівлю та продаж (Order), сигнали (Signal), які рекомендують певні дії з торгівлі, та передбачення (Prediction), із інформацією про прогнозовані значення і пов'язані з певними сигналами.

Кожен користувач (Account) має свій акаунтний портфель (AccountPortfolio), в якому зберігається інформація про баланс та прибуток. Користувачі можуть створювати замовлення на купівлю або продаж крипто валютних активів. Кожне замовлення містить інформацію про символ активу, кількість та напрямок операції.

Діаграма класів UML для МД у системі зображено на рис. 1, де:

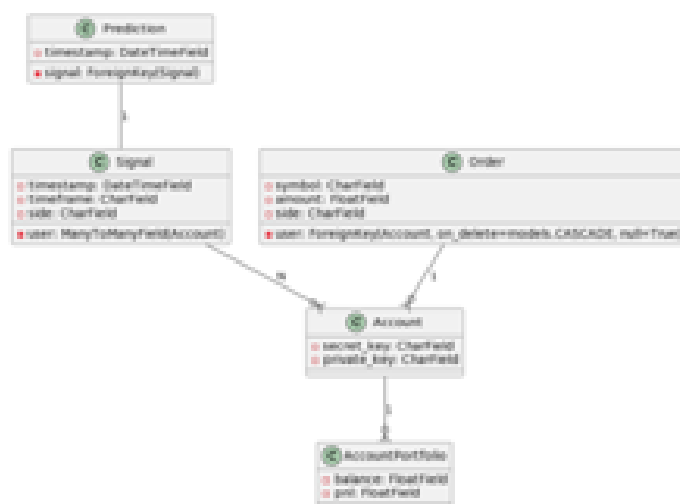


Рисунок 1 – Діаграма класів UML для моделі даних

– відношення між класами Account та AccountPortfolio: 1 вказує на зв'язок "один до одного" (One-to-One) між Account та AccountPortfolio. Один обліковий запис може бути пов'язаний лише з одним портфелем облікового запису;

– відношення між класами Signal та Account: N вказує на зв'язок "багато до багатьох" (Many-to-Many) між Signal і Account. Один сигнал може бути пов'язаний з декількома обліковими записами, і один обліковий

запис може бути пов'язаний з кількома сигналами;

– відношення між класами Order та Account: 1 вказує на зв'язок "один до багатьох" (One-to-Many) між Order та Account. Одне замовлення може бути пов'язане лише з одним обліковим записом, але один обліковий запис може мати кілька замовлень.

– відношення між класами Prediction та Signal: 1 вказує на зв'язок "один до одного" (One-to-One) між Prediction та Signal. Один прогноз може бути пов'язаний лише з одним сигналом.

Отже, МД підтримує зберігання інформації про користувачів, портфелі, сигнали, прогнози та замовлення, що дозволяє системі ефективно прогнозувати і управляти торговими операціями з криптовалютами.

ДОДАТОК Б

Фрагмент вихідного коду модуля прогнозування

```
# Necessary Packages
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt import statsmodels.api as
sm
import statsmodels.formula.api as smf
from statsmodels.tsa.stattools
import adfuller
from sklearn.metrics
import mean_squared_error from pandas import read_csv
from pandas import datetime
from pandas import DataFrame
from statsmodels.tsa.arima_model import SARIMA from
matplotlib
import pyplot
# Bitcoin data
data1 = pd.read_csv('bitcoin.csv', header=0)
# visualize the data frame, first 10 elements
print(data1.head(4150))
#graphical representation data1.plot()
plt.show()
#read data again
series = read_csv('bitcoin.csv', header=3370, index_col=0)

from statsmodels.graphics.tsaplots import acf, pacf
from statsmodels.graphics.tsaplots import plot_acf,
plot_pacf pyplot.figure()
pyplot.subplot(211)
plot_acf(series, ax=pyplot.gca(),lags=30)
pyplot.subplot(212)
plot_pacf(series, ax=pyplot.gca(),lags=30) pyplot.show()
# returns or first difference of the time series
returns=series.diff()
returns.plot() plt.show()
from statsmodels.tsa.arima_model import ARIMA import
pmdarima as pm
model = pm.auto_arima(series, start_p=1,
start_q=1, test='adf', max_p=6, max_q=6, m=1,
d=None, seasonal=False, start_P=0, D=0,
```

```

trace=True,                                error_action='ignore',
suppress_warnings=True,                     stepwise=True)
print(model.summary())
# fit an SARIMA model for the series model = SARIMA
(series, order=(1,1,1)) model_fit = model.fit(dis=0)
print(model_fit.summary())
# plot residual errors
residuals = DataFrame(model_fit.resid) fig, ax =
plt.subplots(1,2) residuals.plot(title="Residuals",
ax=ax[0])
residuals.plot(kind='kde', title='Density', ax=ax[1])
plt.show() print(residuals.describe())
print(model_fit.aic)
# Actual vs Fitted
def forecast_accuracy(forecast, actual):
mape = np.mean(np.abs(forecast - actual)/np.abs(actual))
# MAPE me = np.mean(forecast - actual) # ME
mae = np.mean(np.abs(forecast - actual)) # MAE mpe =
np.mean((forecast - actual)/actual) # MPE rmse =
np.mean((forecast - actual)**2)**.5 # RMSE corr =
np.corrcoef(forecast, actual)[0,1] # corr
mins =
np.amin(np.hstack([forecast[:,None],actual[:,None]]),
axis=1) maxs =
np.amax(np.hstack([forecast[:,None],actual[:,None]]),
axis=1) minmax = 1 - np.mean(mins/maxs) # minmax
# acf1 = acf (fc-test) [1] # ACF1
return({'mape':mape, 'me':me, 'mae': mae, 'mpe':
mpe, 'rmse':rmse, 'corr':corr, 'minmax':minmax})
yy=model_fit.predict(652) xx=series.iloc[652:781,0]
tt=forecast_accuracy(yy, xx) print('forecast
accuracy',tt) pyplot.figure() plt.scatter(xx,yy,
marker='o') pyplot.show() pyplot.figure()
model_fit.plot_predict(652,781)
pyplot.show()
# separate data in training and test data (95% i 5%)
forecast3, stderr, conf_int = model_fit.forecast(10)
print(forecast3) pyplot.plot(forecast3,color='green')
pyplot.show() pyplot.figure()
model_fit.plot_predict(752,791) pyplot.show()
from math import sqrt
from sklearn.metrics import mean_squared_error rms =
sqrt(mean_squared_error(yy, xx)) print('RMSE', rms)
import numpy as np import pandas as pd
import matplotlib.pyplot as plt import statsmodels.api as
sm

```

```

import statsmodels.formula.api as smf
from statsmodels.tsa.stattools import adfuller
from sklearn.metrics import mean_squared_error from pandas
import read_csv
from pandas import datetime from pandas import DataFrame
from statsmodels.tsa.arima_model import ARIMA from
matplotlib import pyplot
from statsmodels.tsa.arima_model import SARIMA
from statsmodels.tsa.arima_model import ARMA
from statsmodels.tsa.seasonal import seasonal_decompose #
Bitcoin data
data = pd.read_csv('bitcoin.csv') #divide into train and
validation set train = data[:int(0.835*(len(data)))] valid
= data[int(0.835*(len(data))):]
#preprocessing (since arima e univariate series as input)
train.drop('Date',axis=1,inplace=True)
valid.drop('Date',axis=1,inplace=True)
#plotting the data train['bprice'].plot()
valid['bprice'].plot() #building the model
from pmdarima import auto_arima
model = auto_arima (train, suppress_warnings = True)
model.fit(train)
trace=True, error_action='ignore',
forecast = model.predict(n_periods=len(valid))
forecast = pd.DataFrame(forecast,index =
valid.index,columns=['Prediction']) #plot the predictions
for validation set
pyplot.figure() plt.figure(figsize=(10, 6)) plt.plot
(train, label = 'Train') plt.plot(valid,
label='Valid')plt.plot(forecast, label='Prediction')
plt.show()
#calculate rmse
from math import sqrt
from sklearn.metrics import mean_squared_error from
sklearn.metrics import mean_absolute_error rms1 =
sqrt(mean_squared_error(valid,forecast))
print('RMSE',rms1)

```

....

Додаток В

Диск з роботою

