

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
Західноукраїнський національний університет
Факультет комп'ютерних інформаційних технологій
Кафедра інформаційно-обчислювальних систем і управління

КУЛИНИЧ Арсен Юрійович

**Модуль прогнозування ціни біткойна на основі методу
ARIMA / Bitcoin price forecasting module based on the ARIMA
method**

Спеціальність 122 – Комп'ютерні науки
Освітньо-професійна програма – Комп'ютерні науки

Дипломний проект

Виконав студент групи КН-41
А. Ю. Кулинич

Науковий керівник:
викладач В.І. Дорош

Дипломний проект допущено до захисту
«__»_____ 2023 р.

Завідувач кафедри
_____ М.П. Комар

Тернопіль – 2023

Факультет комп'ютерних інформаційних технологій
Кафедра інформаційно-обчислювальних систем і управління
Освітній ступінь «бакалавр»
Спеціальність 122 – Комп'ютерні науки
Освітньо-професійна програма – Комп'ютерні науки

ЗАТВЕРДЖУЮ
Завідувач кафедри
М.П. Комар
« ____ » _____ 2022р.

З А В Д А Н Н Я
НА ДИПЛОМНИЙ ПРОЕКТ СТУДЕНТУ
Кулиничу Арсену Юрійовичу
(прізвище, ім'я, по батькові)

1. Тема проекту: Модуль прогнозування ціни біткойна на основі методу ARIMA / Bitcoin price forecasting module based on the ARIMA method
керівник проекту викладач В.І. Дорош

затверджені наказом по університету від 08 грудня 2022 р. № 491.

2. Строк подання студентом закінченого проекту 01 червня 2023 р.

3. Вихідні дані до проекту: технічне завдання.

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

- Провести аналіз поняття біткойну, його особливостей та роль у сучасній економіці.
- Оглянути та проаналізувати існуючі рішення та методи прогнозування ціни біткойна.
- Визначити оптимальний різновид ARIMA для прогнозування ціни біткойна, обґрунтувати вибір та розробити алгоритм його застосування.
- Розробити алгоритм визначення стаціонарності та декомпозиції часового ряду ціни біткойна.
- Реалізувати алгоритм прогнозування ціни біткойна на основі методу SARIMA з урахуванням сезонності.
- Здійснити аналіз набору даних ціни біткойна, включаючи статистичний аналіз, виявлення трендів та сезонності.

- Побудувати модель прогнозування на основі SARIMA з використанням вибраних параметрів та перевірити її точність та надійність.
- Зробити прогноз ціни біткойна на майбутні періоди, використовуючи розроблену модель та провести валідацію результатів.

5. Перелік графічного матеріалу (з точним зазначенням обов’язкових креслень)

- алгоритм визначення стаціонарності та декомпозиції часового ряду ціни біткойна

6. Консультанти розділів проекту

Розділ	Прізвище, ініціали та посада консультанта	Підпис, дата	
		Завдання видав	Завдання прийняв
Н. контроль	викладач В.І. Дорош		

7. Дата видачі завдання 08 грудня 2022 р.

КАЛЕНДАРНИЙ ПЛАН

№ з/п	Назва етапів дипломного проекту	Строк виконання етапів проекту	Примітка
1	Аналіз предметної області і постановка задачі дослідження	30.12.2022	
2	Алгоритмічне забезпечення	24.03.2023	
3	Програмно-технологічне забезпечення	12.05.2023	
4	Повне завершення та оформлення дипломного проекту	01.06.2023	

Студент _____ А. Ю. Кулинич
(підпис)

Керівник проекту _____ В.І. Дорош
(підпис)

РЕФЕРАТ

Пояснювальна записка до дипломного проекту: 64 с., 15 рис., 1 табл., 2 додатки, 36 джерел.

Метою дипломної роботи є розробка та реалізація модуля прогнозування ціни біткойна на основі методу ARIMA.

Об'єктом дослідження є ціна біткойна та її поведінка в часі.

Предметом дослідження є метод ARIMA і його різновиди для прогнозування ціни біткойна.

Розроблено та досліджено програмне забезпечення для прогнозування ціни біткойна на основі методу ARIMA.

Запропонований модуль може бути корисним для інвесторів, трейдерів та дослідників ринку криптовалют. Прогнозування ціни біткойна може допомогти в прийнятті рішень щодо купівлі, продажу або тримання цієї криптовалюти. Дослідження також може допомогти в розумінні динаміки ціни біткойна та факторів, що впливають на неї, що може бути корисним для подальшого розвитку ринку криптовалют та фінансових стратегій.

БІТКОЙН, ПРОГНОЗУВАННЯ, ЧАСОВИЙ РЯД, ARIMA, SARIMA.

ABSTRACT

The bachelor's thesis report: 64 pages, 15 figures, 1 table, 2 appendices, 36 references.

The aim of the diploma project is to develop and implement a module for Bitcoin price prediction based on the ARIMA method.

The object of research is the Bitcoin price and its behavior over time.

The subject of research is the ARIMA method and its variations for Bitcoin price prediction.

Software for Bitcoin price forecasting based on the ARIMA method has been developed and studied.

The proposed module can be useful for investors, traders, and cryptocurrency market researchers. Predicting the Bitcoin price can assist in making decisions regarding buying, selling, or holding this cryptocurrency. The research can also contribute to understanding the dynamics of Bitcoin price and the factors influencing it, which can be valuable for the further development of the cryptocurrency market and financial strategies.

BITCOIN, FORECASTING, TIME SERIES, ARIMA, SARIMA.

ТЕХНІЧНЕ ЗАВДАННЯ

1. НАЙМЕНУВАННЯ ТА ОБЛАСТЬ ЗАСТОСУВАННЯ

1.1 Модуль прогнозування ціни біткойна на основі методу ARIMA.

1.2 Область застосування – фінансовий ринок та криптовалютні інвестиції.

2. ОСНОВА ДЛЯ РОЗРОБЛЕННЯ

Основою для розроблення є завдання на дипломний проект, затверджене кафедрою інформаційно-обчислювальних систем і управління факультету комп'ютерних інформаційних технологій Західноукраїнського національного університету.

3. ПРИЗНАЧЕННЯ РОЗРОБЛЕНОГО КОМПЛЕКСУ

Метою дипломної роботи є розробка та реалізація модуля прогнозування ціни біткойна на основі методу ARIMA.

4. ДЖЕРЕЛА РОЗРОБЛЕННЯ

Джерелами даної розробки є матеріали навчальної і реферативної літератури, технічна документація, науково-дослідні статті, журнали, Інтернет.

5. ТЕХНІЧНІ ВИМОГИ

5.1 Основні функціональні вимоги до програмної системи:

- вибір даних (пакетних даних / даних у реальному часі) з різних архітектур великих даних;

- провести попередню обробку даних для інтелектуального аналізу тексту.

- провести моделювання прогнозування на основі SARIMA.

5.2 Вимоги до апаратних засобів:

– кластер запускається на фізичній машині Macbook Pro з 16 гігабайт оперативної пам'яті та процесором 2.3 GHz Intel Core i5.

5.3 Вимоги до програмних засобів:

- для розробки програмне забезпечення - Python 3.7;
- для створення графічного інтерфейсу користувача використано – tkinter, Adobe XD;
- для реалізації моделей навчання – фреймворк sklearn.

6. ПОРЯДОК КОНТРОЛЮ

6.1 Представлення дипломного проекту на попередній захист.

6.2 Представлення дипломного проекту на захист.

Завдання прийняв до виконання _____ А. Ю. Кулинич
(підпис) (прізвище та ініціали)

Керівник дипломного проекту _____ В.І. Дорош
(підпис) (прізвище та ініціали)

ЗМІСТ

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СКОРОЧЕНЬ І ТЕРМІНІВ	9
Вступ.....	11
1 Аналіз предметної області і постановка задачі дослідження	14
1.1 Поняття біткойну	14
1.2 Аналіз відомих рішень	18
1.3 Вибір оптимального різновиду ARIMA для прогнозування ціни біткойна	23
1.4 Постановка задачі дослідження.....	28
2 Алгоритмічне забезпечення	31
2.1 Алгоритм визначення стаціонарності та декомпозиція часового ряду	31
2.2 Алгоритм прогнозування методом SARIMA	35
2.3 Алгоритм прогнозування ціни біткойна на основі методу SARIMA	38
3 Програмно-технологічне забезпечення.....	41
3.1 Аналіз набору даних	41
3.2 Моделювання на значеннях ціни біткойна.....	50
3.3 Прогнозування ціни біткойна	54
Висновки	58
Список використаних джерел.....	61
Додаток А.....	65
Додаток Б	65

					ДП.КН.9675613.064.ПЗ			
Змн.	Арк.	№ докум.	Підпис	Дат				
Розроб.		Кулинич А. Ю.			Модуль прогнозування ціни біткойна на основі методу ARIMA		Літ.	Аркуш
Перевір.		Дорош В.І.						
Консульт.							8	64
Н. Контр.		Дорош В.І.					ЗУНУ.ФКІТ.КН-42	
Затверд.		Комар М.П.						

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СКОРОЧЕНЬ І ТЕРМІНІВ

- BTC: Bitcoin - перша і найбільш відома криптовалюта, що працює на основі технології блокчейн.
- DLT: Distributed Ledger Technology - технологія розподіленого реєстру, яка використовується для запису і передачі даних через мережу розподілених вузлів.
- ARIMA: Autoregressive Integrated Moving Average - модель аналізу і прогнозування часових рядів, що поєднує авторегресію (AR), інтегрування (I) та ковзне середнє (MA) компоненти.
- GARCH: Generalized Autoregressive Conditional Heteroskedasticity - модель, яка використовується для моделювання та прогнозування волатильності в фінансових часових рядах.
- BNN: Bayesian Neural Network - байєсівська нейронна мережа, модель, що поєднує нейронні мережі з байєсівським підходом для моделювання та прогнозування.
- MAPE: Mean Absolute Percentage Error - середня абсолютна відсоткова похибка, метрика, що використовується для вимірювання відхилень між прогнозованими значеннями і спостережуваними значеннями в задачах прогнозування.
- ШНМ: Штучна нейронна мережа - модель, що імітує роботу нервової системи для обробки і аналізу даних.
- RNN: Recurrent Neural Network - рекурентна нейронна мережа, модель, що здатна зберігати пам'ять про попередні обчислення та працювати з послідовними даними.
- SVM: Support Vector Machines - метод машинного навчання, який використовується для класифікації і регресії, шляхом знаходження границі розділення між класами з найбільшою можливою шириною.
- LSTM: Long Short-Term Memory - тип рекурентних нейронних мереж, що має здатність довготривало запам'ятовувати інформацію та

					ДП.КН. 9675613.064.ПЗ	Арк.
						9
Зм.	Арк.	№ докум.	Підпис	Дата		

уникати проблем втрати градієнту, що часто виникають у звичайних RNN.

- AR: Авторегресія - статистична модель, в якій поточне значення змінної залежить від попередніх значень тієї ж змінної.
- I: Інтегрування - компонент моделі ARIMA, який відповідає за включення різниці між значеннями часового ряду для стабілізації його і надання йому стаціонарності.
- MA: Ковзне середнє - компонент моделі ARIMA, який враховує попередні значення помилок прогнозу для врахування короткострокових змін в часовому ряді.
- SARIMA: Seasonal Autoregressive Integrated Moving Average - модель ARIMA, що додатково враховує сезонність в часовому ряді.
- ARIMAX: Autoregressive Integrated Moving Average with Exogenous Variables - розширена модель ARIMA, яка дозволяє включати вхідні змінні у модель для прогнозування.
- SARIMAX: Seasonal Autoregressive Integrated Moving Average with Exogenous Variables - розширена модель SARIMA, яка дозволяє включати вхідні змінні та сезонність у модель для прогнозування.
- ACF: Autocorrelation Function - автокореляційна функція, яка вимірює статистичну залежність між значеннями часового ряду і його лагами.
- FFT: Fast Fourier Transform - швидке перетворення Фур'є, алгоритм для обчислення перетворення Фур'є, що використовується для аналізу часових сигналів.
- MSE: Mean Squared Error - середня квадратична помилка, метрика, яка використовується для вимірювання відхилень між прогнозованими значеннями і спостережуваними значеннями в задачах прогнозування.
- R^2 : Коефіцієнт детермінації - статистична метрика, яка вимірює відповідність моделі до вихідних даних.

					ДП.КН. 9675613.064.ПЗ	Арк.
						10
Зм.	Арк.	№ докум.	Підпис	Дата		

ВСТУП

Актуальність теми. Модуль прогнозування ціни біткойна на основі методу ARIMA є дуже актуальним завданням у сучасному світі криптовалют. Біткойн, як одна з найпопулярніших та найбільш відомих криптовалют, має велике значення для інвесторів, трейдерів та всього фінансового ринку.

Прогнозування ціни біткойна дозволяє зробити передбачення щодо майбутньої вартості цієї криптовалюти, що має велике значення для прийняття рішень щодо інвестицій та торгівлі. Здатність точно прогнозувати ціну біткойна може допомогти інвесторам отримати вигоду від коливань на ринку та знизити ризики.

Метод ARIMA (Autoregressive Integrated Moving Average) є одним з найпоширеніших та ефективних методів прогнозування часових рядів. Використання ARIMA для прогнозування ціни біткойна дозволяє аналізувати історичні дані, виявляти тренди, сезонні зміни та інші фактори, що впливають на ціну, та робити передбачення на основі цих аналізів.

Оскільки ціна біткойна є вкрай змінною та піддається впливу багатьох факторів, використання методу ARIMA дозволяє отримати об'єктивний та науково обґрунтований прогноз ціни біткойна, що є важливим інструментом для інвесторів та трейдерів.

Отже, розробка модуля прогнозування ціни біткойна на основі методу ARIMA є актуальною задачею, яка може мати велике практичне значення для фінансового ринку та інвестиційного середовища.

Метою дипломної роботи є розробка та реалізація модуля прогнозування ціни біткойна на основі методу ARIMA. Головною метою є створення надійного та точного інструменту, який зможе аналізувати історичні дані ціни біткойна, виявляти тренди та сезонні зміни, а також робити передбачення майбутньої ціни.

Для досягнення поставленої мети дипломної роботи було необхідно провести теоретичні дослідження, визначити основні аспекти та послідовно вирішити наступні завдання:

					ДП.КН. 9675613.064.ПЗ	Арк.
						11
Зм.	Арк.	№ докум.	Підпис	Дата		

1. Провести аналіз поняття біткойну, його особливостей та роль у сучасній економіці.
2. Оглянути та проаналізувати існуючі рішення та методи прогнозування ціни біткойна.
3. Визначити оптимальний різновид ARIMA для прогнозування ціни біткойна, обґрунтувати вибір та розробити алгоритм його застосування.
4. Розробити алгоритм визначення стаціонарності та декомпозиції часового ряду ціни біткойна.
5. Реалізувати алгоритм прогнозування ціни біткойна на основі методу SARIMA з урахуванням сезонності.
6. Здійснити аналіз набору даних ціни біткойна, включаючи статистичний аналіз, виявлення трендів та сезонності.
7. Побудувати модель прогнозування на основі SARIMA з використанням вибраних параметрів та перевірити її точність та надійність.
8. Зробити прогноз ціни біткойна на майбутні періоди, використовуючи розроблену модель та провести валідацію результатів.

Об'єктом дослідження є ціна біткойна та її поведінка в часі.

Предметом дослідження є метод ARIMA і його різновиди для прогнозування ціни біткойна.

Методи дослідження включають аналіз літератури, збір та аналіз статистичних даних про ціну біткойна, реалізацію алгоритмів ARIMA та його різновидів, використання статистичних тестів та метрик для оцінки точності прогнозів, а також порівняння з іншими методами прогнозування.

Практичне значення дослідження полягає в розробці та застосуванні модуля прогнозування ціни біткойна на основі методу ARIMA. Цей модуль може бути корисним для інвесторів, трейдерів та дослідників ринку криптовалют. Прогнозування ціни біткойна може допомогти в прийнятті рішень щодо купівлі, продажу або тримання цієї криптовалюти. Дослідження також може допомогти в розумінні динаміки ціни біткойна та факторів, що

					ДП.КН. 9675613.064.ПЗ	Арк.
Зм.	Арк.	№ докум.	Підпис	Дата		12

впливають на неї, що може бути корисним для подальшого розвитку ринку криптовалют та фінансових стратегій.

Структура та обсяг роботи. Дипломний проект складається з вступу, трьох розділів, висновків і списку використаних джерел. Загальний обсяг роботи становить 64 сторінок комп'ютерного тексту, в якому включено 15 рисунків і 1 таблиця. Список використаних джерел містить 36 найменувань і розташований на 5 сторінках.

					ДП.КН. 9675613.064.ПЗ	Арк.
						13
Зм.	Арк.	№ докум.	Підпис	Дата		

1 АНАЛІЗ ПРЕДМЕТНОЇ ОБЛАСТІ І ПОСТАНОВКА ЗАДАЧІ ДОСЛІДЖЕННЯ

1.1 Поняття біткойну

У 2008 році програміст під псевдонімом на ім'я Сатоші Накамото опублікував 9-сторінковий документ, в якому описується нова децентралізована цифрова валюта. Вони назвали це Bitcoin.

Біткойн, який був введений на ринок в 2009 році Сатоші Накамото, є першою успішною децентралізованою криптовалютою та платіжною системою. Він відноситься до категорії цифрових активів, в яких транзакції захищені та перевірені за допомогою криптографії - науки про кодування та декодування даних. Зазвичай ці транзакції зберігаються на комп'ютерах, розподілених по всьому світу за допомогою технології блокчейн, що є розподіленою книгою записів.

Біткойн може бути поділений на менші одиниці, відомі як "сатоші", які складаються з восьми знаків після коми. Він може використовуватися як засіб платежу, але також вважається засобом збереження вартості, подібно до золота. Це пояснюється тим, що вартість одного біткойна значно зросла з моменту його запуску - від дрібної суми до десятків тисяч доларів. Коли мова йде про біткойн як про ринковий актив, його позначають тикером BTC.

Термін "децентралізований" широко використовується в контексті криптовалюти і означає, що вона не має централізованого розташування або контролю. В разі біткойна та багатьох інших криптовалют, їх технологія та інфраструктура, які відповідають за створення, постачання та безпеку, не залежать від централізованих організацій, таких як банки чи уряди, для їх управління.

Біткойн був розроблений таким чином, щоб користувачі могли обмінюватися цінностями напряму один з одним через однорангову мережу. Це означає, що всі користувачі мають однакову потужність і підключені

					ДП.КН. 9675613.064.ПЗ	Арк.
Зм.	Арк.	№ докум.	Підпис	Дата		14

безпосередньо один до одного без присутності центрального сервера або посередника. Така архітектура дозволяє здійснювати обмін та зберігання даних, а також відправляти та отримувати платежі в біткойнах між сторонами безперешкодно.

Мережа біткойн (з великої літери "В") є повністю публічною і доступною для будь-якого користувача з підключенням до Інтернету і відповідним пристроєм. Вона базується на відкритому вихідному коді, що означає, що будь-хто може переглянути або спільно використовувати код, на якому ґрунтується біткойн.

Біткойн можна сприймати як "Інтернет грошей". Як Інтернет, він цифровий, не належить жодній окремій особі або організації, доступний для всіх (з умовою наявності підключення до електронної мережі), працює неперервно 24/7, і дозволяє людям обмінюватися даними один з одним. Але тепер уявіть, що існує "інтернет-валюта", де будь-хто, хто використовує Інтернет, може сприяти її захисту, емісії та безпосередньо обмінюватися нею без посередництва банків. Саме це, в основному, є біткойн.

Важливо розуміти, що біткойн складається з трьох окремих компонентів, усі з яких об'єднуються, щоб створити децентралізовану платіжну систему:

- Мережа Bitcoin
- Рідна криптовалюта мережі Bitcoin, яка називається біткойн (BTC)
- Блокчейн Bitcoin

Біткойн працює в одноранговій мережі, де користувачам — зазвичай фізичним або юридичним особам, які хочуть обміняти біткоіни з іншими в мережі — не потрібна допомога посередників для виконання та перевірки транзакцій. Користувачі можуть підключити свій комп'ютер безпосередньо до цієї мережі та завантажити її загальнодоступну книгу, у якій записуються всі історичні транзакції біткойнів.

Ця публічна книга використовує технологію, відому як «блокчейн», також відому як «технологія розподіленої книги». Технологія блокчейн — це

					ДП.КН. 9675613.064.ПЗ	Арк.
Зм.	Арк.	№ докум.	Підпис	Дата		15

те, що дозволяє перевіряти, зберігати та впорядковувати транзакції криптовалюти незмінним і прозорим способом. Незмінність і прозорість є життєво важливими обліковими даними для платіжної системи, яка покладається на нульову довіру.

Щоразу, коли нові транзакції підтверджуються та додаються до книги, мережа оновлює копію книги кожного користувача, щоб відобразити останні зміни. Подумайте про це як про відкритий документ Google, який автоматично оновлюється, коли хтось із доступом редагує його вміст.

Як впливає з назви, біткойн-блокчейн — це цифровий рядок упорядкованих у хронологічному порядку «блоків» — фрагментів коду, які містять дані транзакцій біткойнів. Однак важливо зазначити, що перевірка транзакцій і майнінг біткойнів — це окремі процеси. Майнінг все ще може відбуватися незалежно від того, додані транзакції до блокчейну чи ні. Подібним чином вибух транзакцій біткойнів не обов'язково збільшує швидкість, з якою майнери знаходять нові блоки.

Незалежно від обсягу транзакцій, які очікують на підтвердження, біткойн запрограмовано на додавання нових блоків до блокчейну приблизно раз на 10 хвилин.

Завдяки публічному характеру блокчейну всі учасники мережі можуть відстежувати та оцінювати транзакції біткойнів у режимі реального часу. Ця інфраструктура зменшує ймовірність проблеми з онлайн-платежами, відомої як подвійні витрати. Подвійні витрати виникають, коли користувач намагається витратити ту саму криптовалюту двічі.

Боб, який має 1 біткойн, може спробувати надіслати його Ріші та Елізі одночасно та сподіватися, що система його не помітить.

У традиційній банківській системі подвійні витрати запобігають, оскільки звірку виконує центральний орган. Це також не проблема з фізичною готівкою, тому що ви не можете передати двом людям одну і ту саму доларову купюру.

					ДП.КН. 9675613.064.ПЗ	Арк.
Зм.	Арк.	№ докум.	Підпис	Дата		16

Проте біткойн має тисячі копій однієї і тієї ж книги, тому вся мережа користувачів повинна одноставно погодитися щодо дійсності кожної транзакції біткойна, яка відбувається. Ця угода між усіма сторонами називається «консенсусом».

Подібно до того, як банки постійно оновлюють баланси своїх користувачів, кожен, хто має копію реєстру біткойнів, відповідає за підтвердження та оновлення балансів усіх власників біткойнів.

Мережа біткойн автоматично випускає щойно викарбувані біткойни для майнерів, коли вони знаходять і додають нові блоки в блокчейн. Загальна пропозиція біткойнів має обмеження в 21 мільйон монет, тобто коли кількість монет в обігу досягне 21 мільйона, протокол припинить карбувати нові монети. У певному сенсі майнінг біткойнів одночасно виконує функцію перевірки транзакцій і процесу випуску біткойнів (поки не буде видобуто всі монети, він функціонуватиме лише як процес перевірки транзакцій).

Важливо відзначити, що збільшення обсягу обчислювальної потужності, призначеної для майнінгу біткойнів, не означатиме збільшення кількості біткойнів. Майнери з більшою обчислювальною потужністю лише збільшують свої шанси отримати винагороду наступним блоком, тому кількість видобутих біткойнів залишається відносно стабільною з часом.

Мережа біткойн використовує стратегію розподілу монет, відому як «халвінг біткойнів», яка забезпечує зменшення кількості біткойнів, що розподіляються майнерам, з часом. Поступово зменшуючи пропозицію нових біткойнів, що надходять в обіг, ідея полягає в тому, що це допоможе підтримати ціну активу (на основі фундаментальних принципів попиту та пропозиції).

Халвінг біткойна (іноді його називають «халвенінгом») відбувається кожні 210 000 блоків або приблизно чотири роки. Коли протокол біткойн вперше був запущений у 2009 році, кожен успішний майнер отримував 50 біткойнів (BTC) як винагороду за блок. Перенесімося до 2021 року: винагорода за блокування тепер становить 6,25 BTC, що

					ДП.КН. 9675613.064.ПЗ	Арк.
Зм.	Арк.	№ докум.	Підпис	Дата		17

є зменшенням порівняно з 12,5 BTC до удвічі зниження курсу біткойнів у травні 2020 року.

Очікується, що наступний халвінг відбудеться десь у 2024 році, і винагорода за блок знову впаде до 3,125 BTC. Цей процес триватиме до тих пір, поки врешті-решт не залишиться монет для видобутку.

Сьогодні в обігу перебуває понад 18,7 мільйона BTC, тобто залишилося ввести в обіг лише 2,25 мільйона BTC. Однак, беручи до уваги принцип удвічі та інші мережеві фактори, як-от складність майнінгу, передбачається, що останній біткойн буде видобуто приблизно в 2140 році.

Отже, прогнозування ціни біткойна є актуальним завданням з огляду на його унікальні характеристики та майбутні перспективи. Зокрема, очікується наближення наступного халвінгу, що призведе до зменшення винагороди за блок та зменшення обсягу нових монет. Такі події впливають на доступність та ціну біткойна. Зараз у обігу перебуває значна частина біткойнів, і передбачається, що останній біткойн буде видобуто лише у 2140 році. Тому, прогнозування ціни біткойна дозволяє інвесторам, трейдерам та іншим учасникам ринку зробити обґрунтовані рішення та використати можливості, що надає цей цифровий актив.

1.2 Аналіз відомих рішень

Цифрова трансформація економік є найсерйознішим зривом, який зараз відбувається у всіх економіках і фінансових системах. Економіки та фінансові системи світу стають цифровими безпрецедентно швидкими темпами. Згідно з нещодавнім звітом, розмір цифрової економіки у 2025 році оцінюється у 25% (\$23 трлн), що складається з матеріальних та нематеріальних цифрових активів [1]. Найновішою технологією створення та витрачання цифрових активів є технологія розподіленого реєстру (DLT), а найвідомішим її застосуванням є криптовалюта під назвою Bitcoin [2]. Слідом за цими

					ДП.КН. 9675613.064.ПЗ	Арк.
Зм.	Арк.	№ докум.	Підпис	Дата		18

розробками технологія блокчейн знайшла своє місце на перетинах фінтех і мереж наступного покоління [3].

Важливим питанням щодо нематеріальних цифрових активів, і особливо криптовалют, є волатильність цін. Ціну біткоїни (BTC) на період з лютого 2018 року по лютий 2023 року можна побачити на рис. 1.1. Ціни BTC демонстрували надзвичайну волатильність у цей період. Ціна зросла на 1900% у 2017 році, послідовно втративши 72% своєї вартості у 2018 році [4]. До 2013 року популярний інтерес до BTC, його використання у віртуальних транзакціях та ціни були низькими. Цей період не враховується в наших моделях. Хоча ціни BTC демонструють надзвичайну волатильність, BTC як цифровий актив досить стійкий, оскільки він може відновити свою вартість після значних падінь і навіть коли невизначеність на ринку висока, наприклад, під час пандемії COVID-19 [5].

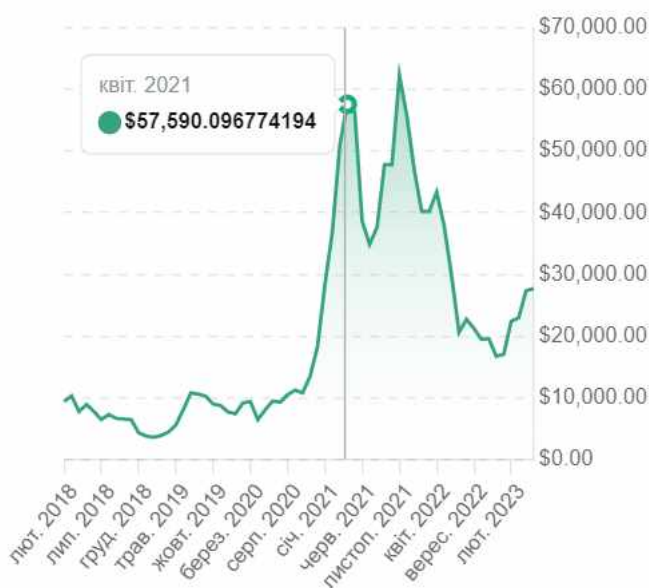


Рисунок 1.1 - Ціни на біткойн (BTC) з лютого 2018 року по лютий 2023 року

Незважаючи на швидко мінливий характер, ціна BTC була областю, де різні дослідники представили зусилля для прогнозування цін. У ряді досліджень обговорювалося, чи передбачувані ціни BTC за допомогою технічних індикаторів і продемонстрували існування значної

передбачуваності повернення [6, 7]. Інші недавні дослідження, такі як [8, 9] і [10], застосовували різні методи, пов'язані з машинним навчанням, для прогнозування цін на кінець дня і прогнозування зростання/зниження цін. [9] повідомляється про максимальну точність до 63% для прогнозування підвищення або зниження цін. [10] повідомив про 98% успіху для щоденного прогнозу цін. Однак часові періоди цих досліджень були обмежені даними — до 1 квітня 2017 року [10] і до 5 березня 2018 року [9].

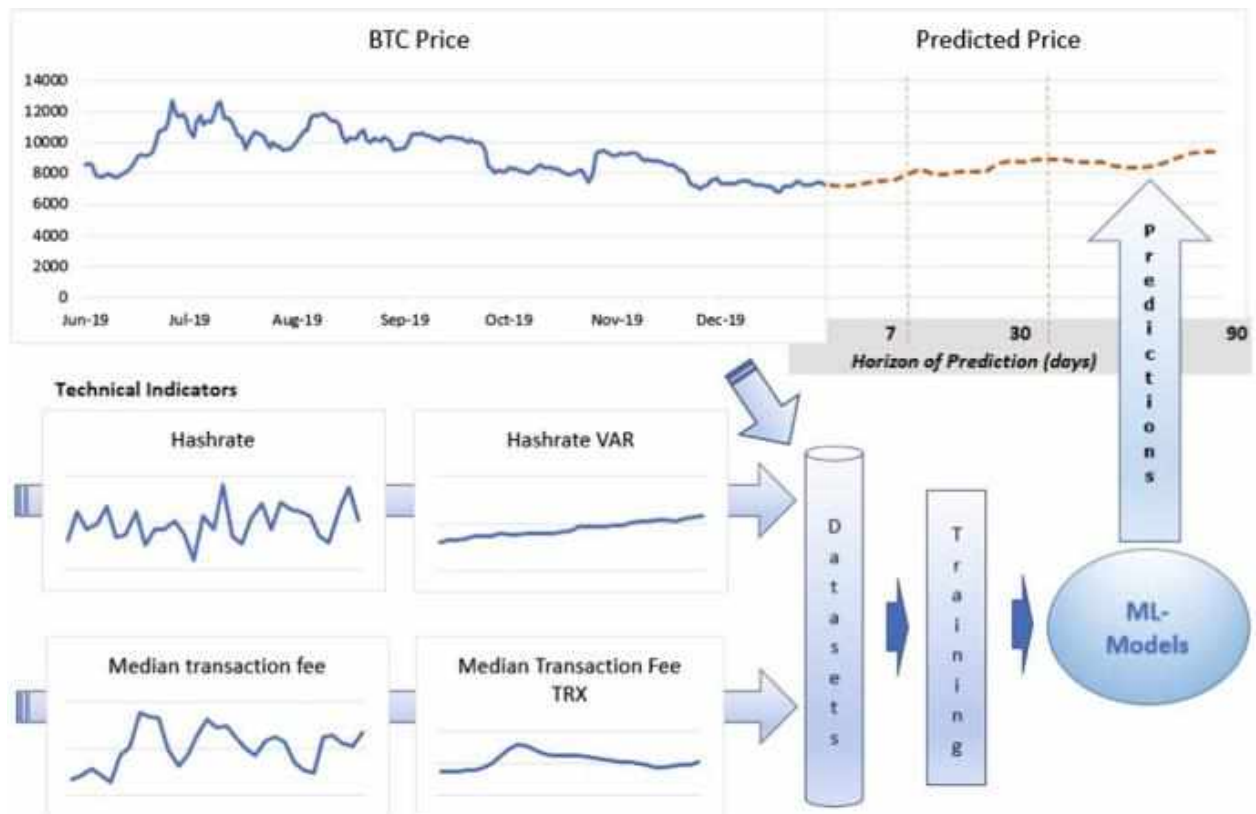


Рисунок 1.2 - Прогноз часових рядів на основі машинного навчання з використанням технічних індикаторів

Коли наприкінці 2013 року біткойн почав привертати увагу всього світу, він став свідком значного коливання його вартості та кількості транзакцій [11]. У літературі розглядалася передбачуваність повернення BTC за допомогою різних параметрів, таких як увага соціальних медіа [12, 13] та історичні технічні показники, пов'язані з BTC [14]. Одна група вивчила період з 4 вересня 2014 року по 31 серпня 2018 року, зафіксувавши, скільки разів термін «біткойн» був опублікований в Twitter. Результати показали, що кількість

твітів у Twitter може вплинути на обсяг торгів BTC на наступний день [15]. Більш того, [16] вивчили вплив коментарів користувачів в онлайн-платформах на коливання цін і кількість транзакцій криптовалют і виявили, що BTC особливо корелює з кількістю позитивних коментарів в соціальних мережах. Вони повідомили про точність 79% разом з тестом причинно-наслідкового зв'язку Грейнджера, який передбачає, що думки користувачів корисні для прогнозування коливань цін.

Що стосується прогнозів часових рядів, то існує три різних типи модельних підходів до прогнозування часових рядів відповідно до [17]. Перший підхід, чисті моделі, використовує лише історичні дані про змінну, яку потрібно передбачити. Прикладами моделей прогнозування чистих часових рядів є Autoregressive Integrated Moving Average (ARIMA) [18] та Generalized AutoRegressive Conditional Heteroskedasticity (GARCH) [19]. [20] представляє прогноз часових рядів на основі ARIMA для цін BTC на наступний день. Однак ми ще не бачили дослідження, заснованого на GARCH.

Моделі чистих часових рядів більше підходять для одновимірних і стаціонарних даних часових рядів. У цій статті ми зосередимося на машинному навчанні з функціями вищого рівня, а не на традиційних моделях з наступних причин. Перш за все, ціни на BTC дуже волатильні і нестаціонарні. Ми демонструємо, що ціни BTC нестаціонарні в наступному розділі. По-друге, в даних є велика кількість функцій, і запропонована методологія машинного навчання обробляє автокореляцію, сезонність і трендові ефекти, в той час як процес навчання моделей чистих часових рядів вимагає ручного налаштування для усунення цих ефектів.

Другий підхід, пояснювальні моделі, використовує функцію предикторних змінних для прогнозування цільової змінної в майбутньому часі. Підходи до прогнозування часових рядів на основі моделей мають недолік попереднього припущення про розподіл даних. Наприклад, [20] і [21] засновані на лог-трансформації цін BTC. Аналогічним чином, [21] використовував щоденні дані BTC з вересня 2011 року по серпень 2017 року

для проведення емпіричного дослідження з моделювання та прогнозування ціни BTC, які порівнюють байєсівську нейронну мережу (BNN) з іншими лінійними та нелінійними еталонними моделями. Вони виявили, що BNN добре прогнозує ціну, перетворену в журнал BTC, пояснюючи високу волатильність ціни BTC. Однак вищезгадані дослідження використовували перетворені в лог ціни для звітування про показники ефективності, які вводять в оману, оскільки такі значення, як правило, нижчі, ніж показники ефективності, обчислені з використанням реальних цін. Ми проаналізували це, розрахувавши показники продуктивності з використанням лог-нормалізованих значень та порівнявши їх із ненормалізованими журналами для наших власних результатів, і виявили, що, хоча нормалізований прогноз ціни звітує про набагато менше значення MAPE, фактична похибка може бути до 10 разів вищою.

Оскільки ціни на криптовалюту є нелінійними і нестационарними, припущення про розподіл даних можуть мати несприятливий вплив на прогнозні показники. Нестационарні моделі часових рядів демонструють статистичні розподіли, що розвиваються з плином часу, що призводить до зміни поведінки залежностей між вхідними та вихідними змінними. Підходи, засновані на машинному навчанні, використовують властиві нелінійні та нестационарні аспекти даних. Вони також можуть скористатися пояснювальними особливостями, беручи до уваги основні фактори, що впливають на прогнозовану змінну. Існує кілька досліджень з моделювання та прогнозування ціни BTC за допомогою машинного навчання,

[22] використовував байєсівський метод регресії, який використовує модель латентного джерела, яка була розроблена [23] для прогнозування варіації цін з використанням історичних даних BTC. [24] використовував машинне навчання та інженерію функцій для дослідження того, як функції мережі BTC можуть впливати на рух ціни BTC. Вони отримали точність класифікації 55%. [9] використовував штучну нейронну мережу (ШНМ) для досягнення точності класифікації 65%. Крім того, [25] передбачив ціну BTC,

					ДП.КН. 9675613.064.ПЗ	Арк.
Зм.	Арк.	№ докум.	Підпис	Дата		22

використовуючи байєсівську оптимізовану рекурентну нейронну мережу (RNN) і довгу короткочасну пам'ять (LSTM). Точність класифікації, якої вони досягли, становила 52% при використанні LSTM з RMSE 8%. Вони також повідомили, що в прогнозуванні нелінійні моделі глибокого навчання показали кращі результати, ніж ARIMA. [10] використовував алгоритми ANN та SVM у регресійних моделях для прогнозування мінімальних, максимальних та кінцевих цін BTS та повідомив, що алгоритм SVM показав найкращі результати з MAPE 1.58%. Одним з останніх досліджень прогнозування щоденних цін BTC є [8], який використовував високовимірні функції з різними алгоритмами машинного навчання, такими як SVM, LSTM і випадковий ліс. Для прогнозу цін на наступний день з липня 2017 року по січень 2018 року найвищої точності в 65,3% досягла SVM.

На основі розглянутих рішень прогнозування біткойну, найкращим методом є ARIMA через свою простоту. Хоча деякі дослідження використовують складніші моделі, такі як штучні нейронні мережі і байєсівські методи, ARIMA все ж досягає задовільної точності в прогнозуванні ціни біткойна. Інші методи можуть мати вищу точність в деяких випадках, але вимагають складнішої підготовки даних та використання більш складних алгоритмів. Отже, метод ARIMA є простим і ефективним інструментом для прогнозування ціни біткойна.

1.3 Вибір оптимального різновиду ARIMA для прогнозування ціни біткойна

ARIMA (Autoregressive Integrated Moving Average) є популярним методом прогнозування часових рядів, зокрема цін біткойна. ARIMA складається з трьох основних компонентів: авторегресії (AR), інтегрування (I) та ковзного середнього (MA).

					ДП.КН. 9675613.064.ПЗ	Арк.
Зм.	Арк.	№ докум.	Підпис	Дата		23

1. AR (Авторегресія): Включає в себе використання попередніх значень ряду для прогнозування майбутніх значень. Параметр AR показує, скільки попередніх значень використовується для прогнозу.
2. I (Інтегрування): Включає в себе диференціювання ряду для отримання стаціонарного ряду. Інтегрування може включати перший (диференціювання одного разу) або більше порядків (диференціювання кілька разів), щоб забезпечити стаціонарність ряду.
3. MA (Ковзне середнє): Включає в себе використання попередніх помилок прогнозування для прогнозування майбутніх значень. Параметр MA показує, скільки попередніх помилок використовується для прогнозу.

ARIMA може мати різні варіанти, наприклад, $ARIMA(p, d, q)$, де p , d і q - це параметри, що визначають кількість попередніх значень, рівень диференціювання та кількість попередніх помилок, відповідно. Варіанти ARIMA використовуються для моделювання і прогнозування різних типів часових рядів, включаючи ціни біткойна.

ARIMA (Autoregressive Integrated Moving Average) є основним класом моделей для прогнозування часових рядів. Окрім базового ARIMA, існують його різновиди, такі як SARIMA, ARIMAX, SARIMAX і низка інших. Вони розширюють базову модель ARIMA, додаючи додаткові функціональності для більш точного та гнучкого прогнозування.

1. SARIMA (Seasonal ARIMA): Цей варіант ARIMA дозволяє моделювати сезонність в часовому ряді. Включає в себе параметри, що дозволяють ураховувати сезонні зміни та їх вплив на прогнозування.

Переваги SARIMA (Seasonal ARIMA) моделі:

- Моделювання сезонності: SARIMA дозволяє ефективно моделювати сезонні зміни в часовому ряді. Це особливо корисно, коли сезонні фактори мають значний вплив на поведінку ряду.
- Гнучкість: SARIMA може враховувати різні типи сезонності, такі як добова, тижнева, місячна або інші. Це дозволяє моделі адаптуватися до різноманітних часових структур.

					ДП.КН. 9675613.064.ПЗ	Арк.
Зм.	Арк.	№ докум.	Підпис	Дата		24

- Управління трендом: SARIMA може моделювати і управляти трендом в часовому ряді. Вона дозволяє враховувати лінійні або нелінійні зміни тренду, що допомагає вдосконалити прогнози.

Недоліки SARIMA моделі:

- Складність підбору параметрів: SARIMA має багато параметрів, включаючи коефіцієнти авторегресії (AR), коефіцієнти рухомого середнього (MA) і сезонні коефіцієнти. Вибір оптимальних значень цих параметрів може бути витратним процесом, особливо для складних часових рядів.
- Обробка аномалій: SARIMA модель припускає, що часовий ряд має стаціонарну структуру. Аномалії або викиди в ряді можуть вплинути на точність прогнозів, оскільки модель може неправильно враховувати такі незвичайні події.
- Вимоги до обсягу даних: SARIMA може вимагати значної кількості даних для ефективного моделювання сезонності. Якщо у вас обмежена кількість спостережень, можуть виникнути проблеми з точністю прогнозів.

Незважаючи на недоліки, SARIMA є потужним інструментом для прогнозування часових рядів з сезонністю та варіабельністю. Вона дозволяє ураховувати складні сезонні залежності і тренди, що забезпечує точність прогнозування. Крім того, SARIMA може бути використана для прогнозування на довгі періоди, включаючи сезонні коливання.

Однак, важливо враховувати, що SARIMA модель потребує правильного підбору параметрів, що може бути викликом. Також, при наявності аномалій або викидів в ряді, модель може давати неточні прогнози. Крім того, залежно від обсягу даних, можуть виникати обмеження в її застосуванні.

Загалом, SARIMA є одним з найефективніших методів прогнозування часових рядів з сезонністю і є популярним в галузі фінансів, економіки та інших сферах, де сезонні коливання мають суттєвий вплив.

					ДП.КН. 9675613.064.ПЗ	Арк.
Зм.	Арк.	№ докум.	Підпис	Дата		25

2. ARIMAX (ARIMA with Exogenous Variables): Цей варіант ARIMA дозволяє включати додаткові екзогенні змінні, які не є частиною самого ряду, але можуть впливати на його поведінку. Це можуть бути такі фактори, як економічні показники, погодні умови, події тощо.

ARIMAX є потужним варіантом ARIMA, оскільки дозволяє враховувати зовнішні фактори, які можуть впливати на часовий ряд. Основні переваги ARIMAX включають:

- Урахування зовнішніх факторів: ARIMAX дозволяє включати екзогенні змінні, що дозволяє моделювати і враховувати їх вплив на ряд. Це дозволяє отримати більш точні та адаптивні прогнози.
- Більш гнучкий прогноз: Додавання екзогенних змінних дозволяє моделі адаптуватися до змінних умов та факторів, що впливають на ряд. Це покращує гнучкість та робить ARIMAX більш універсальним для прогнозування в різних сферах.
- Управління ендегенними та екзогенними залежностями: ARIMAX дозволяє аналізувати та керувати взаємозв'язком між ендегенними (внутрішніми) та екзогенними (зовнішніми) змінними. Це дозволяє виявити та врахувати складні взаємодії між цими змінними.

Незважаючи на переваги, ARIMAX також має деякі недоліки:

- Складність визначення екзогенних змінних: Вибір та включення екзогенних змінних може бути складним процесом, особливо якщо немає чіткої теоретичної бази для вибору цих змінних.
- Більший обсяг даних: Включення екзогенних змінних вимагає наявності відповідних даних для цих змінних. Це може збільшити обсяг даних, які потрібно аналізувати та обробляти.
- Підбір параметрів: Включення екзогенних змінних у модель ARIMAX вимагає підбору відповідних параметрів моделі. Це може бути складним завданням, особливо при великій кількості екзогенних змінних.

- Пояснювальна сила: Додавання екзогенних змінних може покращити точність прогнозу, але може ускладнити інтерпретацію результатів. Розуміння взаємозв'язку між ендегенними та екзогенними змінними може стати складним завданням.

Усупереч недолікам, ARIMAX залишається потужним інструментом для прогнозування, особливо в ситуаціях, де зовнішні фактори можуть впливати на часовий ряд. Він дозволяє отримати більш точні та адаптивні прогнози, що допомагає у прийнятті кращих рішень та плануванні майбутніх подій.

3. SARIMAX (Seasonal ARIMA with Exogenous Variables): Це поєднання SARIMA і ARIMAX, яке дозволяє моделювати сезонність та включати екзогенні змінні одночасно.

Переваги SARIMAX:

- Урахування сезонності: SARIMAX дозволяє ефективно моделювати сезонні зміни у часовому ряді, що дозволяє отримати більш точні прогнози в умовах сезонної варіації даних.
- Врахування впливу екзогенних змінних: SARIMAX дозволяє включати додаткові фактори, які можуть впливати на поведінку часового ряду. Це дозволяє враховувати зовнішні впливи та отримувати більш точні та адаптивні прогнози.
- Гнучкість моделювання: SARIMAX надає гнучкість у виборі параметрів моделі, таких як ступінь сезонності, порядок авторегресії, порядок диференціювання тощо. Це дозволяє налаштовувати модель під конкретні потреби досліджуваного часового ряду.

Недоліки SARIMAX:

- Складність моделювання: SARIMAX вимагає підбору великої кількості параметрів, зокрема стосовно сезонності та екзогенних змінних. Це може бути викликом, особливо при обробці складних даних з багатьма взаємозв'язаними факторами.
- Обмежена пояснювальна сила: Додавання сезонності та екзогенних змінних може ускладнити інтерпретацію результатів моделі.

					ДП.КН. 9675613.064.ПЗ	Арк.
Зм.	Арк.	№ докум.	Підпис	Дата		27

Розуміння взаємозв'язку між різними компонентами моделі може стати складним завданням.

Незважаючи на недоліки, SARIMAX є потужним інструментом для прогнозування, особливо в ситуаціях, коли на поведінку часового ряду впливають як сезонні фактори, так і зовнішні екзогенні змінні. Він надає можливість отримувати більш точні та адаптивні прогнози, що є важливим для багатьох досліджень та прийняття рішень. Враховуючи сезонність і екзогенні фактори, SARIMAX дозволяє моделювати складні залежності та покращує якість прогнозування. Проте, необхідно пильно відбирати та налаштовувати параметри моделі для досягнення оптимальних результатів. В цілому, SARIMAX є корисним інструментом для прогнозування з урахуванням сезонності та екзогенних факторів.

Кожен з цих варіантів ARIMA має свої особливості і може бути використаний для прогнозування різних типів часових рядів залежно від їхніх характеристик та наявності додаткових факторів, що впливають на ряд. Вибір конкретного варіанту ARIMA залежить від конкретної задачі прогнозування та вимог до точності та гнучкості моделі.

Отже, SARIMA є відмінним вибором для прогнозування ціни біткойна в Модулі прогнозування. Оскільки біткойн має сезонні залежності і складну динаміку, SARIMA дозволяє враховувати ці сезонності та зміни у часовому ряді. При цьому модель SARIMA є відносно простою у реалізації і має добрі результати у прогнозуванні. Це забезпечує нам зручний і ефективний інструмент для прогнозування ціни біткойна, здатний враховувати його особливості і забезпечувати достатню точність прогнозів.

1.4 Постановка задачі дослідження

Актуальність дипломної роботи полягає в тому, що біткойн є однією з найпопулярніших криптовалют у світі, і його ціна підлягає значним

					ДП.КН. 9675613.064.ПЗ	Арк.
						28
Зм.	Арк.	№ докум.	Підпис	Дата		

коливанням. Прогнозування ціни біткойна має велике значення для інвесторів, трейдерів та фахівців у галузі криптовалют, оскільки це дозволяє їм приймати обґрунтовані рішення щодо купівлі, продажу або тримання біткойна.

Також актуальність полягає у тому, що прогнозування ціни біткойна є складною задачею через його високу волатильність та залежність від різних факторів, таких як новини, регулятивні рішення, технічні та фундаментальні показники. Розробка ефективного модуля прогнозування ціни біткойна на основі методу ARIMA дозволить отримувати більш точні та надійні прогнози, що сприятиме зниженню ризиків та збільшенню доходів.

Додатково, дослідження прогнозування ціни біткойна на основі ARIMA є важливим з точки зору розвитку фінансових стратегій та аналітики ринку криптовалют. Це може допомогти в розумінні динаміки цін, виявленні трендів та паттернів, а також встановленні ефективних торгових стратегій.

Метою дипломної роботи є розробка та реалізація модуля прогнозування ціни біткойна на основі методу ARIMA. Головною метою є створення надійного та точного інструменту, який зможе аналізувати історичні дані ціни біткойна, виявляти тренди та сезонні зміни, а також робити передбачення майбутньої ціни.

Для досягнення поставленої мети дипломної роботи було необхідно провести теоретичні дослідження, визначити основні аспекти та послідовно вирішити наступні завдання:

1. Провести аналіз поняття біткойну, його особливостей та роль у сучасній економіці.
2. Оглянути та проаналізувати існуючі рішення та методи прогнозування ціни біткойна.
3. Визначити оптимальний різновид ARIMA для прогнозування ціни біткойна, обґрунтувати вибір та розробити алгоритм його застосування.
4. Розробити алгоритм визначення стаціонарності та декомпозиції часового ряду ціни біткойна.

5. Реалізувати алгоритм прогнозування ціни біткойна на основі методу SARIMA з урахуванням сезонності.

6. Здійснити аналіз набору даних ціни біткойна, включаючи статистичний аналіз, виявлення трендів та сезонності.

7. Побудувати модель прогнозування на основі SARIMA з використанням вибраних параметрів та перевірити її точність та надійність.

8. Зробити прогноз ціни біткойна на майбутні періоди, використовуючи розроблену модель та провести валідацію результатів.

Завдання з 1 до 3 описані у розділах 1.1-1.4, а завдання з 3 до 6 розглядаються у параграфах 2.1-3.3 відповідно.

					ДП.КН. 9675613.064.ПЗ	Арк.
						30
Зм.	Арк.	№ докум.	Підпис	Дата		

2 АЛГОРИТМІЧНЕ ЗАБЕЗПЕЧЕННЯ

2.1 Алгоритм визначення стаціонарності та декомпозиція часового ряду

Задача прогнозування часового ряду є важливою складовою багатьох досліджень і прикладних областей. Для досягнення точного та надійного прогнозу необхідно попередньо визначити стаціонарність ряду та розкласти його на складові. Алгоритм визначення стаціонарності та декомпозиції часового ряду надає процедуру для перевірки стаціонарності, виявлення тренду, сезонності та залишкової складової. Цей підхід дозволяє краще розуміти структуру ряду і підготувати дані для подальшого прогнозування з використанням відповідних моделей.

Алгоритм визначення стаціонарності та декомпозиції часового ряду включає наступні кроки:

Крок 1. Перевірка стаціонарності: Виконати статистичний тест, такий як тест Діккі-Фуллера, для перевірки наявності стаціонарності в часовому ряді. Якщо ряд нестаціонарний, застосувати методи для зроблення ряду стаціонарним, наприклад, знімання тренду або різницю між послідовними значеннями.

Математично алгоритм перевірки стаціонарності часового ряду за допомогою тесту Діккі-Фуллера можна описати наступним чином:

1. Припущення:

- Нульова гіпотеза (H_0): Часовий ряд має одиничний корінь, тобто нестаціонарний.
- Альтернативна гіпотеза (H_1): Часовий ряд є стаціонарним.

2. Обчислення статистики тесту:

- Вираховується тестова статистика, яка залежить від ряду і має специфічний розподіл при виконанні нульової гіпотези.
- Зазвичай використовується статистика t , яка обчислюється як співвідношення зміни ряду до стандартного відхилення.

					ДП.КН. 9675613.064.ПЗ	Арк.
Зм.	Арк.	№ докум.	Підпис	Дата		31

3. Визначення критичної області:

- Встановлюється критичний рівень значущості (наприклад, $\alpha = 0.05$).
- За допомогою критичного рівня значущості обчислюється критична значущість тесту.

4. Прийняття рішення:

- Якщо значення тестової статистики менше критичної значущості, то нульова гіпотеза відхиляється, і часовий ряд вважається стаціонарним.
- Якщо значення тестової статистики більше критичної значущості, то немає достатніх доказів для відхилення нульової гіпотези, і часовий ряд вважається нестаціонарним.

У випадку, коли часовий ряд виявляється нестаціонарним, можна застосувати методи для зроблення ряду стаціонарним, наприклад:

- Знімання тренду: Використовуючи методи, такі як згладжування ковзним середнім або вагове згладжування, можна видалити трендову складову з ряду.
- Різницю між послідовними значеннями: Цей підхід включає обчислення різниці між поточним значенням ряду і попереднім значенням. Це допомагає усунути сезонну складову та інші нелінійні залежності.
- Логарифмування: Використання логарифму може допомогти зменшити варіацію величини ряду та нормалізувати його.
- Сезонне згладжування: Застосування методів, таких як згладжування експоненційним зваженим середнім (Exponential Weighted Moving Average, EWMA) або сезонного декомпозиційного згладжування, може допомогти усунути сезонну складову з ряду.

Ці методи допомагають зробити ряд стаціонарним, тобто знизити варіацію, усунути тренди та сезонність, що дозволяє застосовувати моделі прогнозування, такі як SARIMA, які передбачають стаціонарність ряду. Після

застосування цих методів можна використовувати SARIMA для прогнозування майбутніх значень часового ряду.

Крок 2. Декомпозиція ряду: Використати методи декомпозиції, такі як STL (Seasonal and Trend decomposition using Loess), для розбиття часового ряду на компоненти: тренд, сезонність та залишкову складову.

Крок 3. Визначення тренду: Використовуючи післядекомпозиційні методи, визначити трендову складову ряду, яка представляє довгострокову зміну.

Метод декомпозиції STL (Seasonal and Trend decomposition using Loess) використовується для розбиття часового ряду на трендову, сезонну та залишкову складові. Цей метод базується на локально згладженій регресії, де кожна з компонент моделюється окремо.

Нехай маємо часовий ряд Y з n спостереженнями. Метод STL виконує такі кроки:

1. Застосування локально згладженої регресії для вилучення трендової складової:
 - Вибирається параметр ширини вікна, який визначає, скільки спостережень використовується для згладжування на кожному кроці.
 - На кожному кроці обчислюється локальна регресія між спостереженнями із підвибірки, використовуючи наприклад метод найменших квадратів.
 - Отримуємо трендову складову шляхом віднімання прогнозованих значень від вихідного ряду.
2. Вилучення сезонної складової:
 - Виконується сгладжування ряду трендових значень за допомогою низькочастотного фільтра.
 - Отримуємо сезонну складову шляхом віднімання згладжених трендових значень від вихідного ряду.
3. Отримання залишкової складової:

					ДП.КН. 9675613.064.ПЗ	Арк.
						33
Зм.	Арк.	№ докум.	Підпис	Дата		

- Залишкова складова отримується шляхом віднімання суми трендової та сезонної складових від вихідного ряду.
- Результатом декомпозиції є трендова, сезонна та залишкова складові, які можна використовувати для аналізу та прогнозування часового ряду.

Крок 4. Визначення сезонності: Використати методи для виявлення сезонності в часовому ряді, такі як автокореляційна функція (ACF) або періодограма. Визначити періодичність та амплітуду сезонної складової.

Для визначення сезонності в часовому ряді можна використовувати методи, такі як автокореляційна функція (ACF) і періодограма. Автокореляційна функція (ACF): Нехай маємо часовий ряд Y з n спостереженнями. Автокореляційна функція вимірює кореляцію між значеннями ряду Y , зсувом на певну кількість періодів. Для виявлення сезонності, ми розглядаємо зсуви, які відповідають можливим періодам сезонності.

1. Автокореляційна функція обчислюється наступним чином:

- Обчислити середнє значення ряду Y .
- Обчислити коваріацію між Y та його зсувом на 1 період, коваріацію між Y та зсувом на 2 періоди і так далі.
- Нормалізувати коваріації, поділивши їх на дисперсію ряду Y .
- Отримуємо послідовність автокореляційних коефіцієнтів, що відображають кореляцію між рядом Y та його зсувами.

Періодограма: Періодограма вимірює силу амплітуди сигналу на різних частотах. Для визначення сезонності, ми шукаємо піки в періодограмі, які вказують на домінуючі періоди в ряді.

2. Періодограма обчислюється наступним чином:

- Виконати перетворення Фур'є (FFT) на часовий ряд Y , щоб отримати його спектральний вміст.

					ДП.КН. 9675613.064.ПЗ	Арк.
Зм.	Арк.	№ докум.	Підпис	Дата		34

- Обчислити квадрат модуля комплексних значень FFT для отримання спектральної щільності потоку.
- Отримуємо послідовність значень спектральної щільності, яка відображає амплітуду сигналу на різних частотах.

Крок 5. Визначення залишкової складової: Залишкова складова отримується відніманням тренду та сезонної складової від оригінального ряду. Це представляє непояснену варіацію в часовому ряді.

Алгоритм визначення стаціонарності та декомпозиції часового ряду допомагає розкрити структуру ряду та виділити різні компоненти, що дає змогу краще розуміти його поведінку та застосовувати відповідні моделі для прогнозування.

2.2 Алгоритм прогнозування методом SARIMA

SARIMA (Seasonal Autoregressive Integrated Moving Average) є розширенням моделі ARIMA для прогнозування часових рядів з сезонною складовою. Вона враховує як залежність від попередніх значень (авторегресійну складову), так і залежність від попередніх помилок (рухомий середній) в поєднанні з сезонною складовою.

Математично модель SARIMA може бути представлена наступним чином:

$$\text{SARIMA}(p, d, q)(P, D, Q, s):$$

- p : ступінь авторегресійної складової (AR order)
- d : ступінь різницевої складової (порядок диференціювання)
- q : ступінь складової ковзного середнього (MA order)
- P : ступінь сезонної авторегресійної складової (Seasonal AR order)

- D: ступінь сезонної різницевої складової (порядок сезонного диференціювання)
- Q: ступінь сезонної складової ковзного середнього (Seasonal MA order)
- s: період сезонності

Алгоритм прогнозування методом SARIMA (Seasonal Autoregressive Integrated Moving Average) базується на комбінації ARIMA моделі з сезонністю і включає наступні кроки та представлено у вигляді блок-схеми (див.дод.А):

Крок 1. Перетворення даних: Початкові дані про часовий ряд можуть потребувати перетворень для зроблення їх стаціонарними. Це може включати різницю між значеннями ряду або логарифмування.

Крок 2. Визначення параметрів моделі:

- AR (Autoregressive) - параметр, що відображає залежність між поточним значенням ряду і його попередніми значеннями на певних затримках.
- I (Integrated) - параметр, що відображає кількість різницювань, необхідних для зроблення ряду стаціонарним.
- MA (Moving Average) - параметр, що відображає залежність між поточним значенням ряду і помилками прогнозу на попередніх кроках.

Крок 3. Вибір оптимальних параметрів:

- Параметри моделі SARIMA можна вибрати за допомогою методів, таких як сітка пошуку, які оцінюють показники якості моделі (наприклад, AIC або BIC) для різних комбінацій параметрів.
- Критерії якості оцінюють, наскільки добре модель підходить до даних і як точно вона прогнозує майбутні значення.

Крок 4. Побудова моделі:

- Застосування підібраних параметрів до моделі SARIMA для побудови прогнозу.

Крок 5. Оцінка адекватності моделі:

					ДП.КН. 9675613.064.ПЗ	Арк.
Зм.	Арк.	№ докум.	Підпис	Дата		36

- Проведення аналізу залишків, щоб переконатися, що вони випадкові та некорельовані.
- Графіки автокореляції залишків можуть допомогти виявити наявність систематичних закономірностей.

Крок 6. Генерація прогнозу:

- Застосування побудованої моделі SARIMA до вихідних даних, включаючи останні доступні спостереження.
- Модель SARIMA використовує попередні значення ряду, помилки прогнозу та сезонну компоненту для генерації прогнозу на майбутній період.
- Прогноз може бути зроблений на довільну кількість майбутніх кроків.

Крок 7. Оцінка точності прогнозу:

- Порівняння прогнозованих значень з фактичними спостереженнями для оцінки точності моделі.
- Використання метрик, таких як середня абсолютна помилка (MAE), середня квадратична помилка (MSE) або коефіцієнт детермінації (R^2), для кількісної оцінки точності прогнозу.

Крок 8. Оновлення моделі:

- При отриманні нових спостережень можлива потреба в оновленні моделі.
- Проведення оновленого аналізу даних, підбір нових оптимальних параметрів та побудова нового прогнозу.

Цей алгоритм покроково описує процес прогнозування ціни біткойна за допомогою моделі SARIMA. Виконання кожного кроку вимагає використання відповідних статистичних методів та алгоритмів.

Алгоритм прогнозування методом SARIMA включає підготовку часового ряду, вибір оптимальних параметрів, побудову моделі, підганяння до тренувальної вибірки, прогнозування та оцінку точності. Цей алгоритм

підходить для прогнозування ціни біткойна з використанням моделі SARIMA, що дозволяє урахувати як сезонні, так і несезонні фактори в часовому ряді.

2.3 Алгоритм прогнозування ціни біткойна на основі методу SARIMA

Алгоритм прогнозування ціни біткойна на основі методу SARIMA (сезонний авторегресійний інтегрований ковзний середній) можна описати наступним чином та у вигляді блок-схеми (див.дод.Б):

Крок 1. Відбір даних: Збирання і підготовка часового ряду цін біткойна, який включає значення цін на регулярних інтервалах часу, наприклад, щомісячно.

Крок 2. Аналіз стаціонарності: Перевірка стаціонарності часового ряду шляхом визначення показників, таких як середнє значення і дисперсія, а також використання статистичних тестів, наприклад, тесту Діккі-Фуллера. Якщо ряд нестаціонарний, проводиться його перетворення для досягнення стаціонарності.

Крок 3. STL-декомпозиція: Застосування методу STL (Seasonal and Trend decomposition using Loess) для розкладання часового ряду на складові: тренд, сезонність і залишкову складову. Це дозволяє виявити структуру і закономірності в ряді.

Крок 4. Вибір моделі: Вибір оптимальної моделі SARIMA, що найкраще описує дані, з врахуванням показників, таких як AIC (Akaike Information Criterion) і BIC (Bayesian Information Criterion). Підбір оптимальних параметрів моделі SARIMA, включаючи показники сезонності, здійснюється шляхом перебору різних комбінацій параметрів і оцінки їх ефективності.

Крок 5. Побудова моделі: Побудова обраної моделі SARIMA на основі вибраних параметрів.

					ДП.КН. 9675613.064.ПЗ	Арк.
Зм.	Арк.	№ докум.	Підпис	Дата		38

Крок 6. Перевірка адекватності: Аналіз залишкових значень моделі SARIMA, включаючи тести на автокореляцію (ACF) і нормальність розподілу, для перевірки адекватності моделі.

Крок 7. Прогнозування: Використання побудованої моделі SARIMA для прогнозування майбутніх значень ціни біткойна. Зазвичай прогноз проводиться на певний часовий горизонт, наприклад, на місяць або кілька місяців вперед.

Крок 8. Інверсія перетворень: Якщо під час аналізу стаціонарності або побудови моделі були застосовані перетворення даних, необхідно здійснити інверсію цих перетворень для отримання прогнозованих значень в початкових одиницях виміру.

Крок 9. Візуалізація результатів: Побудова графіка, на якому відображені реальні значення ціни біткойна і прогнозовані значення, що дозволяє порівняти їх та оцінити ефективність моделі. Зазвичай на графіку відображаються також довірчі інтервали для прогнозованих значень.

Крок 10. Оцінка точності: Оцінка точності прогнозів шляхом порівняння прогнозованих значень з фактичними даними і використання різних метрик, таких як середня квадратична помилка (RMSE), середня абсолютна помилка (MAE) і коефіцієнт детермінації (R^2).

Крок 11. Виділення основних факторів: Аналіз основних факторів, що впливають на ціну біткойна, здійснюється на основі аналізу залишкових значень моделі. Це дозволяє зрозуміти, які фактори впливають на зміну ціни і як їх можна врахувати при подальшому прогнозуванні.

Постійне оновлення моделі: Цей алгоритм прогнозування ціни біткойна на основі методу SARIMA вимагає постійного оновлення моделі з урахуванням нових даних. Оскільки ціни біткойна постійно змінюються, важливо періодично оновлювати модель, враховуючи нові спостереження. Нові дані додаються до існуючих даних, а потім проводиться аналіз стаціонарності, перетворення даних, вибір моделі і прогнозування. Цей процес

дозволяє підтримувати актуальність і точність прогнозів на основі оновленої і більш актуальної інформації.

Алгоритм прогнозування ціни біткойна на основі методу SARIMA є потужним і ефективним інструментом для прогнозування цінових змін у криптовалютному ринку. Він враховує як сезонні, так і несезонні коливання ціни, а також використовує історичні дані для побудови моделі та зробити прогноз на майбутнє.

Прогнозування ціни біткойна може бути корисним для інвесторів, трейдерів і дослідників ринку, які прагнуть зрозуміти тенденції і ризики на ринку криптовалют і приймати обґрунтовані рішення. Однак варто зауважити, що прогнозування цін криптовалют є складним завданням, оскільки ринок досить волатильний і піддається впливу багатьох факторів, таких як регуляторні зміни, новини та загальна економічна ситуація. Тому прогнози слід розглядати як інструмент для отримання загального уявлення про тенденції, а не як точні передбачення майбутніх цін.

					ДП.КН. 9675613.064.ПЗ	Арк.
						40
Зм.	Арк.	№ докум.	Підпис	Дата		

3 ПРОГРАМНО-ТЕХНОЛОГІЧНЕ ЗАБЕЗПЕЧЕННЯ

3.1 Аналіз набору даних

Для реалізації модуля прогнозування ціни біткойна на основі методу ARIMA, обрано набір даних, що містить історичні дані ринку Bitcoin [32] з 2012 до 2021 року з 1-хвилинними інтервалами для вибраних бірж. Дані включають ціни відкриття, максимуму, мінімуму та закриття (OHLC), об'єм в BTC та вказану валюту, та вагову ціну Bitcoin. Мітки часу в Unix форматі. Він містить понад 4.8 мільйони записів.

У наборі даних є сім колонок, кожна з яких має тип даних **float64**. Ці колонки включають:

1. Open: Ціна відкриття Bitcoin у вказаний час.
2. High: Найвища ціна Bitcoin протягом цього періоду.
3. Low: Найнижча ціна Bitcoin протягом цього періоду.
4. Close: Ціна закриття Bitcoin у вказаний час.
5. Volume_(BTC): Обсяг торгівлі Bitcoin за цей період.
6. Volume_(Currency): Вартість обсягу торгівлі в вказаній валюті.
7. Weighted_Price: Середньозважена ціна Bitcoin за цей період.

Загальний об'єм пам'яті, який використовує цей DataFrame, складає приблизно 425,5 МБ.

Створимо 4 графіки (Рис.3.1), які відображають середню вартість біткоїну (в USD) за різними проміжками часу: за дні, за місяці, за квартали та за роки.

1. Перший графік (За дні) відображає динаміку ціни біткоїну кожного дня.
2. Другий графік (За місяці) відображає середню ціну біткоїну за місяць.
3. Третій графік (За квартали) відображає середню ціну біткоїну за квартал (три місяці).
4. Четвертий графік (За роки) відображає середню ціну біткоїну за рік.

					ДП.КН. 9675613.064.ПЗ	Арк.
Зм.	Арк.	№ докум.	Підпис	Дата		41

Обмін Біткоїну, середня вартість в USD

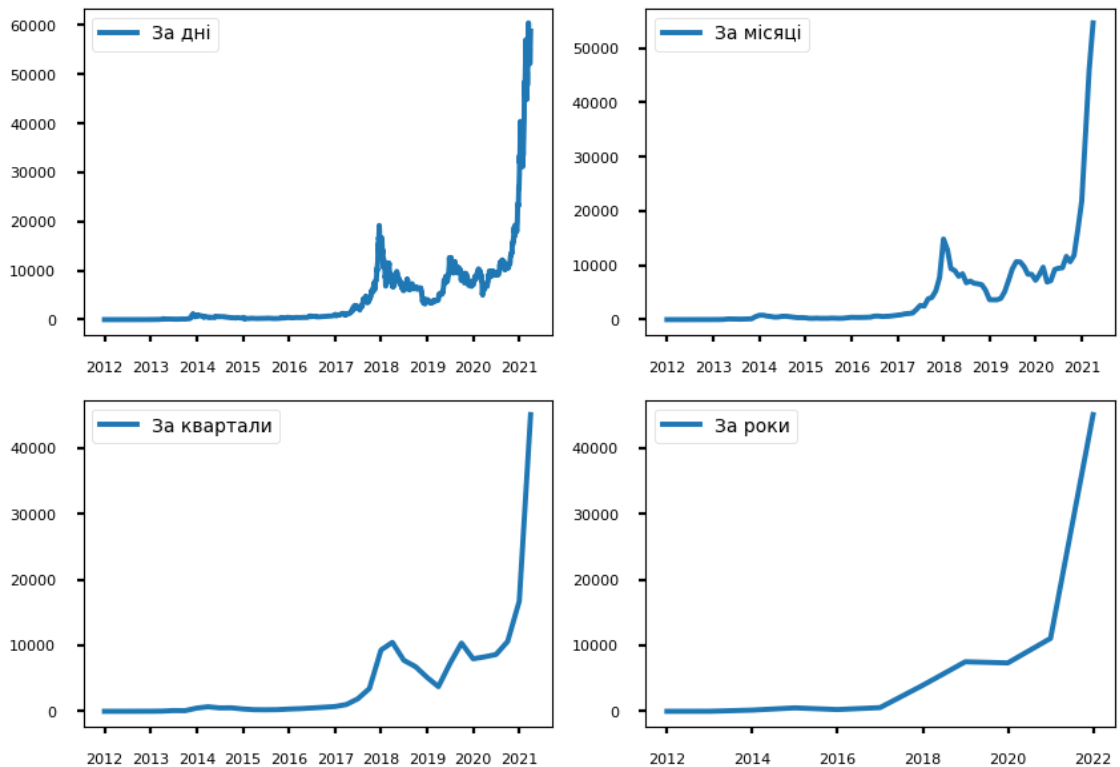


Рисунок 3.1 - Обмін Біткоїну, середня вартість в USD

Далі визначено кількість пропущених значень. Ця функція `missing_values_table(df)` приймає `DataFrame` як вхідний параметр, розраховує загальну кількість і відсоток пропущених значень для кожного стовпця. Вона повертає таблицю, впорядковану за відсотком пропущених значень. `df.interpolate()` - це функція `pandas`, яка використовується для заповнення пропущених значень у `DataFrame`. Вона виконує лінійну інтерполяцію в точках, де маються пропущені дані.

```

1 def missing_values_table(df):
2     # Загальна кількість пропущених значень
3     mis_val = df.isnull().sum()
4
5     # Відсоток пропущених значень
6     mis_val_percent = 100 * df.isnull().sum() / len(df)
7
8     # Створення таблиці з результатами
9     mis_val_table = pd.concat([mis_val, mis_val_percent], axis=1)
10
11     # перейменування стовпців
12     mis_val_table_ren_columns = mis_val_table.rename(
13         columns={0: 'пропущені значення', 1: '% від загальної кількості'})
14
15     # Сортування таблиці за відсотком пропущених значень у низхідному порядку
16     mis_val_table_ren_columns = mis_val_table_ren_columns[
17         mis_val_table_ren_columns.iloc[:, 1] != 0].sort_values(
18         '% від загальної кількості', ascending=False).round(1)
19
20     # Виведення деякої загальної інформації
21     print("Обрана вами таблиця має " + str(df.shape[1]) + " стовпців.\n"
22           " " + str(mis_val_table_ren_columns.shape[0]) +
23           " стовпців є пропущені значення.")
24
25     # Повернення таблиці з інформацією про пропущені значення
26     return mis_val_table_ren_columns
27
28 missing_values_data1 = missing_values_table(df)
29 print(missing_values_data1.head(30))
30

```

Обрана вами таблиця має 7 стовпців.

У 7 стовпцях є пропущені значення.

	пропущені значення	% від загальної кількості
Open	1243608	25.6
High	1243608	25.6
Low	1243608	25.6
Close	1243608	25.6
Volume_(BTC)	1243608	25.6
Volume_(Currency)	1243608	25.6
Weighted_Price	1243608	25.6

Статистика (Таблиця 3.1) описує розподіл даних для кожного стовпця у DataFrame. Для кожного стовпця вказано середнє (mean), стандартне відхилення (std), мінімальне (min), медіанне (50%), 25-й та 75-й перцентилі та максимальне (max) значення. Ці статистичні дані допомагають зрозуміти характеристики даних в кожному стовпці:

1. Open, High, Low, Close, Weighted_Price: Середні ціни відкриття, максимуму, мінімуму, закриття та вагової ціни складають приблизно 4591.96, 4595.21, 4588.55, 4591.95 та 4591.89 відповідно. Максимальні ціни досягають високого рівня біля 61763.56.
2. Volume_(BTC): Середній обсяг торгів складає 9.54 BTC з максимальним обсягом торгів 5853.85 BTC.
3. Volume_(Currency): Середній обсяг торгів у валюті становить приблизно 31433.56 з максимальним в значенні 13900670.

Ці дані вказують на значний розкид у цінах та обсязі торгів Bitcoin протягом вказаного періоду. Це може бути пов'язано з волатильністю криптовалюти.

Таблиця 3.1. Описова статистика

	Open	High	Low	Close	Volume_(BTC)	Volume_(Currency)	Weighted_Price
count	4.857377e+06	4.857377e+06	4.857377e+06	4.857377e+06	4.857377e+06	4.857377e+06	4.857377e+06
mean	4.591960e+03	4.595217e+03	4.588553e+03	4.591953e+03	9.535767e+00	3.143356e+04	4.591894e+03
std	8.157752e+03	8.164283e+03	8.151038e+03	8.157843e+03	2.979570e+01	1.321901e+05	8.157532e+03
min	3.800000e+00	3.800000e+00	1.500000e+00	1.500000e+00	0.000000e+00	0.000000e+00	3.800000e+00
25%	2.450500e+02	2.451500e+02	2.449900e+02	2.450400e+02	5.000000e-01	1.721672e+02	2.450430e+02
50%	6.996000e+02	6.999800e+02	6.991100e+02	6.995850e+02	2.213777e+00	1.425636e+03	6.995569e+02
75%	7.235270e+03	7.239690e+03	7.230770e+03	7.235570e+03	8.006946e+00	1.414915e+04	7.235224e+03
max	6.176356e+04	6.178183e+04	6.167355e+04	6.178180e+04	5.853852e+03	1.390067e+07	6.171621e+04

Візуалізуємо (Рис.3.2) розподіл числових значень у вашому DataFrame. Кожний графік представляє розподіл значень окремого числового стовпця. Для кожного числового стовпця створюється графік розподілу, який показує, як часто зустрічається кожне значення або діапазон значень. Це може допомогти виявити аномалії, виявити скошеність даних та визначити, чи є дані нормально розподілені. Розподіл відображає результати таблиці 1.1.

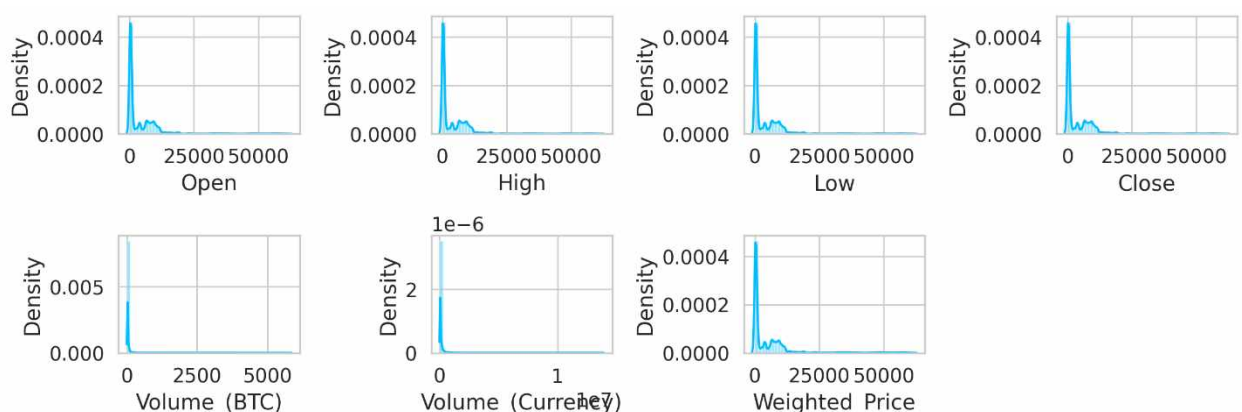


Рисунок 3.2 - Розподіл числових значень

Далі побудуємо графік розсіювання, який відображає відношення між "Low" (низькою ціною) і "High" (високою ціною) в DataFrame df1. Графік

розсіювання допомагає виявити будь-які залежності між двома змінними. На осі x відображається низька ціна, а на осі y - висока ціна. Це може допомогти розумінню, чи є відносна кореляція між цими двома параметрами.

З діаграми (Рис.3.3) розсіювання видно пряму залежність між високою та низькою ціною, це означає, що коли низька ціна зростає, висока ціна також зростає, і навпаки. Це вказує на сильну позитивну кореляцію між цими двома змінними.

Це зрозуміло, оскільки в межах одного торгового інтервалу (у даному випадку 1 хвилина), висока та низька ціни, як правило, будуть близькі одна до одної. Великі відхилення можуть відбутися лише у разі значних коливань ціни.

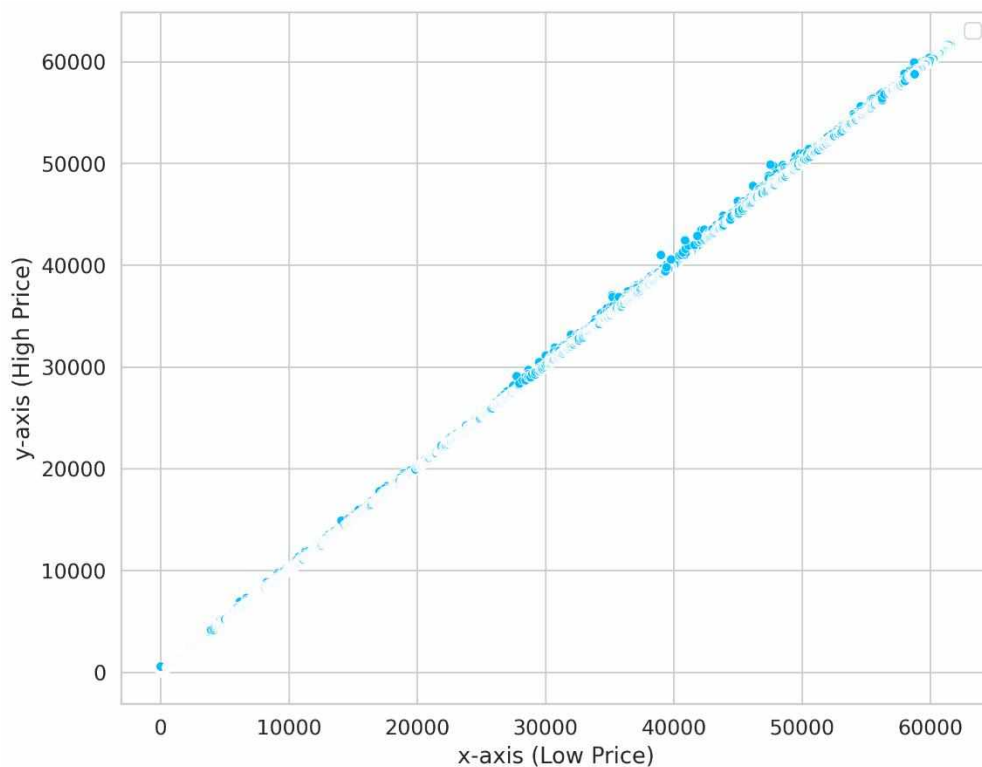


Рисунок 3.3 - Графік розсіювання між "Low" (низькою ціною) і "High" (високою ціною)

Діаграма (Рис.3.4) розсіювання показує відношення між обсягом торгів в BTC та обсягом торгів в валюті. Розподіл відображається у вигляді логарифмічної функції, це може означати, що збільшення обсягу торгів в BTC відповідає експоненціальному зростанню обсягу торгів у валюті. Таке

відношення може свідчити про те, що навіть невеликі зміни в обсязі торгів BTC можуть мати значний вплив на обсяг торгів у валюті. Це може бути пов'язано з волатильністю Bitcoin або з особливостями ринку криптовалют.

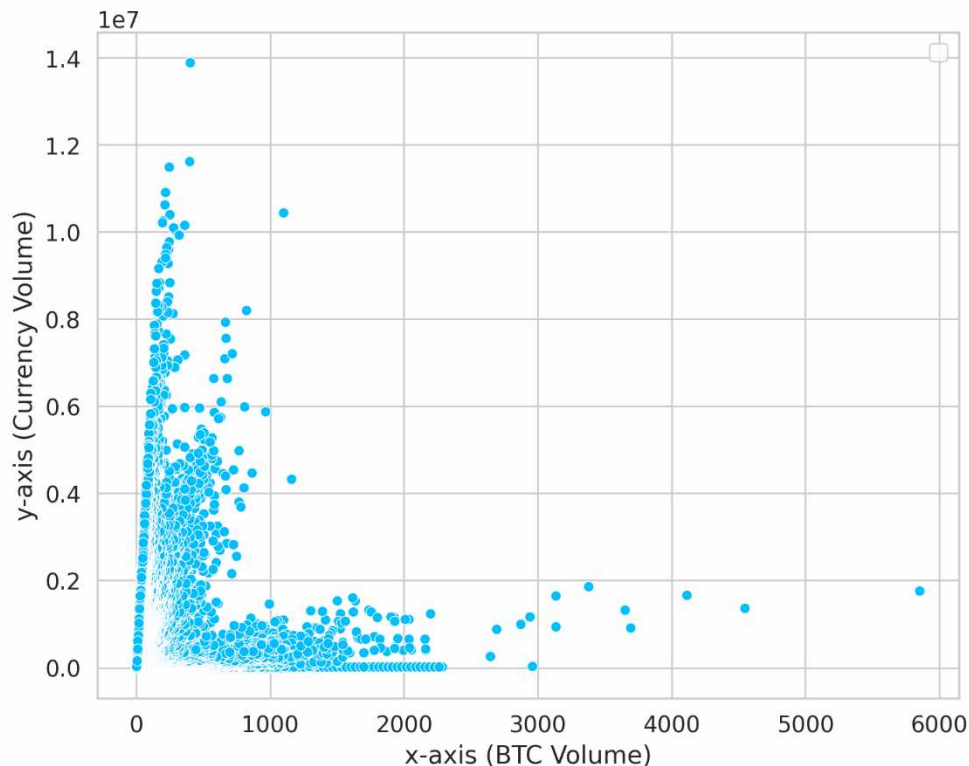


Рисунок 3.4 - Графік розсіювання між між обсягом торгів в BTC та обсягом торгів в валюті

Матриця (Рис.3.5) кореляції показує, наскільки змінні взаємно пов'язані між собою. "Open", "High", "Low", "Close" і "Weighted_Price" мають майже ідеальну позитивну кореляцію (близько до 1). Це означає, що коли одне з цих значень зростає, інші також зростають, і навпаки. Всі інші значення показують слабку міру залежності.

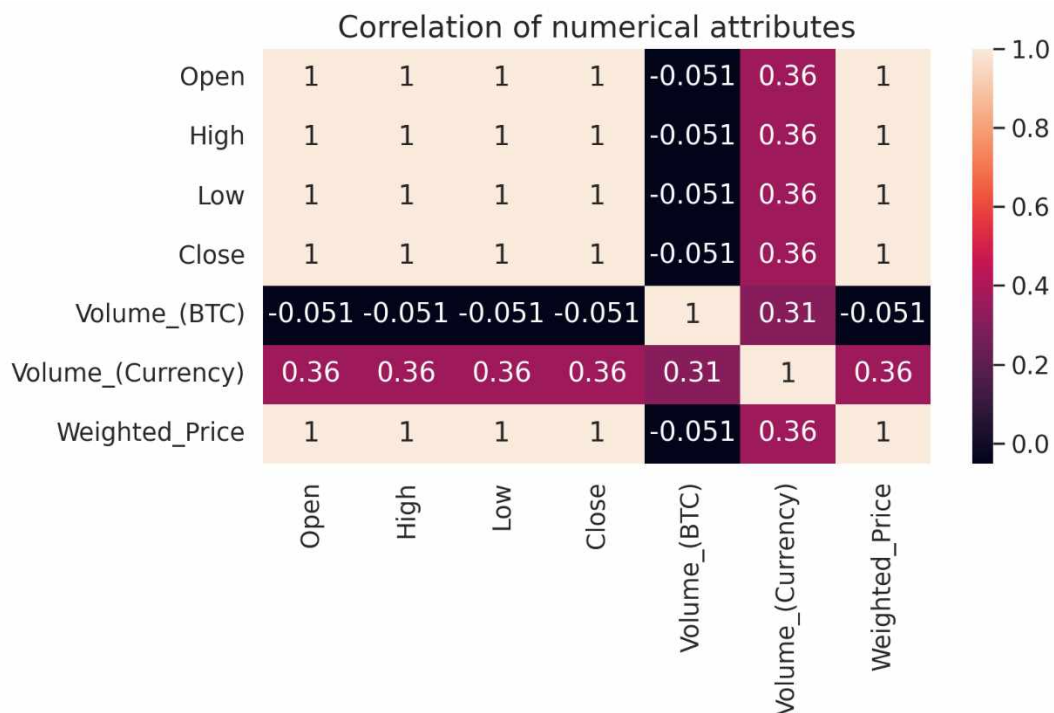


Рисунок 3.5 - Кореляція числових атрибутів

Виконаємо сезонний розклад часового ряду `df_month.Weighted_Price` і виводить результат на графіку. Крім того, вона виконує тест Діккі-Фуллера для перевірки стаціонарності ряду. Опис результату коду:

- На графіку відображаються чотири підграфіки: оригінальний часовий ряд, трендова складова, сезонна складова та залишки.
- Оригінальний часовий ряд показує зміну значень цін з часом.
- Трендова складова представляє довгострокову зміну цін.
- Сезонна складова відображає періодичність або сезонні коливання в ряді.
- Залишки представляються як нерегулярні відхилення від тренду та сезонної складової.
- Результат тесту Діккі-Фуллера виводиться як р-значення (p-value). Це значення використовується для перевірки гіпотези про наявність стаціонарності в ряді. Нижче графіку виводиться це значення.

В цілому, дана частина коду допомагає визначити наявність трендів, сезонності та стаціонарності в часовому ряді `df_month.Weighted_Price` і візуально оцінити їх характеристики.

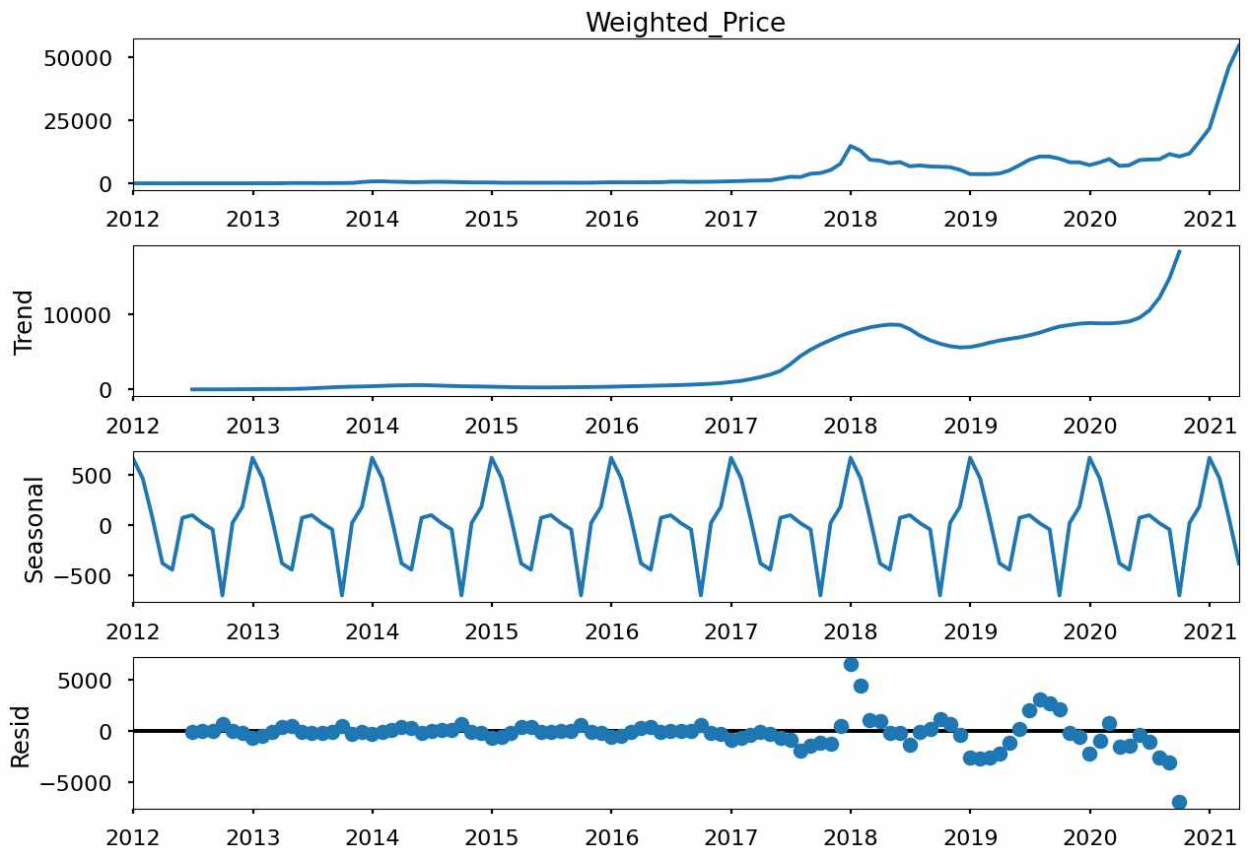


Рисунок 3.6 - Сезонний розклад часового ряду

Результат тесту Діккі-Фуллера, який виводиться у вашому випадку, є р-значенням (p-value) і дорівнює 0.998863. Тест Діккі-Фуллера використовується для перевірки гіпотези про наявність стаціонарності в часовому ряді. Зазвичай, для прийняття рішення про стаціонарність часового ряду, порогове значення значення p становить 0.05 або менше. Якщо р-значення менше порогового значення, то гіпотеза про наявність стаціонарності приймається, в іншому випадку - відхиляється. У нашому випадку, р-значення (0.998863) є дуже великим, що свідчить про те, що немає достатніх доказів відхилення від нульової гіпотези про наявність нестационарності. Це означає, що часовий ряд, ймовірно, не є стаціонарним і має тренд або сезонні

коливання. Важливо враховувати цей результат при подальшому аналізі та моделюванні вашого часового ряду.

Після застосування перетворення Вох-Сох до часового ряду `df_month.Weighted_Price`, виконується тест Діккі-Фуллера ще раз для оцінки стаціонарності ряду. Результат тесту Діккі-Фуллера, який виводиться, має те саме р-значення (p-value) як і раніше: $p=0.998863$. Це означає, що навіть після застосування перетворення Вох-Сох до ряду, немає достатніх доказів для відхилення від нульової гіпотези про наявність нестационарності. Це може вказувати на те, що перетворення Вох-Сох не вирішило проблему нестационарності ряду. Варто розглянути інші методи або підходи для досягнення стаціонарності в даному ряді.

Після виконання сезонної диференціації за допомогою зсуву на 12 місяців (`df_month['prices_box_diff'] = df_month.Weighted_Price_box - df_month.Weighted_Price_box.shift(12)`), ви виконуєте тест Діккі-Фуллера для оцінки стаціонарності нового ряду `df_month.prices_box_diff[12:]`, який враховує зсув на 12 місяців. Результат тесту Діккі-Фуллера, який виводиться, має р-значення (p-value) рівне 0.444282. Це значення більше за стандартний поріг значущості 0.05. Отже, на даному етапі немає достатніх доказів для відхилення від нульової гіпотези про наявність нестационарності. Це означає, що навіть після сезонної диференціації ряду, не вдалося досягти стаціонарності. Є можливість спробувати інші методи або підходи для отримання стаціонарного ряду.

Отримане значення р-значення (0.000024) свідчить про те, що після регулярної диференціації вдалося досягти стаціонарності в ряду `df_month.prices_box_diff2`. Це означає, що тренд та сезонні коливання були ефективно видалені, і ряд тепер може використовуватися для подальшого аналізу та моделювання.

На графіку показано STL-декомпозицію нового ряду `df_month.prices_box_diff2[13:]`, яка відображає тренд, сезонність та залишки.

					ДП.КН. 9675613.064.ПЗ	Арк.
Зм.	Арк.	№ докум.	Підпис	Дата		49

З тренду та сезонної складової були видалені, і залишки вже не мають помітних трендів чи сезонності.

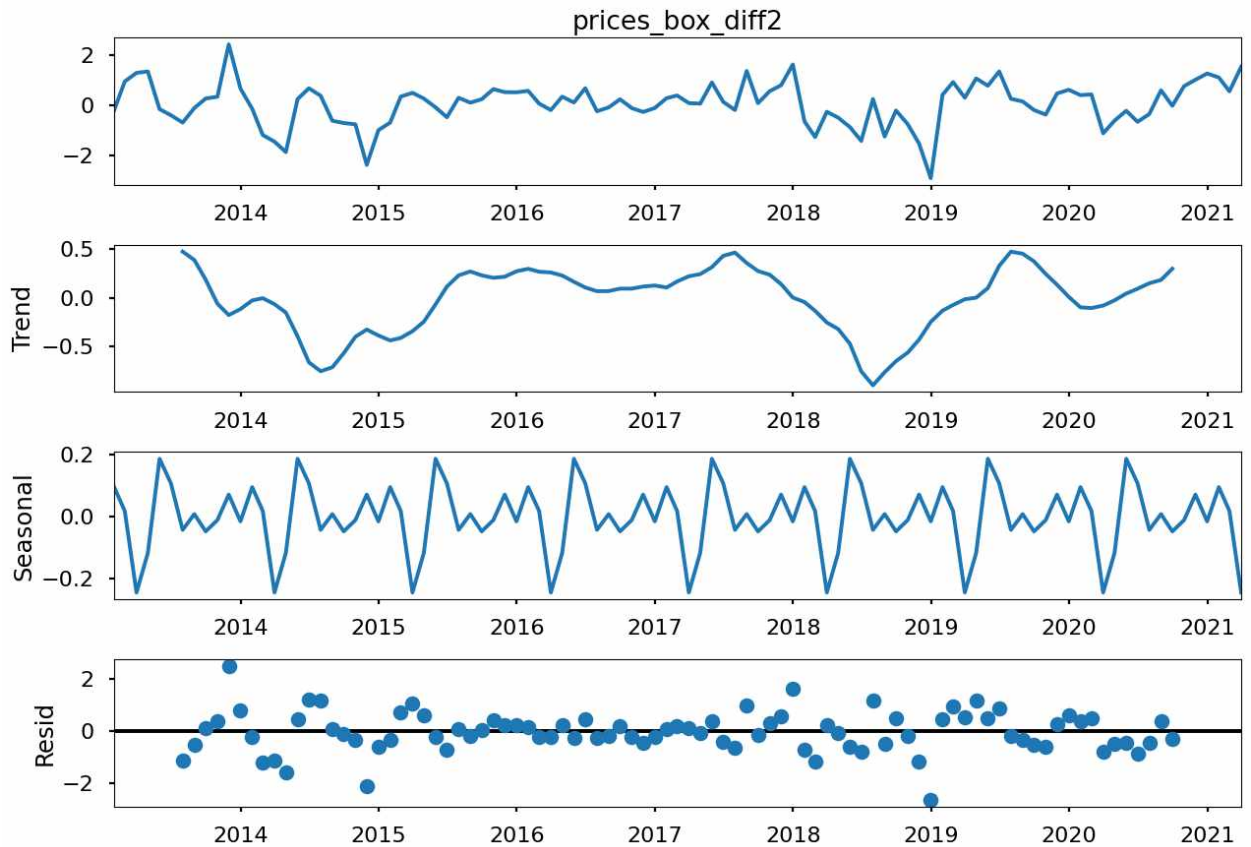


Рисунок 3.7 - STL-декомпозиція нового ряду

Отже, після регулярної диференціації вдалося отримати стаціонарний ряд, що підтверджується як результатом тесту Діккі-Фуллера, так і графічним відображенням STL-декомпозиції.

3.2 Моделювання на значеннях ціни біткойна

Аналізуючи результати рисунку 3.8, можна зробити наступні спостереження:

- Автокореляція (Autocorrelation): Значення автокореляції показують залежність між значеннями ряду на різних лагах. Наприклад, на лагу 1 автокореляція складає 0.473002, що свідчить про помірну позитивну

залежність між попереднім значенням та поточним значенням ряду. Значення автокореляції понад порогове значення можуть вказувати на наявність сезонності в ряді.

- Часткова автокореляція (Partial Autocorrelation): Значення часткової автокореляції відображають пряму залежність між значеннями ряду на певному лазі, усунувши вплив інших лагів. Наприклад, на лагу 1 часткова автокореляція складає 0.477829, що свідчить про пряму залежність між першим і другим лагами. Значення часткової автокореляції, які виходять за межі довірчого інтервалу, можуть вказувати на наявність сезонної складової.

Загалом, глядячи на автокореляцію та часткову автокореляцію на різних лагах, можна виявити можливі сезонні залежності та добре підходящі моделі ARMA або ARIMA для моделювання ряду. Значення автокореляції та часткової автокореляції, що виходять за межі довірчого інтервалу, можуть бути ознакою значимих залежностей та підказувати про використання певних лагів у моделі.

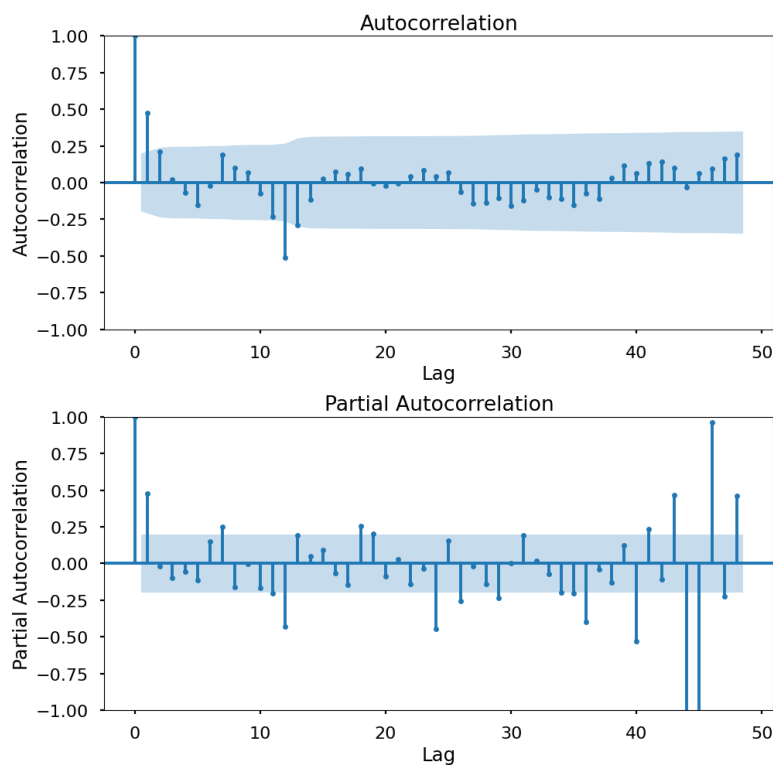


Рисунок 3.8- Результат Autocorrelation та Partial Autocorrelation

```

1 # Initial approximation of parameters
2 Qs = range(0, 2)
3 qs = range(0, 3)
4 Ps = range(0, 3)
5 ps = range(0, 3)
6 D=1
7 d=1
8 parameters = product(ps, qs, Ps, Qs)
9 parameters_list = list(parameters)
10 len(parameters_list)
11
12 # Model Selection
13 results = []
14 best_aic = float("inf")
15 warnings.filterwarnings('ignore')
16 for param in parameters_list:
17     try:
18         model=sm.tsa.statespace.SARIMAX(df_month.Weighted_Price_box, order=(param[0], d, param[1]),
19                                         seasonal_order=(param[2], D, param[3], 12)).fit(dispatch=-1)
20     except ValueError:
21         print('wrong parameters:', param)
22         continue
23     aic = model.aic
24     if aic < best_aic:
25         best_model = model
26         best_aic = aic
27         best_param = param
28     results.append([param, model.aic])

```

Результати коду включають оцінку параметрів моделі SARIMA для вибору найкращої моделі на основі інформаційного критерію Акаїке (AIC).

Основні кроки включають:

1. Визначення діапазонів значень параметрів для обробки: Qs, qs, Ps, ps.
2. Створення всіх можливих комбінацій параметрів за допомогою функції product з модуля itertools.
3. Ініціалізація списку results для зберігання результатів.
4. Перебірка кожної комбінації параметрів:
 - Спроба побудувати модель SARIMAX з використанням поточних параметрів.
 - Обчислення значення AIC для моделі.
 - Порівняння AIC з найкращим значенням AIC.
 - Збереження параметрів та AIC для поточної моделі.
5. Виведення найкращої моделі (моделі з найменшим значенням AIC) та її параметрів.

Цей процес допомагає вибрати найкращу модель SARIMA на основі найменшого значення AIC, що свідчить про найкращу прогностичну здатність

					ДП.КН. 9675613.064.ПЗ	Арк.
Зм.	Арк.	№ докум.	Підпис	Дата		52

моделі. Загальна кількість комбінацій параметрів для перебору може бути отримана зі списку `parameters_list`.

Найкращою моделлю (Рис.3.9), яка має найменше значення AIC (173.616), є SARIMAX(1, 1, 0)x(0, 1, [1], 12).

Проаналізувавши результат, можемо побачити, що коефіцієнт `ar.L1` дорівнює 0.4253, що свідчить про наявність авторегресійного ефекту на попередньому шагу. Коефіцієнт `ma.S.L12` дорівнює -0.9947 і не є значимим ($P > |z| = 0.840$), що може вказувати на відсутність сезонної складової в моделі. Значення `sigma2` вказує на оцінку дисперсії шумової складової моделі.

Тест Лjung-Бокса (Ljung-Box) демонструє значення 0.02, що свідчить про відсутність автокореляції першого порядку в залишках моделі. Тест Харке-Бера (Jarque-Bera) має значення 3.59, що показує, що розподіл залишків може не бути абсолютно нормальним.

```

parameters      aic
19 (1, 0, 0, 1) 173.616296
21 (1, 0, 1, 1) 174.766384
25 (1, 1, 0, 1) 175.547021
37 (2, 0, 0, 1) 175.554321
7  (0, 1, 0, 1) 175.589433

SARIMAX Results
=====
Dep. Variable:      Weighted_Price_box      No. Observations:      112
Model:              SARIMAX(1, 1, 0)x(0, 1, [1], 12)      Log Likelihood      -83.808
Date:               Thu, 11 May 2023      AIC      173.616
Time:               17:26:50      BIC      181.402
Sample:             12-31-2011      HQIC      176.766
                  - 03-31-2021
Covariance Type:    opg
=====
              coef      std err      z      P>|z|      [0.025      0.975]
-----
ar.L1          0.4253      0.085      5.029      0.000      0.260      0.591
ma.S.L12       -0.9947      4.917     -0.202      0.840     -10.631      8.642
sigma2          0.2439      1.185      0.206      0.837      -2.079      2.567
=====
Ljung-Box (L1) (Q):      0.02      Jarque-Bera (JB):      3.59
Prob(Q):                0.89      Prob(JB):      0.17
Heteroskedasticity (H):  1.16      Skew:      0.29
Prob(H) (two-sided):     0.68      Kurtosis:      3.73
=====

Warnings:
[1] Covariance matrix calculated using the outer product of gradients (complex-step).
```

Рисунок 3.9 - Результати вибору найкращої моделі

Загалом, отримана модель SARIMA є найкращою на основі критерію AIC і може бути використана для подальших прогнозів часового ряду. Однак, важливо провести подальший аналіз та перевірити всі припущення моделі перед її використанням.

3.3 Прогнозування ціни біткойна

Спершу виконує STL-декомпозицію резидуальних значень, отриманих з найкращої моделі SARIMA. Результати декомпозиції відображаються на графіках (Рис.3.10).

У першому підграфіку (subplot) відображаються резидуальні значення моделі, що залишилися після використання SARIMA для прогнозу часового ряду. В другому підграфіку відображається автокореляційна функція (ACF) для резидуальних значень.

Крім того, виводиться результат тесту Дікі-Фуллера (Dickey-Fuller test) для перевірки на стаціонарність резидуалів моделі.

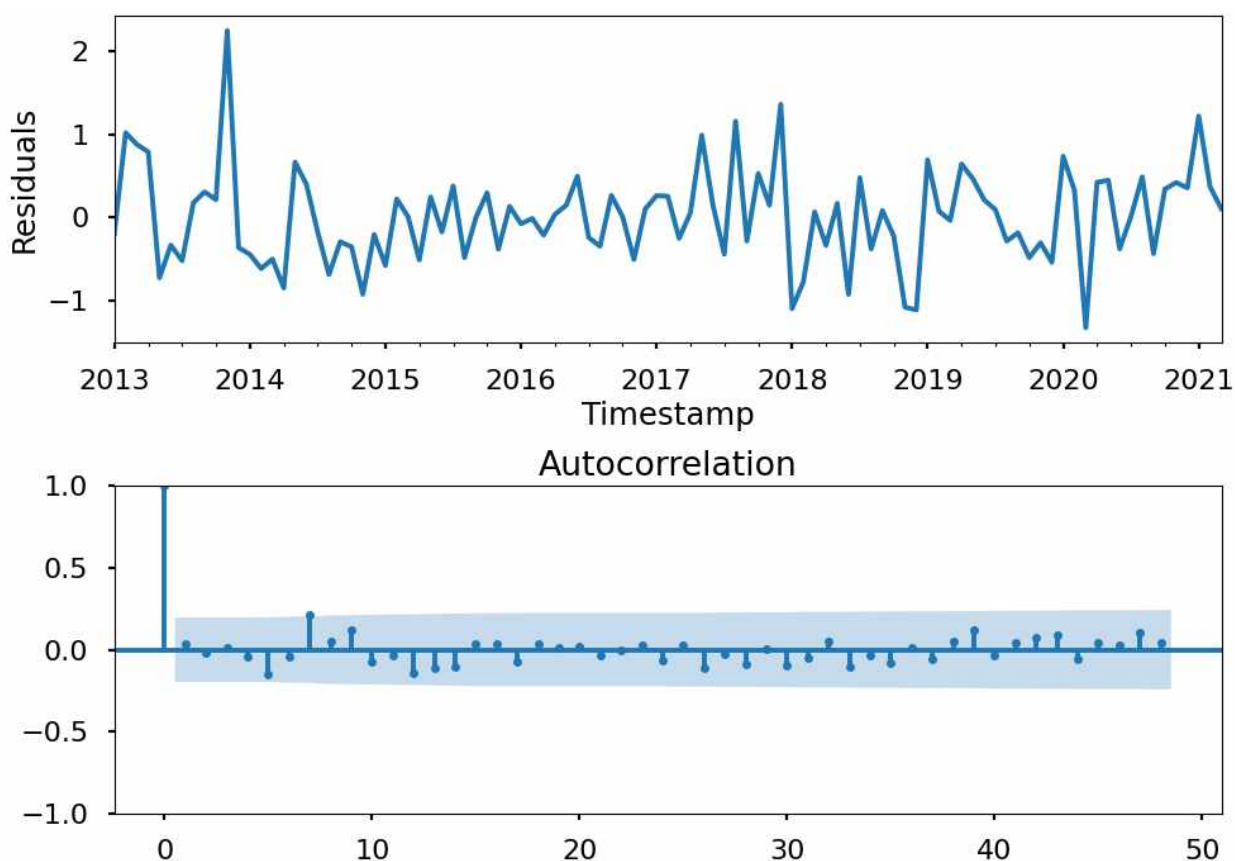


Рисунок 3.10 - Результати декомпозиції SARIMA

Результат тесту Дікі-Фуллера показує, що $p\text{-value}$ (p) дорівнює 0.000000, що набагато менше за стандартний рівень значущості 0.05. Це означає, що ми маємо достатній доказ для відкидання нульової гіпотези про нестационарність резидуалів моделі. Таким чином, результати тесту підтверджують, що резидуали моделі є стаціонарними. Це важливий фактор для правильного моделювання та прогнозування часового ряду.

```

1 # Prediction
2 df_month2 = df_month[['Weighted_Price']]
3 date_list = [datetime(2017, 6, 30), datetime(2017, 7, 31), datetime(2017, 8, 31), datetime(2017, 9, 30),
4              datetime(2017, 10, 31), datetime(2017, 11, 30), datetime(2017, 12, 31), datetime(2018, 1, 31),
5              datetime(2018, 1, 28)]
6 future = pd.DataFrame(index=date_list, columns= df_month.columns)
7 df_month2 = pd.concat([df_month2, future])
8 df_month2['forecast'] = invboxcox(best_model.predict(start=0, end=75), lambda)
9 plt.figure(figsize=(15,7))
10 df_month2.Weighted_Price.plot()
11 df_month2.forecast.plot(color='r', ls='--', label='Predicted Weighted_Price')
12 plt.legend()
13 plt.title('Bitcoin exchanges, by months')
14 plt.ylabel('mean USD')
15 plt.show()

```

В цьому коді створюється новий DataFrame з назвою `df_month2`, який містить колонку "Weighted_Price" з оригінальними значеннями цін. Далі створюється список дат `date_list`, який включає майбутні дати, для яких ми хочемо здійснити прогноз. За допомогою цього списку створюється новий DataFrame `future`, який містить пусті значення для всіх колонок, включаючи "Weighted_Price" і "forecast". Потім виконується об'єднання двох DataFrame - `df_month2` і `future`, за допомогою функції `pd.concat()`.

Далі, в `df_month2` створюється нова колонка "forecast", в яку зберігаються зворотно перетворені прогнозні значення. Це робиться за допомогою функції `invboxcox()`, використовуючи `best_model.predict()` для отримання прогнозних значень з моделі SARIMA. Параметр `start=0, end=75` вказує, що хочемо отримати прогнози для перших 76 значень (включаючи наявні дані та майбутні дати).

Нарешті, виконується візуалізація результатів. Графік показує оригінальні значення "Weighted_Price" (синій колір) і прогнозовані значення (червоний колір) на вказаних датах. Цей графік дозволяє порівняти прогнози з реальними значеннями і оцінити точність моделі.



Рисунок 3.11 - Прогноз

Прогнози показують очікувані значення середньої ціни Bitcoin для кожного місяця. Наприклад, прогнозована середня ціна Bitcoin на 31 січня 2023 року складає 189,958.42 USD, на 28 лютого 2023 року - 193,431.39 USD, а на 31 березня 2023 року - 198,815.73 USD.

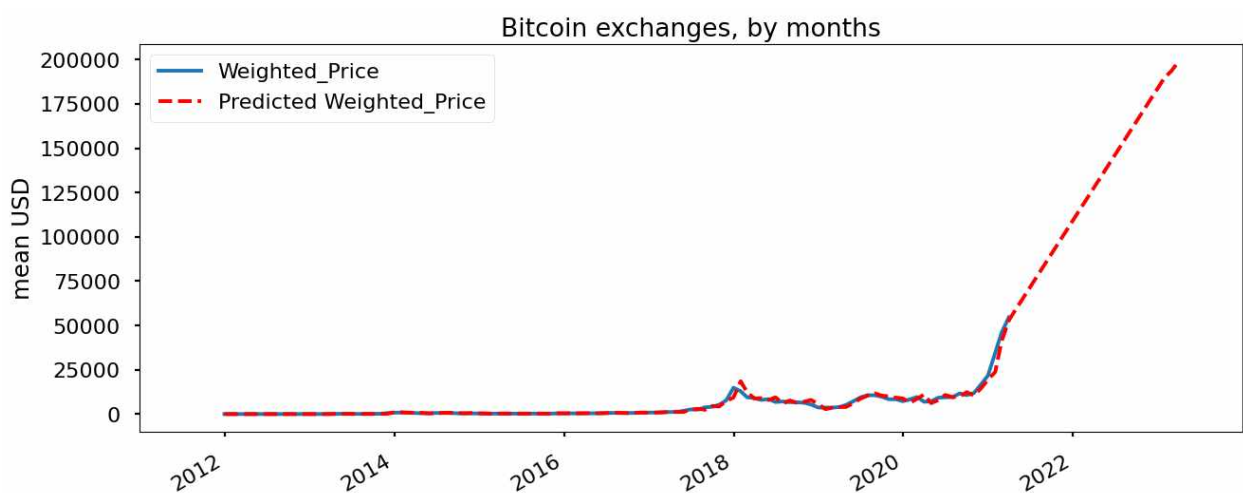


Рисунок 3.14 - Прогноз до березня 2023 року

Важливо зазначити, що це прогнози, які ґрунтуються на побудованій моделі SARIMA та попередніх спостереженнях. Прогнози можуть бути

піддані похибці, оскільки майбутні події та фактори, які можуть впливати на ціну Bitcoin, невідомі. Тому важливо розглядати ці прогнози як інструмент для розуміння можливих тенденцій, а не як точні передбачення.

Отже, для побудови моделі ARIMA було використано метод SARIMA, який дозволяє врахувати сезонність в даних. Було проведено підбір оптимальних параметрів моделі, використовуючи критерій AIC (інформаційний критерій Акаїке), і було обрано найкращу модель з найменшим значенням AIC.

Після побудови моделі ARIMA було здійснено прогнозування ціни Bitcoin на майбутні періоди. Прогнозовані значення були зворотно перетворені з використанням оберненої функції Box-Cox, яка була застосована до вихідних даних для стабілізації їх дисперсії.

Отримані прогнози були візуалізовані на графіку, який показує оригінальні дані ціни Bitcoin та прогнозовані значення. Зазначимо, що прогнози мають певну ступінь невизначеності, оскільки майбутні події та фактори, що впливають на ціну Bitcoin, не можуть бути повністю передбачені.

У цілому, модель ARIMA показалася корисним інструментом для прогнозування ціни Bitcoin. Прогнози можуть бути використані для отримання загального уявлення про можливі тенденції в ціні та допомогти в прийнятті інформованих рішень. Проте, важливо пам'ятати, що прогнози підлягають похибці і потребують постійного оновлення та коригування залежно від нових даних та розвитку ринку.

					ДП.КН. 9675613.064.ПЗ	Арк.
						57
Зм.	Арк.	№ докум.	Підпис	Дата		

ВИСНОВКИ

Прогнозування ціни біткойна є актуальним завданням з огляду на унікальні характеристики та майбутні перспективи цього цифрового активу. Зменшення винагороди за блок та наближення наступного халвінгу, а також обмежена кількість біткойнів у обігу, створюють зміни в доступності та ціні біткойна. Прогнозування ціни біткойна надає інвесторам, трейдерам та іншим учасникам ринку інформацію для прийняття обґрунтованих рішень та використання можливостей, що надає цей актив.

У контексті прогнозування ціни біткойна, ARIMA є найкращим методом на основі розглянутих рішень. Він володіє простотою в застосуванні, але при цьому досягає задовільної точності прогнозів. Хоча деякі складніші моделі можуть мати вищу точність у певних ситуаціях, вони вимагають більшої складності у підготовці даних та використанні. ARIMA забезпечує зручний і ефективний інструмент для прогнозування ціни біткойна.

З огляду на сезонну залежність та складну динаміку біткойна, SARIMA є відмінним вибором для прогнозування ціни в модулі прогнозування. Цей метод дозволяє враховувати сезонність та зміни у часовому ряді, при цьому маючи простоту у реалізації та добрі результати у прогнозуванні. SARIMA є зручним та ефективним інструментом для прогнозування ціни біткойна, здатним враховувати його особливості та забезпечувати достатню точність прогнозів.

У роботі були розроблені алгоритми, які значно сприяють прогнозуванню ціни біткойна. Алгоритм визначення стаціонарності та декомпозиції часового ряду дозволяє розкрити структуру ряду і виділити його компоненти, що поліпшує розуміння його поведінки і використання відповідних прогнозуючих моделей.

Розроблений алгоритм прогнозування з використанням методу SARIMA включає підготовку часового ряду, вибір оптимальних параметрів, побудову моделі, прогнозування та оцінку точності. Цей алгоритм спеціально

					ДП.КН. 9675613.064.ПЗ	Арк.
Зм.	Арк.	№ докум.	Підпис	Дата		58

адаптований для прогнозування ціни біткойна, дозволяючи враховувати як сезонні, так і несезонні фактори в часовому ряді.

Отриманий алгоритм прогнозування ціни біткойна на основі методу SARIMA є потужним і ефективним інструментом, який враховує як сезонні, так і несезонні коливання ціни. Використовуючи історичні дані, цей алгоритм дозволяє побудувати модель та зробити прогноз на майбутнє, що забезпечує цінну інформацію для учасників криптовалютного ринку.

Після проведення регулярної диференціації часового ряду вдалося отримати стаціонарний ряд. Це було підтверджено як результатом тесту Діккі-Фуллера, так і графічним відображенням STL-декомпозиції.

Загалом, модель SARIMA, яка була побудована, виявилася найкращою відповідно до критерію AIC і може бути використана для подальшого прогнозування часового ряду. Однак, перед використанням моделі важливо провести додатковий аналіз та перевірити всі припущення, що лежать в її основі.

Для побудови моделі ARIMA був використаний метод SARIMA, який дозволяє враховувати сезонність в даних. Було проведено підбір оптимальних параметрів моделі за допомогою критерію AIC, і була обрана модель з найменшим значенням AIC.

Після побудови моделі ARIMA було здійснено прогнозування ціни біткойна на майбутні періоди. Прогнозовані значення були зворотно перетворені за допомогою оберненої функції Вох-Сох, яка була застосована до вихідних даних для стабілізації їх дисперсії.

Отримані прогнози були візуалізовані на графіку, який показує оригінальні дані ціни біткойна та прогнозовані значення. Важливо зазначити, що прогнози мають певну ступінь невизначеності, оскільки майбутні події та фактори, що впливають на ціну біткойна, не можуть бути повністю передбаченими.

У цілому, модель ARIMA виявилася корисним інструментом для прогнозування ціни біткойна. Прогнози можуть бути використані для

					ДП.КН. 9675613.064.ПЗ	Арк.
						59
Зм.	Арк.	№ докум.	Підпис	Дата		

отримання загального уявлення про можливі тенденції в ціні та допомогти приймати інформовані рішення. Проте важливо пам'ятати, що прогнози підлягають похибці і потребують постійного оновлення та коригування залежно від нової інформації та розвитку ринку.

У висновку, розроблений алгоритм прогнозування на основі моделі SARIMA є потужним і ефективним інструментом для прогнозування цінових змін на криптовалютному ринку, зокрема для ціни біткойна. Використання SARIMA дозволяє враховувати як сезонні, так і несезонні коливання ціни та залежності між попередніми значеннями. Цей підхід дозволяє отримати більш точні та надійні прогнози.

					ДП.КН. 9675613.064.ПЗ	Арк.
						60
Зм.	Арк.	№ докум.	Підпис	Дата		

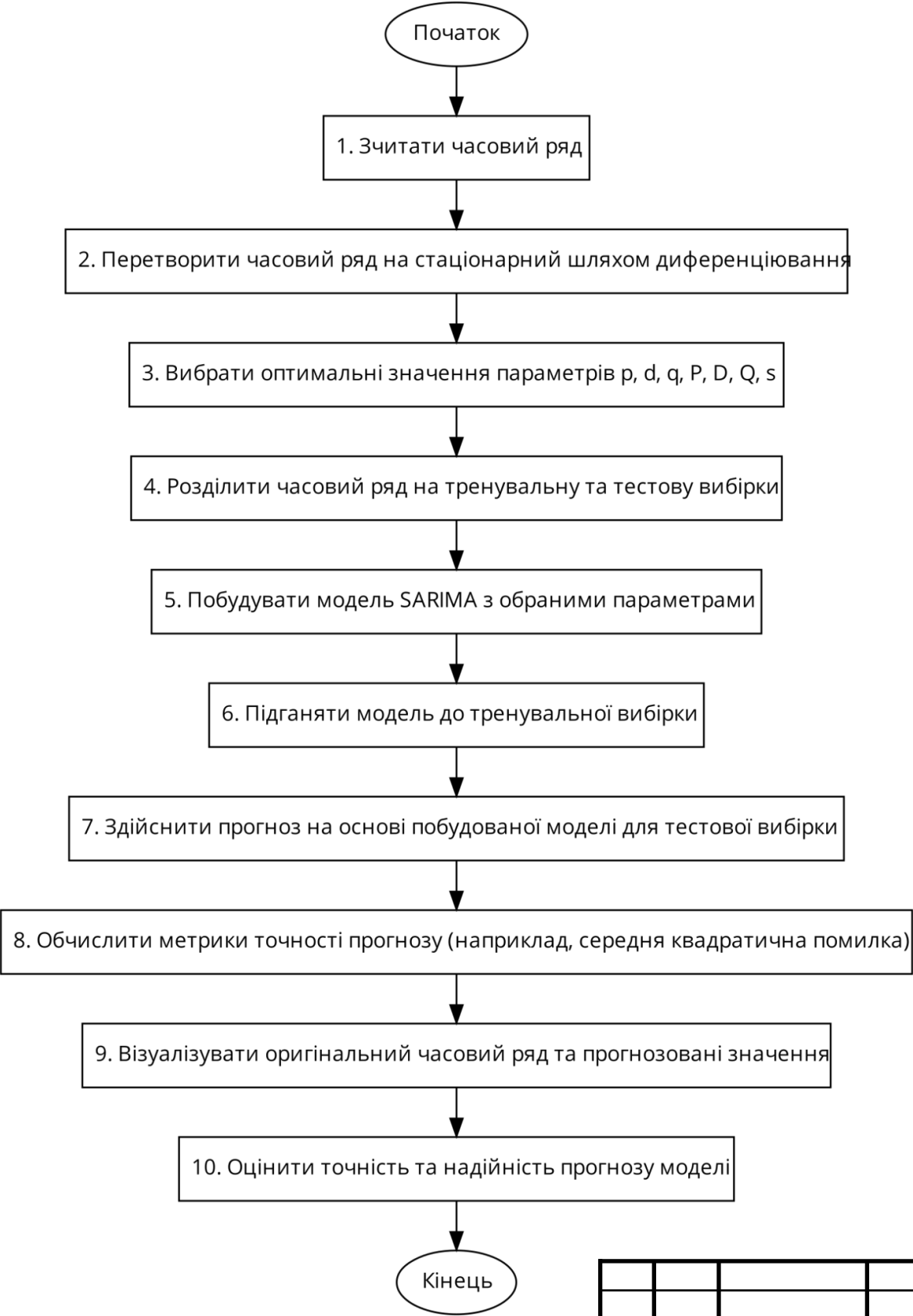
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Xu, W., & Cooper, A. (2017). Digital spillover: measuring the true impact of the digital economy. Huawei and Oxford Economics. Retrieved from <https://www.huawei.com/minisite/gci/en/digital-spillover/files/gci-digital-spillover.pdf>
2. Oliver, J. (2013). Mastering blockchain distributed ledger technology, and smart contracts explained-packt publishing, vol 53. Packt Publishing Ltd. <https://doi.org/10.1017/CBO9781107415324.004>
3. Unal, D., Hammoudeh, M., & Kiraz, M. S. (2020). Policy specification and verification for blockchain and smart contracts in 5G networks. ICT Express, 6(1), 43–47. <https://doi.org/10.1016/j.ict.2019.07.002>
4. Morris, D. Z. (2017). Bitcoin hits a new record high, but stops short of \$20,000. Fortune. Retrieved from <https://fortune.com/2017/12/17/bitcoinrecord-high-short-of-20000/>
5. Shukla, S., & Dave, S. (2020). Bitcoin beats coronavirus blues. The Economic Times. Retrieved from <https://economictimes.indiatimes.com/markets/stocks/news/bitcoinbeats-coronavirus-blues/articleshow/75049718.cms>
6. Liu, L. (2019). Are Bitcon returns predictable?: Evidence from technical indicators. Physica A: Statistical Mechanics and its Applications, 533, 121950. <https://doi.org/10.1016/j.physa.2019.121950>
7. Huang, J. Z., Huang, W., & Ni, J. (2018). Predicting Bitcoin returns using high-dimensional technical indicators. Journal of Finance and Data Science, 5(3), 140–155. <https://doi.org/10.1016/j.jfds.2018.10.001>
8. Chen, Z., Li, C., & Sun, W. (2020). Bitcoin price prediction using machine learning: an approach to sample dimension engineering. Journal of Computational and Applied Mathematics, 365, 112395.

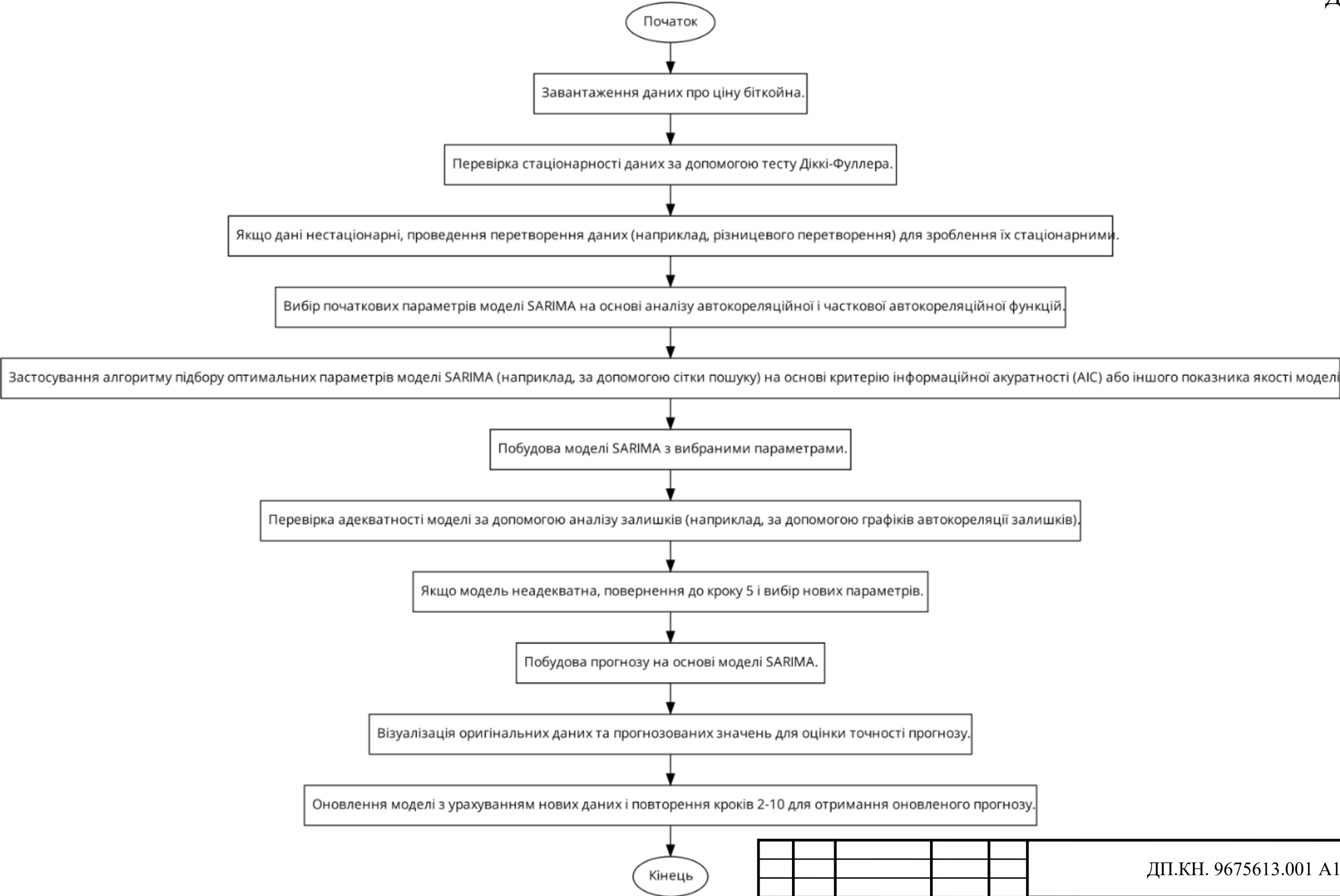
9. Adcock, R., & Gradojevic, N. (2019). Non-fundamental, non-parametric Bitcoin forecasting. *Physica A: Statistical Mechanics and its Applications*, 531, 121727. <https://doi.org/10.1016/j.physa.2019.121727>
10. Mallqui, D. C., & Fernandes, R. A. (2019). Predicting the direction, maximum, minimum and closing prices of daily Bitcoin exchange rate using machine learning techniques. *Applied Soft Computing*, 75, 596–606.
11. Urquhart, A. (2018). What causes the attention of Bitcoin? *Economic Letters*, 166, 40–44.
12. Philippas, D., Rjiba, H., Guesmi, K., & Goutte, S. (2019). Media attention and Bitcoin prices. *Finance Research Letters*, 30, 37–43.
13. Abraham, J., Higdon, D., Nelson, J., & Ibarra, J. (2018). Cryptocurrency price prediction using tweet volumes and sentiment analysis. *SMU Data Science Review*, 1(3), 1.
14. Huang, J. Z., Huang, W., & Ni, J. (2019). Predicting Bitcoin returns using high-dimensional technical indicators. *Journal of Finance and Data Science*, 5(3), 140–155.
15. Shen, D., Urquhart, A., & Wang, P. (2019). Does twitter predict Bitcoin? *Economic Letters*, 174, 118–122.
16. Kim, Y. B., Kim, J. G., Kim, W., Im, J. H., Kim, T. H., Kang, S. J., & Kim, C. H. (2016). Predicting fluctuations in cryptocurrency transactions based on user comments and replies. *PloS one*, 11(8), e0161197. <https://doi.org/10.1371/journal.pone.0161197>
17. Hyndman, R. J., & Athanasopoulos, G. (2018). *Forecasting: principles and practice* (2nd ed.). OTexts. Retrieved from <https://otexts.com/fpp2/>
18. Bakar, N. A., & Rosbi, S. (2017). Autoregressive integrated moving average (ARIMA) model for forecasting cryptocurrency exchange rate in high volatility environment: A new insight of bitcoin transaction. *International Journal of Advanced Engineering Research and Science*, 4(11), 130–137. <https://doi.org/10.22161/ijaers.4.11.20>

19. Garcia, R. C., Contreras, J., Van Akkeren, M., & Garcia, J. B. C. (2005). A GARCH forecasting model to predict day-ahead electricity prices. *IEEE Transactions on Power Systems*, 20(2), 867–874.
20. Munim, Z. H., Shakil, M. H., & Alon, I. (2019). Next-day Bitcoin price forecast. *Journal of Risk and Financial Management*, 12(2), 103. <https://doi.org/10.3390/jrfm12020103>
21. Jang, H., & Lee, J. (2017). An empirical study on modeling and prediction of Bitcoin prices with Bayesian neural networks based on blockchain information. *IEEE Access*, 6, 5427–5437.
22. Shah, D., & Zhang, K. (2014). Bayesian regression and Bitcoin. In 2014 52nd Annual Allerton Conference on Communication, Control, and Computing (Allerton). IEEE. pp. 409–414.
23. Chen, G. H., Nikolov, S., & Shah, D. (2013). A latent source model for nonparametric time series classification. In *Advances in Neural Information Processing Systems*, pp. 1088–1096.
24. Greaves, A., & Au, B. (2015). Using the Bitcoin transaction graph to predict the price of Bitcoin. <https://doi.org/10.1109/CEC.2010.5586007>
25. McNally, S., Roche, J., & Caton, S. (2018). Predicting the price of Bitcoin using machine learning. In *Proceedings of the 26th Euromicro International Conference on Parallel, Distributed, and Network-Based Processing, PDP 2018*, pp. 339–343. <https://doi.org/10.1109/PDP2018.2018.00060>
26. Liu, F. T., Ting, K. M., & Zhou, Z. H. (2012). Isolation-based anomaly detection. *ACM Transactions on Knowledge Discovery from Data*, 6(1), 1–39. <https://doi.org/10.1145/2133360.2133363>
27. Cybenko, G. (1989). Approximation by superpositions of a sigmoidal function. *Mathematics of Control, Signals and Systems*, 2(4), 303–314. <https://doi.org/10.1007/BF02551274>
28. Kingma, D. P., & Ba, J. (2014). Adam: a method for stochastic optimization. *arXiv preprint. arXiv:1412.6980*.

29. Hanin, B., & Sellke, M. (2017). Approximating continuous functions by ReLU nets of minimal width. arXiv preprint. arXiv:1710.11278.
30. Patel, H., Rajput, D. S., Reddy, G. T., Iwendi, C., Bashir, A. K., & Jo, O. (2020). A review on classification of imbalanced data for wireless sensor networks. *International Journal of Distributed Sensor Networks*, 16(4), 1550147720916404. <https://doi.org/10.1177/1550147720916404>
31. Mandrekar, J. N. (2010). Receiver operating characteristic curve in diagnostic test assessment. *Journal of Thoracic Oncology*, 5(9), 1315–1316.
32. Bitcoin Historical Data. Kaggle: Your Machine Learning and Data Science Community. URL: <https://www.kaggle.com/datasets/mczielinski/bitcoin-historical-data>.
33. V. Krylov et al., "Multiple Regression Method for Analyzing the Tourist Demand Considering the Influence Factors," 2019 10th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS), Metz, France, 2019, pp. 974-979, doi: 10.1109/IDAACS.2019.8924461.
34. Koziuk, V., & Lipyanina-Goncharenko, H. (2021, September). Intelligent Method of Predicting the Discount Rate Trend. In 2021 11th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS) (Vol. 2, pp. 1132-1140). IEEE.
35. Komar, M., Savenko, O., Sachenko, A., Lendiuk, T., Lipianina-Honcharenko, K., Hladiy, G., & Vasylykiv, N. (2022). Evaluation the Efficiency of Information Technology of Big Data Intelligence Analysis and Processing.
36. Комар М.П., Саченко А.О., Васильків Н.М., Гладій Г.М., Коваль В.С. Методичні рекомендації до виконання кваліфікаційної роботи з освітньо-професійної програми «Комп'ютерні науки» спеціальності 122 «Комп'ютерні науки» за першим (бакалаврським) рівнем вищої освіти. – Тернопіль: ЗУНУ, 2021. – 56 с.



					ДП.КН. 9675613.001 А1										
					Алгоритм прогнозування методом SARIMA	Літера			Маса		Масштаб				
Змн.	Арк.	№ докум.	Підпис	Дата											
Розроб.		Кулинич А. Ю.													
Перевір.		Дорош В.І.													
Консультант															
Т. Контр.															
Н. Контр.		Дорош В.І.									ЗУНУ.ФКІТ.КН-42				
Затверд.		Комар М.П.													



					ДП.КН. 9675613.001 А1						
					Алгоритм прогнозування ціни біткойна на основі методу SARIMA	Літера			Маса	Масштаб	
Змн.	Арк.	№ докум.	Підпис	Дата							
Розроб.		Кулинич А. Ю.									
Перевір.		Дорош В.І.									
Консультант											
Т. Контр.						Аркуш 1			Аркушів 1		
Н. Контр.		Дорош В.І.				ЗУНУ.ФКІТ.КН-42					
Затверд.		Комар М.П.									