

**КИЇВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ
ІМЕНІ ТАРАСА ШЕВЧЕНКА
Економічний факультет
Кафедра економічної кібернетики**

КВАЛІФІКАЦІЙНА РОБОТА БАКАЛАВРА

«Аналіз впливу соціальних мереж на курс криптовалют»

студентки 4 курсу
спеціальності 051 «Економіка»
ОПП «Економічна кібернетика»
денної форми навчання
Пасик Світлани Михайлівни

Науковий керівник

кандидат фізико-математичних наук,
доцент кафедри економічної
кібернетики
Банна Оксана Леонідівна

Засвідчую, що у цій кваліфікаційній
роботі немає запозичень із праць інших
авторів без відповідних посилань
Студент _____

(підпис)

Роботу допущено до захисту перед ЕК
рішенням кафедри економічної кібернетики
від 12 червня 2023 р., протокол № 17

Завідувач кафедри:

доктор економічних наук, професор
Ляшенко Олена Ігорівна

(підпис)

ЗМІСТ

ВСТУП.....	3
РОЗДІЛ I. КРИПТОВАЛЮТИ ТА СОЦІАЛЬНІ МЕРЕЖІ.....	6
1.1. Поняття, призначення криптовалют	6
1.2. Процес купівлі/продажу та зберігання криптовалюти	8
1.3. Ринок криптовалют та вплив соціальних мереж на нього	13
РОЗДІЛ II. МЕТОДИ ДЛЯ АНАЛІЗУ ЧАСОВИХ РЯДІВ	23
2.1. Огляд часових рядів та їх властивостей.....	23
2.2. Модель ARIMA	26
2.3. Розширена модель ARIMA (ARIMAX).....	30
2.4. Пакет Prophet	31
РОЗДІЛ III. АНАЛІЗ ВПЛИВУ СОЦІАЛЬНОЇ МЕРЕЖІ TWITTER НА BITCOIN ...	34
3.1. Підготовка даних.....	34
3.2. Візуальний аналіз	35
3.3. Аналіз за допомогою ARIMA	39
3.4. Аналіз за допомогою Prophet.....	44
ВИСНОВКИ	49
СПИСОК ДЖЕРЕЛ ТА ЛІТЕРАТУРИ	51
ДОДАТОК А	55

ВСТУП

За останні десятиліття світовий фінансовий сектор переживає інтенсивний розвиток та трансформацію, в основі яких лежать глобалізація та цифровізація економіки. Фінансові системи активно пристосовуються до нових економічних та технологічних вимог, створюючи умови для ефективного функціонування торговельних та фінансових відносин, які є невід'ємною частиною глобальних економік.

Особливу роль у цьому процесі відіграють глобалізація та цифровізація, що сприяють фінансовим інноваціям. Платіжні системи та інструменти зіштовхуються зі зростаючими регуляторними вимогами, оскільки ризики стають все більш розподіленими між світовими економіками.

Цифрові валюти, зокрема криптовалюти, стають новим засобом розвитку існуючих платіжних систем і фінансових установ. Вони пропонують нову концепцію грошей та безпечності транзакцій. Наприклад, США вже активно досліджують та планують впровадження криптовалюти в свої внутрішні платіжні системи. Відповідний законопроект, відомий як «Crypto-Currency Act of 2020», був розглянутий Конгресом США у 2019 році. Цей законопроект встановлює процедури для визнання, ліцензування та реєстрації цифрових валют як засобу платежу, а також визначає відповідальні державні органи, що регулюють та контролюють нові валюти [23].

З моменту створення Bitcoin криптовалюти значно зросли не лише з погляду капіталізації, а й кількості. 2021 рік став рекордним для криптовалютного ринку та блокчейну. У листопаді 2021 року загальна капіталізація ринку досягла піку майже в 3 трильйони доларів США. Проте всього за два місяці вона впала на 40%, знизившись більш ніж на 1,2 трильйона доларів США. Таким чином, в січні 2022 року загальна ринкова капіталізація криптовалют становила приблизно 1,8 трильйона доларів США [5]. Станом на травень 2023 року загальна капіталізація знизилася до 1,1 трильйона доларів США, що є значно менше, ніж було майже півтора роки тому [18]. Ці цифри свідчать про велику волатильність та нестабільність на ринку криптовалют, де значні зміни можуть відбуватися у короткі проміжки часу. Такі

різкий зміни на ринку підкреслюють необхідність в глибокому розумінні динаміки його ціни та факторів, що впливають на неї.

У сучасному світі соціальні мережі стали необхідною частиною нашого повсякденного життя. Вони стали потужним інструментом для комунікації та обміну інформацією між користувачами з усього світу. Соціальні мережі впливають на різні аспекти нашої діяльності, включаючи економіку та фінансові ринки. Інформація, яка поширюється через соціальні мережі, може мати значний вплив на настрої та рішення інвесторів, що впливає на курс криптовалют. У зв'язку із цим важливо дослідити, як саме соціальні мережі впливають на курс криптовалют. Ця тема зацікавила численних вчених, які досліджують взаємозв'язок між соціальними мережами та фінансовими ринками.

Наприклад, у роботі М. Савченко, А. Завидовської та В. Краєвської «Кореляційний аналіз волатильності криптовалют світу» відзначається, що вплив соціальних мереж на криптовалютний ринок є очевидним у повсякденному житті. Дослідники також вказують, що аналіз ефективності криптовалют неможливий без врахування впливу новин і соціальних мереж, а також їх взаємозалежності [2].

У дослідженні Й. Боллена, Х. Мао та С. Цзена «Twitter mood predicts the stock market» досліджується зв'язок між настроєм громадськості, виміряним на основі великомасштабного збору твітів на платформі twitter.com, та значеннями DJIA [11].

Також дослідженнями цієї та схожих тем займалися такі науковці: А. Бутеска, П. Хаєк, М. З. Абедін [20], А. Ф. Юстіндра та В. Б. Кластер [43], Т. Рао та С. Срівастава [36], Р. Джайн, Р. Нгуєн та інші [10].

Метою даної кваліфікаційної роботи є дослідження впливу соціальних мереж на ціновий ряд Bitcoin за допомогою різних моделей прогнозування, зокрема ARIMAX і Prophet.

Для досягнення поставленої мети, кваліфікаційна робота передбачає виконання наступних завдань:

- вивчення теоретичних основ цифрових валют, включаючи їх сутність, принципи функціонування і особливості;

- аналіз технології блокчейн та її роль у розвитку криптовалют, дослідження основних принципів блокчейн-технології;
- оцінка поточного стану ринку криптовалют;
- дослідження факторів, що впливають на коливання курсу;
- аналіз курсів криптовалют з використанням підходів, які базуються на аналізі історичних даних та можливому впливі соціальних мереж

Об'єктом дослідження є вплив соціальних мереж на курс криптовалют.

Предметом дослідження є теоретичні і практичні засади аналізу взаємодії соціальних мереж та курсу криптовалют.

Під час написання даної роботи використовувалися наступні **методи**: теоретичний метод, метод порівнянь та узагальнення, метод аналізу, метод комп'ютерного моделювання, економіко-математичний та аналітичний методи.

Структура роботи. Кваліфікаційна робота складається зі вступу, трьох розділів, висновків, списку джерел та літератури та додатку.

Результати проведених досліджень були представлені та обговорювались на міжнародній науково-практичній конференції «Шевченківська весна 2023. Повоєнне відновлення економіки України: проблеми та перспективи» [1].

РОЗДІЛ І

КРИПТОВАЛЮТИ ТА СОЦІАЛЬНІ МЕРЕЖІ

1.1. Поняття, призначення криптовалют

Під криптовалютою розуміється будь-яка система електронних грошей, що використовується для купівлі та продажу в Інтернеті без необхідності взаємодії з центральним банком [34]. Загалом криптовалюти – це електронні цифрові валюти, які використовують криптографічні методи для забезпечення безпеки та анонімності транзакцій. Їх створено з метою забезпечення можливості здійснення безпечних та анонімних операцій, що робить їх популярними серед користувачів, які цінують приватність та конфіденційність.

У 1960-х роках професійні криптографи розпочали дискусію про можливість створення глобальної інформаційної мережі, яка вперше отримала практичне втілення в 1980-х роках. Завдяки цій мережі з'явилася можливість обмінюватися брокерськими даними, необхідними для проведення торгівельних операцій на фондових біржах. Саме в цей час з'явилася ідея цифрових грошей, які пропонували швидкий доступ до покупки акцій, фінансових активів та ін. Основна мета цієї концепції полягала у спрощенні та прискоренні фінансових операцій на ринку.

Уперше криптовалюта була згадана в 1989 році, а через кілька років американський криптограф Девід Чаум винайшов цифрову готівку, яка покладалася на криптографію для захисту та перевірки транзакцій. Однак, лише на початку 1990-х років почали розроблятися криптографічні протоколи та програмні забезпечення, які дозволили б створити справді децентралізовану цифрову валюту [39].

До криптовалют було багато спроб створити кілька варіантів цифрових активів. Однак, основною проблемою більшості спроб була проблема подвійних витрат. Цифровий актив повинен бути використаним тільки один раз, щоб уникнути його копіювання та підроблення.

Більш як за 10 років до появи криптовалют, інженер Вей Дай вперше запропонував концепцію «В-money» у своїй публікації 1998 року, де обговорював ідею цифрової валюти, яку можна було б надсилати за допомогою групи

неідентифікованих цифрових псевдонімів [19]. Також у тому ж році, блокчейн-піонер Нік Сзабо спробував створити децентралізовану цифрову валюту, яку він назвав Віт Голд [33]. Ідея Сзабо була спровокована недоліками традиційної фінансової системи, таких як зменшення кількості довіри, необхідного для здійснення транзакцій. Хоча обидва проєкти так і не були запущені офіційно, вони були джерелом натхнення для створення Bitcoin.

Термін криптовалюта увійшов у загальнодоступне використання зі сплеском Bitcoin у 2008 році. У жовтні цього року в статті людини чи групи людей під псевдонімом Сатоші Накамото «Bitcoin: A Peer-to-Peer Electronic Cash System» було описано систему створення цифрової валюти під назвою Bitcoin, яка використовує технологію блокчейн [32]. Цей день в історії Bitcoin проклав шлях для наступних подій. Після більше 10 років з моменту виникнення технології блокчейн, на ринку з'явилося безліч криптовалют та низки інших застосувань технології блокчейну. Стаття Накамото фактично запустила криптовалютну революцію.

Через чотири місяці Сатоші Накамото, чия справжня особистість залишається таємницею донині, видобув перший блок мережі Bitcoin, фактично пілотуючи технологію блокчейн. У 2010 році Сатоші Накамото призупинив свою участь у проєкті та передав справи своїм колегам. Спілкування з ними він підтримує і досі через сайт P2P Foundation. Там же вказана особиста інформація про вік та місце проживання: Японія, 37 років. Однак у користувачів коректність персональних даних викликає сумніви. Статок Сатоші Накамото оцінюються в 1,1 млн Bitcoin (станом на травень 2022 року це складало трохи більше, ніж 29 млрд USD) [29].

Основне призначення всіх криптовалют – полегшити та прискорити процес транзакцій. Усі вони побудовані на основі технології блокчейн. Щоб розібратися в механізмі роботи криптовалют, слід розуміти суть терміну «блокчейн».

«Блокчейн» – це програмний код, розроблений Сатоші Накамото в 2009 році для безпечного зберігання та передачі інформації в мережі Інтернет, включаючи фінансові транзакції, документи та іншу інформацію. «Блокчейн» складається з ланцюжка блоків і дає можливість проводити транзакції та добувати валюту Bitcoin усередині мережі. Користувач отримує валюту за обчислення комп'ютером певних

математичних алгоритмів, заданих блоком. Особливість «блокчейн» полягає у записі всіх проведених операцій на блоках без права зміни, видалення чи фальсифікації вже внесеної інформації, оскільки записи зберігаються на кожному вузлі мережі. Слід зазначити, що всі дані передаються згідно з механізмом Peer-to-Peer (від одного користувача до іншого або P2P), що передбачає відсутність посередника між відправником та одержувачем монети [26].

З точки зору безпеки, блокчейн має кілька переваг порівняно з традиційними централізованими системами. Оскільки дані розподіляються по всій мережі, зламування або викрадення всієї інформації стає набагато складніше, ніж у випадку з централізованою базою даних. Крім того, кожен блок в ланцюзі містить хеш попереднього блоку, що робить зміну будь-якого блоку майже неможливою без перерахунку всієї ланки блоків. Також блокчейн не потребує посередників, тому транзакції можуть бути здійснені без втручання третьої сторони. Це забезпечує швидші та більш ефективні операції. Проте, технологія блокчейн також має свої недоліки. Найбільшим з них є швидкість обробки транзакцій, оскільки кожна транзакція повинна бути підтверджена мережею. Окрім цього, розмір блокчейну постійно зростає, що може викликати проблеми зі зберіганням і обробкою даних.

1.2. Процес купівлі/продажу та зберігання криптовалюти

Перша зареєстрована покупка товарів за допомогою Bitcoin була здійснена у травні 2010 року чоловіком, на ім'я Ласло Хейніц. Програміст Ласло Хейніц попросив у спільноті BitcoinTalk доставити йому піцу за 10 000 Bitcoin. Він отримав відгук від користувача під ніком «jercos», який доставив йому дві піци в обмін на ці 10 000 Bitcoin. На той момент це було дуже незвичною подією, оскільки кількість людей, які знали про Bitcoin, була дуже обмежена. Але з часом ця історія стала дуже популярною і отримала статус легенди. Цей день досі відзначається як Bitcoin Pizza Day [9].

Після народження Bitcoin як першої криптовалюти потрібно було знайти рішення для торгівлі ними. У березні 2010 року з'явилася перша біржа криптовалют, відома як bitcoinmarket.com (на сьогоднішній день вже не функціонує). У липні того ж року була запущена Mt.Gox. Протягом 2013 року з'явилося кілька інших

конкуруючих криптовалют: Станом на травень 2013 року на ринку криптовалют налічувалось 10 цифрових активів, включаючи Litecoin. Ще одним значним криптоактивом, що приєднався до ринку у серпні того ж року, став XRP (Ripple).

У міру того, як вартість Bitcoin зростала, активувалися перші хакери на крипторинку. У червні 2011 року Mt.Gox зламали: було вкрадено 2000 BTC, вартість яких на той момент становила близько 30 000 доларів. У 2013 році Mt.Gox стала найбільшою криптовалютною біржею, на піку якої оброблялося 70% усіх транзакцій з Bitcoin. Проте, у 2014 році Mt.Gox стала першою великою біржею криптовалют, з якої було вкрадено 850 000 BTC. Це найбільший випадок крадіжки Bitcoin в історії криптовалюти, який на той час оцінювався в 475 мільйонів доларів США (поточна вартість близько 23,5 мільярдів доларів) [31]. Після цієї безпрецедентної ситуації ціна Bitcoin впала на 50% і не відновилася до початкової вартості до кінця 2016 року. Хакерські атаки на криптовалютні біржі залишаються актуальними і поширеними явищами, хоча досить рідко мають масштаб Mt.Gox. Тому важливо пам'ятати, що для забезпечення безпеки криптовалютних активів необхідно дотримуватись деяких простих правил.

Для того, щоб купити криптовалюту, необхідно пройти три етапи, але сам процес може бути досить простим і зручним. Першим важливим кроком необхідно здійснити вибір платформи, на якій буде проводитись купівля/продаж криптовалюти. Зазвичай, можна вибрати між традиційним брокером та спеціальними біржами криптовалюти.

Традиційні онлайн-брокери представляють собою платформи, які надають можливість придбати та продати криптовалюту, а також інші фінансові активи, такі як ETF, акції та облігації. Ці платформи, здебільшого, пропонують нижчі витрати на торгівлю, але обмеженіший набір криптовалютних функцій. Щодо криптовалютних бірж, то їх є доволі багато на вибір, вони пропонують більш широкий спектр криптовалютних функцій, включаючи зберігання гаманців та відсоткові рахунки, але їх комісії можуть бути вищі. Більшість бірж стягують комісію в залежності від транзакцій чи способу оплати. Порівнюючи різні платформи при виборі, необхідно враховувати, які криптовалюти пропонуються, які комісії стягуються, безпеку брокерів чи бірж, варіанти зберігання та виведення активів.

Після того, як платформа була обрана, наступним кроком буде поповнення рахунку для початку торгівлі. Більшість криптовалютних бірж дозволяють користувачам придбати криптовалюту за допомогою фіатних (тобто випущених державою) валют, таких як долар США чи євро, за допомогою карток (дебетових чи кредитних), хоча це залежить від платформи.

Варто зазначити, що придбання криптовалют за допомогою кредитних карт вважається ризикованим. Це пов'язано з тим, що криптовалюти дуже нестабільні, і не рекомендується ризикувати заборгованістю або потенційно сплачувати високі комісійні за транзакції з кредитної картки. Деякі криптовалютні біржі можуть навіть не підтримувати цей вид оплати, тому слід заздалегідь це перевіряти для кожної платформи.

Окрім того, важливими факторами, які слід враховувати, є комісії, які можуть різнитися в залежності від обраного способу оплати та використовуваної платформи. Комісії можуть стягуватися за потенційні депозити та зняття коштів з рахунку, а також за торгівлю. Крім того, слід враховувати, що на деяких криптовалютних біржах можуть бути додаткові комісії за виведення коштів з гаманця.

Після того, як була обрана платформа та поповнений рахунок, наступним кроком є оформлення замовлення на купівлю або продаж криптовалюти. Його можна розмістити через веб- або мобільну платформу брокера або біржі. Для того, щоб купити чи продати криптовалюту, необхідно вибрати тип замовлення (наприклад, ринкове замовлення або лімітне замовлення), ввести кількість криптовалют для купівлі/продажу та підтвердити замовлення.

Після оформлення замовлення, воно буде відображено у списку відкритих замовлень на платформі. Залежно від типу замовлення, воно може бути виконано негайно або за певних умов. Наприклад, якщо користувач розмістив ринкове замовлення на купівлю криптовалюти, то воно буде виконано негайно за поточну ринкову ціну.

Якщо ж користувач вирішив розмістити лімітоване замовлення на купівлю чи продаж, то воно буде виконано тільки тоді, коли ціна криптовалюти досягне вказаного рівня. Якщо ж ціна не досягне вказаного рівня, замовлення залишиться

відкритим і може бути виконано в майбутньому, якщо ціна відповідатиме вказаному рівню.

Крім традиційного способу купівлі криптовалюти через біржі та брокерські платформи, існують інші способи, які дозволяють інвестувати в цей цифровий актив. Наприклад, деякі платіжні послуги дозволяють своїм клієнтам здійснювати операції з купівлі, продажу та зберігання криптовалюти безпосередньо на їхніх платформах. Це дуже зручний спосіб для тих, хто не бажає мати справу зі складним процесом купівлі та зберігання криптовалюти на біржі.

Ще один з таких способів – майнінг криптовалюти. Цей метод вимагає великих витрат на обладнання та електроенергію, але може бути дуже прибутковим для тих, хто готовий вкласти значні зусилля та кошти. Ідея майнінгу полягає в тому, щоб створювати нові блоки та підтверджувати транзакції в мережі криптовалюти. За ці зусилля майнери отримують винагороду в криптовалюті.

Також, ще одним популярним способом інвестування в криптовалюту є ICO (Initial Coin Offering). Первинне розміщення монет (ICO) можна порівняти з первинним публічним розміщенням (IPO) у сфері криптовалют [22]. Компанія, яка має намір залучити фінансування для створення нової криптовалюти, програми або послуги, може запустити ICO як спосіб залучення коштів. Інвестори можуть купувати токени та, якщо проект успішний, отримувати високий прибуток.

Крім того, існують інші інструменти інвестування, які дають доступ до криптовалют через традиційний фондовий ринок. Наприклад, Bitcoin trusts – це можливість інвестувати в акції Bitcoin-трастів через звичайний брокерський рахунок. Взаємні фонди Bitcoin (Bitcoin mutual funds) також є зручним інструментом, який дозволяє інвестувати в криптовалюту через пайові фонди Bitcoin ETF.

Ще один спосіб інвестування в криптовалюту – це блокчейн-акції або ETF. Це дає можливість опосередковано інвестувати в криптовалюту шляхом блокчейн-компаній, що спеціалізуються на криптотехнологіях для транзакцій. Крім того, можна придбати акції або ETF компаній, що займаються розробкою та використанням технології блокчейн [41].

Після придбання криптовалюти, важливо забезпечити її безпечне зберігання, щоб захистити від крадіжки або злому. Один із способів забезпечення безпеки криптовалют – використання крипто-гаманців, які можуть бути фізичними пристроями або онлайн-програмами, що дозволяють зберігати приватні ключі криптовалют безпечно. Для використання гаманця необхідно мати номер гаманця та код доступу, що забезпечує можливість здійснення безпечних та анонімних операцій. Деякі криптовалютні біржі надають власні гаманці для зручного зберігання прямо на платформі.

Існує ще кілька варіантів зберігання криптовалюти. Перш за все, це може бути сторонній сервіс: можна зберігати криптовалюту у третьої сторони, наприклад біржі, яка надає послуги гаманця. Такий спосіб зберігання криптовалюти потребує повної довіри до біржі. Для того, щоб налаштувати гаманець на біржі, потрібно створити обліковий запис, використовуючи інформацію про власника, таку як ім'я, паспорт або ідентифікаційний номер.

Також є можливість самостійно створити свій гаманець. За використання такого способу можна обійти вимоги щодо ідентифікації сторонніх постачальників. При такому підході до зберігання криптовалюти, безпека активів буде повністю залежати від власника гаманця. Вибір типу криптовалютного гаманця буде залежати від конкретних потреб та бажаних функцій користувача. Важливо мати на увазі, що деякі криптовалютні гаманці можуть підтримувати лише певні види криптовалют або мати обмежену функціональність. Іноді це може означати компроміс між безпекою та зручністю використання [17].

Зберігання криптовалюти – це насправді безпечне зберігання приватного ключа або рядка даних, схожих на пароль. Закриті ключі можуть мати різні формати, наприклад, представлені у вигляді двійкового коду, QR-коду, мнемонічної фрази або інших форматів. Закриті ключі можуть зберігатися або в мобільних додатках, або програмах для настільних комп'ютерів, які переважно використовують «гаряче зберігання», оскільки ці гаманці регулярно підключаються до Інтернету. Також існує можливість зберігання закритих ключів на спеціалізованому окремому апаратному пристрої, який не підключений до Інтернету, це так зване «холодне зберігання».

Холодні гаманці, зазвичай, стягують комісію, а гарячі – ні. Існує також можливість використання гаманця з кількома підписами, який вимагає декількох закритих ключів для схвалення транзакції перед перенесенням активів. Теоретично це може підвищити безпеку коштів.

Щоб здійснити транзакцію, знадобиться кілька речей: доступ до криптовалюти, адреса гаманця (або відкритий ключ) вашого контрагента – це може мати форму рядка або QR-коду. Після введення суми та адреси контрагента, система має підтвердити операцію та зберегти її. Це означає, що інформація про цю операцію має увійти до складу коду нового згенерованого блоку. Для повного підтвердження системою інформація про транзакцію повинна бути включена до всіх 6 блоків. Це надає надзвичайно високий рівень безпеки для користувачів криптовалюти. Цей процес також забезпечує, що будь-яка транзакція повинна бути підтверджена кількома вузлами, перш ніж вона буде занесена до блокчейну. Таким чином, будь-яка спроба здійснити несанкціоновану транзакцію або змінити даний блок буде миттєво виявлена та зупинена громадськими вузлами, оскільки будь-яка зміна в будь-якому блоку буде суперечити вже існуючим записам в блокчейні.

Однак, проведення операції може зайняти 5-7 годин, що може бути не дуже зручним для деяких користувачів. Тому, деякі криптовалютні гаманці дозволяють користувачам налаштувати рівень підтверджень, необхідних для проведення транзакції. Наприклад, якщо користувач відправляє меншу суму, то він може вибрати швидкість проведення транзакції, яка потребує меншої кількості підтверджень. Це може скоротити час проведення операції, але при цьому може зменшити рівень безпеки операції.

1.3. Ринок криптовалют та вплив соціальних мереж на нього

Криптовалюти повільно, але впевнено стають глобальним явищем. Зростаюча тенденція до криптовалют привела до більшої прийнятності і використання їх у різних випадках. Навіть створення цифрової валюти Центрального Банку вже є реальністю, а великі компанії показують зростаючий інтерес до блокчейну та криптовалют, інвестуючи в них. Очевидно, такі події сприятимуть швидкому розширенню ринку.

У липні 2015 року було запущено криптовалюту Ethereum [21]. Зараз це другий за обсягом ринкової капіталізації криптовалютний актив, який приніс у світ криптовалют «розумні» контракти та, у кінцевому рахунку, децентралізовану фінансову систему. Це дозволяє блокчейну Ethereum управляти власною екосистемою на своєму блокчейні, а також мати свою власну валюту – Ether (ETH). Найменша одиниця Етера також відома як Вей.

Криптовалюти продовжують залишатися однією з найбільш захоплюючих тем для інвесторів, підприємців, регуляторів та широкої громадськості. Зміни їх цін, спекулятивні коливання та дискусії про справедливую цінність криптовалют спровокували багато глобальних обговорень. У цих обговореннях висловлюються як позитивні, так і негативні думки, включаючи твердження про те, що ринок криптовалют може бути бульбашкою без реальної фундаментальної цінності, а також постає стурбованість щодо потреби регулювання та правового контролю в цій сфері.

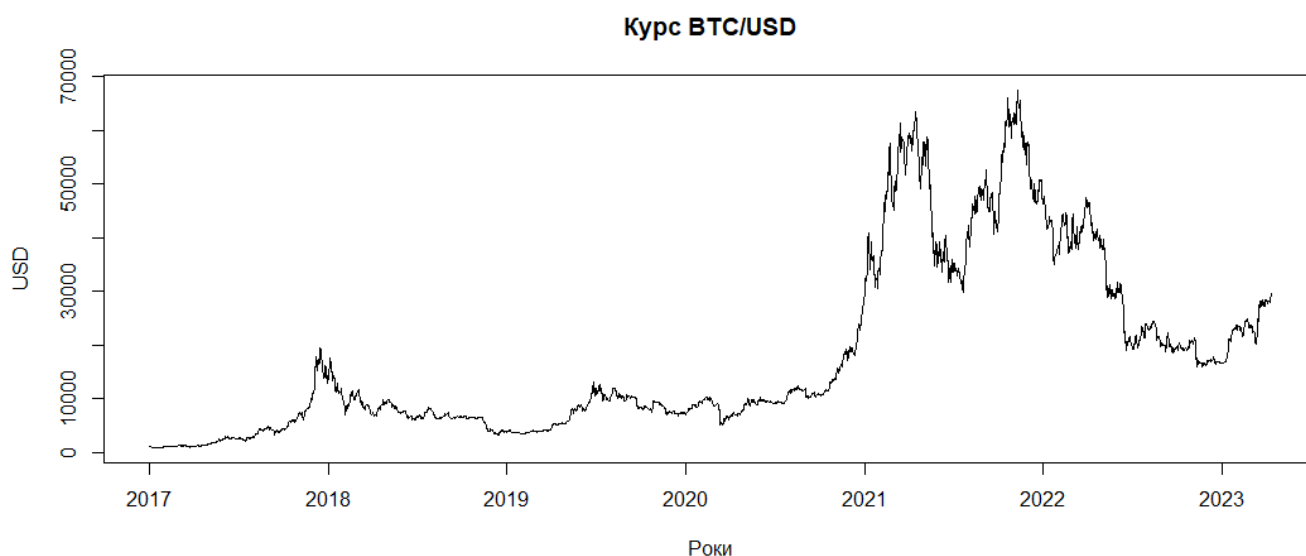


Рис. 1.1. Курс криптовалюти Bitcoin від січня 2017 року до квітня 2023 року. Джерело: [42].

Проте, на сьогоднішній день, оцінка реальної економічної вартості криптовалют, таких як Bitcoin BTC і Ethereum, залишається складною задачею. Однак, щоб отримати уявлення про ринкову вартість криптовалют, можна проаналізувати їх курси відносно існуючих фіатних валют (наприклад, долара США, євро, і т.д.), які використовуються у повсякденному житті (рис. 1.1, рис. 1.2). Хоча курси

криптовалют є дуже нестабільними, вони демонструють, що криптовалюти мають цінність для тих, хто готовий платити валюту, щоб їх придбати.

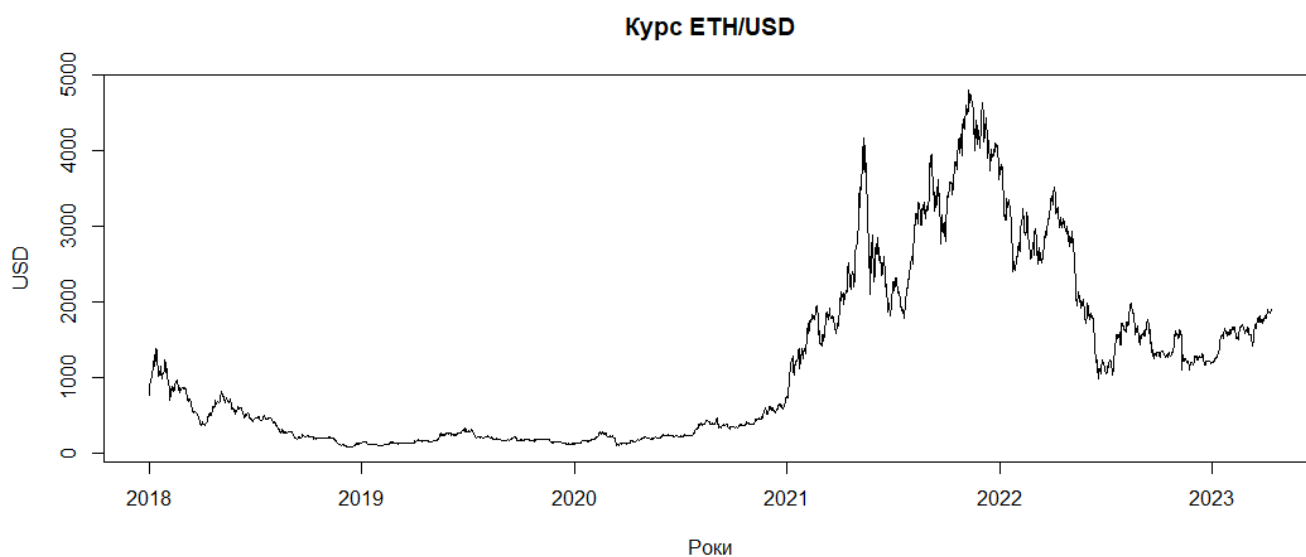


Рис. 1.2. Курс криптовалюти Ethereum від 2018 року до квітня 2023 року.
Джерело: [42].

У 2023 році загальна ринкова капіталізація криптовалют оцінюється приблизно в 1,1 трильйони доларів США, але це значно нижче, ніж у листопаді 2021 року. Тоді капіталізація крипторинку сягала приблизно 3 трильйони доларів (рис. 1.3).

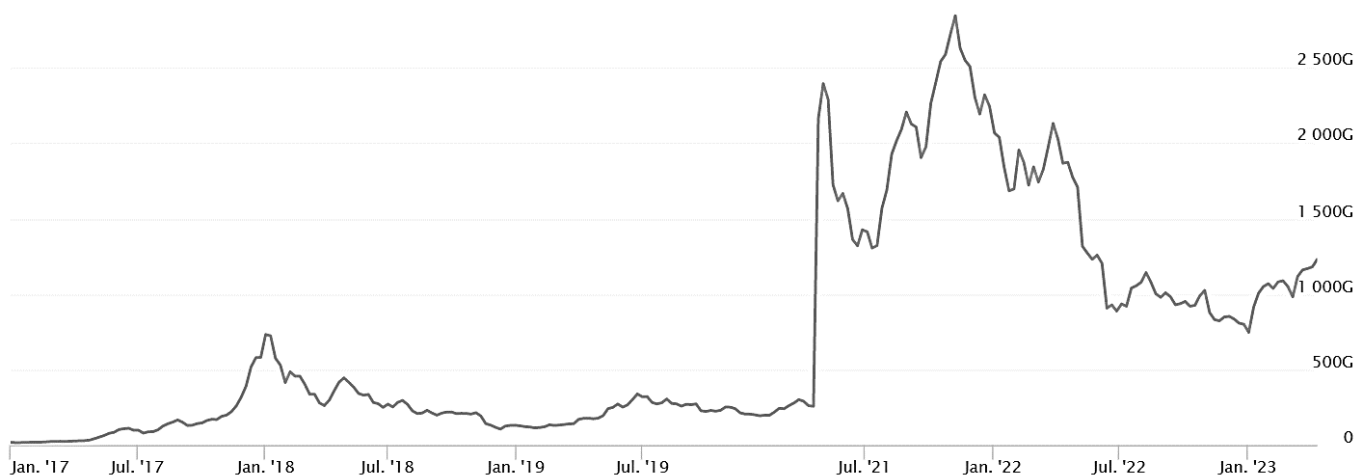


Рис. 1.3. Загальна капіталізація ринку криптовалют з 2017 по 2023 роки.
Джерело: [16].

Кілька років тому про стан ринку криптовалют свідчив, у більшості, курс Bitcoin. Сьогодні все трохи змінилося. Ринок криптовалют зріс з кількох криптовалют до сотень тисяч різноманітних криптоактивів. Хоча Bitcoin все ще є найпомітнішою та

однією із найпопулярніших криптовалют з точки зору ціни за одиницю та загальної ринкової капіталізації, але разом із зростанням кількості криптовалютних проєктів, доля Bitcoin в загальній ринковій капіталізації становить значно менше 50% всієї вартості усіх криптоактивів [18].

Станом на квітень 2023 року існує 22 904 криптовалют. Однак не всі криптовалюти є активними. Тобто безпосередньо в обігу перебуває лише близько 8 832 активних криптовалют [24]. Також варто зазначити, що не всі ці проєкти мають реальну економічну цінність і перспективи розвитку. Багато з них можуть бути класифіковані як «маски» або «шахрайські» проєкти, які створюються з метою залучення коштів від недосвідчених інвесторів. Проте, на ринку також є проєкти, які розвиваються і мають реальну вартість для споживачів та інвесторів.

Ринок криптовалют не тільки розширюється, але й стає більш різноманітним, з'являються нові криптовалюти, які використовують різні технології, протоколи та функції. Наприклад, розвивається сегмент криптовалют, пов'язаних з екологічно чистим майнінгом, мережами з низькою комісією, заснованих на блокчейні з розширеними функціями «розумних контрактів» та іншими інноваційними технологіями.

Найбільші криптовалюти за загальною вартістю наявних монет у доларах, тобто за ринковою капіталізацією, знаходяться в табл. 1.1. Можемо помітити, що загальна ринкова капіталізація Bitcoin складає 480 мільярдів доларів США. Ethereum, у свою чергу, має ринкову капіталізацію більше 200 мільярдів доларів США, що в 2,3 разів менше, ніж у Bitcoin, але все ще значно вище, ніж будь-яка інша криптовалюта на ринку. Ці дані підтверджують важливість Bitcoin та Ethereum на ринку криптовалют, а також показують, що вони займають значну частку ринку.

Деякі економічні аналітики прогнозують великі зміни для криптовалюти в майбутньому та вважають, що майбутнє криптовалюти залежить від того, як вона буде інтегрована в реальність і які регулятивні заходи будуть введені. Наприклад, існує можливість того, що криптовалюта може бути включена до списку торговельних активів на NASDAQ, що ще більше підвищить довіру до блокчейну та його роль як альтернативи традиційним валютам. Деякі експерти передбачають, що

існування перевіреного біржового фонду (ETF) для криптовалюти може полегшити процес інвестування в Bitcoin, однак попит на цю криптовалюту залежить від бажання інвесторів і може не виникнути автоматично після введення фонду.

Таблиця 1.1

Топ-20 криптовалют за капіталізацію станом на березень 2023 року

№	Назва криптовалюти	Абревіатура	Ціна (дол. США)	Капіталізація (млрд дол. США)
1	Bitcoin	BTC	24852,93	480,23
2	Ethereum	ETH	1690,57	206,85
3	Tether	USDT	1	73,26
4	BNB	BNB	307,87	48,63
5	USD Coin	USDC	1	39,48
6	XRP	XRP	0,37	18,88
7	Cardano	CAR	0,34	11,92
8	Polygon	MATIC	1,16	10,16
9	Dogecoin	DOGE	0,07	9,60
10	Binance USD	BUSD	1	8,39
11	Solana	SOL	20,12	7,71
12	Polkadot	DOT	6,11	7,12
13	Dai	DAI	1	6,45
14	TRON	TRX	0,07	6,08
15	SHIBA INU	SHIB	0,00	6,05
16	Litecoin	LTC	80,71	5,85
17	Avalanche	AVAX	16,38	5,33
18	Uniswap	UNI	6,07	4,63
19	Wrapped Bitcoin	WBTC	24570,19	3,77
20	Cosmos	ATOM	12,13	3,47

Джерело: [18].

Як вже відомо, ціна криптовалют є дуже нестабільною і може змінюватися дуже швидко. Багато різних факторів можуть впливати на ціну криптовалют, включаючи технічні, фундаментальні та соціальні чинники. Розглянемо деякі з них.

- Регулювання. Криптовалюти є децентралізованими, але все ж можуть бути регульованими урядами та фінансовими інституціями, хоча уряди все ще не мають найкращих практик регулювання криптовалюти, що робить її доволі ризикованою та нестабільною інвестицією. Такі продукти, як

біржові фонди (ETF) або ф'ючерсні контракти, можуть забезпечити більший доступ для інвесторів, тим самим збільшити їхню вартість. Регулювання також може дозволити інвесторам робити ставки проти ціни криптовалюти з ф'ючерсними контрактами або опціонами, що може зменшити волатильність цін на криптовалюту. Регулювання також може негативно вплинути на попит на криптовалюту, зокрема, заборона криптовалют в певних країнах може призвести до значного зниження попиту на ці активи.

- Попит та пропозиція. Як і з будь-яким іншим активом, ціна криптовалюти залежить від співвідношення попиту та пропозиції на ринку. Якщо попит на криптовалюту зростає, а пропозиція залишається стабільною або зменшується, ціна може зрости. Bitcoin, наприклад, має фіксований запас у 21 мільйон одиниць. Інші, такі як Ethereum, не мають обмеження пропозиції.
- Вартість виробництва. Процес «майнінгу» використовується для створення нових токенів криптовалюти. Майнери використовують комп'ютер, щоб підтвердити наступний блок на блокчейні. Чим більше конкуренції є для добування певної криптовалюти, тим складніше її добувати, оскільки майнери змагаються, щоб вирішити складну математичну задачу для підтвердження блоку. Тому вартість майнінгу зростає, оскільки команда потребує потужнішого обладнання, такого як комп'ютери, для успішного добування. При зростанні витрат на майнінг криптовалюта також збільшує свою вартість. Майнери не будуть добувати криптовалюту, якщо її вартість, яку вони майнять, недостатньо висока, щоб відшкодувати їх витрати. Отже, в умовах попиту на блокчейн та його застосування, очікується зростання ціни.
- Біржі криптовалют. Майже будь-яка біржа криптовалют перелічує найпопулярніші токени та проводить торги з основними криптовалютами, такими як Bitcoin та Ethereum, на кількох біржах одночасно. Для деяких

менших токенів може бути обмежений доступ, оскільки вони торгуються лише на кількох біржах, що може ускладнити можливості інвестування. Якщо криптовалюта також слабо торгується на невеликій біржі, комісія, яку забирає біржа, може бути занадто високою для деяких інвесторів. Якщо криптовалюта перелічена на кількох біржах, це може збільшити кількість інвесторів, готових купувати її, тим самим збільшуючи попит, а з ним зростає і ціна.

- Конкуренція. Існує безліч криптовалют, і конкуренція між ними може впливати на ціни. Наприклад, якщо нова криптовалюта виходить на ринок з більш покращеними функціями, вона може залучити більше інвесторів та популярності, що може знизити попит на інші криптовалюти. Проте, хоча їх доволі легко запустити, важко зробити їх життєздатними, оскільки потрібно створити мережу користувачів для цієї криптовалюти.
- Внутрішнє управління. Криптовалютні мережі рідко працюють за статичними правилами, і їх розвиток залежить від спільноти, яку вони обслуговують. Деякі мережі надають власникам можливість приймати рішення щодо видобутку або використання токена. Для внесення будь-яких змін в управління токеном необхідний консенсус між зацікавленими сторонами. Інвесторам зазвичай подобається стабільне управління, яке також забезпечує більш стабільне ціноутворення. Однак, повільний процес оновлення програмного забезпечення для поліпшення протоколів може обмежувати цінність криптовалюти. Якщо оновлення забезпечує нові можливості для власників криптовалюти, але його впровадження займає багато часу, це може негативно вплинути на зацікавлені сторони [12].
- Соціальні медіа. Криптовалюти завжди були на передовій новин у засобах масової інформації. Новини про успіхи та провали в крипторинку можуть вплинути на попит та ціну криптовалют. Розглянемо саме цей фактор детальніше.

Соціальні мережі відіграють все більш важливу роль у визначенні цін на криптовалюти. Вони забезпечують платформу для спільноти людей, які обговорюють різні аспекти ринку криптовалют, включаючи новини, технічні аналізи, прогнози цін та багато іншого. З появою таких платформ, як Twitter і Reddit, інформація та думки про цифрові валюти можуть швидко поширюватися, впливаючи на їх вартість на ринку. Наприклад, новини про розробки великих компаній або урядів можуть вплинути на ціну певної криптовалюти. Крім того, соціальні мережі дозволяють розповсюджувати інформацію про нові ICOs (Initial Coin Offerings) та інші проєкти, що можуть впливати на ціни криптовалют.

Один із способів впливу соціальних медіа на криптовалютний ринок – це вплив інфлюенсерів. Криптоінфлюенсери – це люди, які зібрали велику кількість підписників на платформах соціальних мереж, часто завдяки обміну досвідом і думками щодо криптовалют. Коли ці впливові особи поділяють позитивні чи негативні настрої щодо певного токена чи проєкту, це може мати значний вплив на його ціну.

Twitter є однією з найбільш популярних соціальних мереж, де криптоінвестори та експерти виражають свої думки щодо ринку. Це може включати прогнози цін, поради щодо інвестування та інформацію про нові ринки та технології. Багато людей вважають, що Twitter є дуже важливим джерелом інформації про криптовалюти та ринок в цілому. Наприклад, якщо популярний криптовалютний інфлюенсер твітне про певний токен, це може призвести до різкого зростання попиту на цей токен і відповідного зростання його вартості.

Прикладом впливу соціальних мереж на криптовалюту може бути ситуація, коли Ілон Маск, генеральний директор Tesla, зробив пост у Twitter про Dogecoin. Цей пост мав значний вплив на вартість цієї альтернативної криптовалюти. Внаслідок цього посту, ціна Dogecoin різко зросла до свого найвищого рівня в історії [4]. Цікаво, що в біографії Маска на Twitter з'явилася фраза «колишній генеральний директор Dogecoin». Це призвело до подальшого зростання вартості Dogecoin на 17%. Також, після того, як Ілон Маск у квітні 2023 року змінив логотип Twitter на зображення собаки Shiba Inu, ціна Dogecoin раптово підскочила [7]. Такий приклад демонструє,

як одне повідомлення в соціальній мережі може миттєво вплинути на ціну криптовалюти і викликати значні коливання на ринку.

Соціальні мережі також можуть допомогти вивчити поведінку користувачів на ринку. Аналіз поведінки користувачів, які обговорюють криптовалюти на соціальних мережах, може допомогти вивчити їхні настрої та передбачити майбутні тенденції на ринку. Відстежуючи ключові слова та хештеги, пов'язані з криптовалютами, аналітики можуть оцінити громадську думку щодо різних токенів і проєктів. Це може бути корисним для прогнозування ринкових рухів і визначення потенційних інвестиційних можливостей. Наприклад, якщо багато користувачів в соціальних мережах розміщують повідомлення про те, що вони купують певну криптовалюту, це може вказувати на зростання попиту на цю криптовалюту та спричинити зростання її цін.

Однак вплив соціальних мереж на ціни криптовалюти не завжди є позитивним. Дезінформація та фейкові новини, чутки та неправдиві твердження про криптовалюти також можуть швидко поширюватися у соціальних мережах. Це може призвести до паніки продажів чи покупок, викликаючи різкі та непередбачувані коливання цін.

Крім того, соціальні мережі можуть бути використані для маніпулювання цінами криптовалют. Деякі групи користувачів можуть об'єднатися для спроби змусити ціну криптовалюти зрости або знизитися, розповсюджуючи недостовірну інформацію або «фейкові новини» на соціальних мережах. Також, відомо, що інфлюенсери та знаменитості, які мають велику кількість прихильників в соціальних мережах, можуть використовувати свій вплив, щоб рекламувати або підбадьорювати своїх шанувальників купувати певну криптовалюту.

З одного боку, це може призвести до значного зростання ціни криптовалют, якщо впливова особистість має велику кількість прихильників, які довіряють її рекомендаціям. З іншого боку, це також може привести до шахрайства та обману, особливо якщо люди купують криптовалюту на основі недостовірної інформації або рекомендацій.

Прикладом цього може бути криптовалютна афера на Twitter (Crypto Twitter Scam) – це шахрайська схема, яка відбувається на платформі Twitter, пов'язана з

обманом користувачів щодо криптовалют. Шахраї виступають під іменами відомих осіб, експертів чи проєктів у сфері криптовалют та блокчейну і пропонують швидке збагачення або виграш. Вони надсилають обмежений часовий пропозиційний твіт, в якому просять користувачів відправити певну кількість криптовалюти на певний гаманець з обіцянкою отримати значно більшу суму назад. Однак, після надсилання коштів, користувачі не отримують жодного повернення або виграшу, а шахраї зникають [35].

Відповідно до усього вище сказаного, перспективним видом електронних грошей можна назвати криптовалюту. Загалом криптовалюта є видом цифрової валюти, в основі якої лежить цифровий код. Її можна зберігати безпосередньо в електронних гаманцях, переводити з якогось одного гаманця в інший. Соціальні медіа стали важливим інструментом для інвесторів і трейдерів на ринку криптовалют. Надаючи доступ до інформації та статистичних даних у реальному часі, платформи соціальних мереж можуть впливати на вартість криптовалют. Однак інвестори повинні бути обережними та критичними, оцінюючи інформацію в соціальних мережах, і завжди проводити ретельні дослідження, перш ніж приймати інвестиційні рішення.

РОЗДІЛ II

МЕТОДИ ДЛЯ АНАЛІЗУ ЧАСОВИХ РЯДІВ

2.1. Огляд часових рядів та їх властивостей

Часовий ряд – це сукупність спостережень за чітко визначеними елементами даних, отриманими шляхом повторних вимірювань у часі. Дані, які збираються нерегулярно або лише один раз, не є динамічними рядами [40].

Часовий ряд можна розглядати як послідовність значень досліджуваної величини, зібраних у часі відповідно до певного регулярного або нерегулярного розкладу. Кожен елемент цієї послідовності називається рівнем часового ряду і відповідає певному моменту в часі. Якщо досліджувана величина є одновимірною, тобто кожному моменту часу відповідає значення лише одного показника, ряд є одновимірним. В іншому випадку, якщо кожному моменту часу відповідає декілька показників, то ряд є багатовимірним.

Основні характеристики часових рядів, які можуть бути використані для їхньої аналізу, включають такі компоненти:

- Тренд – це загальна тенденція ряду до зростання або спаду протягом тривалого періоду. Визначення тренду може бути корисним для прогнозування майбутніх значень ряду.
- Сезонність – це регулярні циклічні зміни в ряді, які повторюються через певні періоди часу, зазвичай в течії року. Наявність сезонності може мати значний вплив на поведінку ряду.
- Коливання – це випадкові зміни, які не можуть бути пояснені трендом або сезонністю. Коливання можуть бути корисними для прогнозування короткострокових змін в ряді.
- Циклічність – це зміна значень ряду, що повторюється з певною періодичністю, але не пов'язана з сезонністю. Циклічність може бути результатом економічних циклів, природних катаклізмів тощо.

Додатковою характеристикою часового ряду може бути стаціонарність, що вказує на те, що статистичні властивості ряду не змінюються з часом. Стаціонарність

є важливою для багатьох методів аналізу часових рядів, таких як моделі ARIMA та ARIMAX.

Аналіз часових рядів є складною і багатогранною задачею, що вимагає від дослідника вміння визначати і описувати властивості рядів, будувати математичні моделі та прогнозувати майбутні значення. Основні задачі аналізу часових рядів можна сформулювати наступним чином:

1. Опис часового ряду: дослідження тенденції, сезонності, циклів, випадкових коливань та інших властивостей ряду, що дозволяє зрозуміти його поведінку в часі та зробити висновки про його динаміку.
2. Моделювання часового ряду: визначення математичної моделі, що краще відображає властивості ряду та дозволяє побудувати прогноз майбутніх значень.
3. Валідація моделі: оцінка точності моделі, її адекватності та здатності до прогнозування, що дозволяє зробити висновки про її ефективність.
4. Прогнозування майбутніх значень: використання побудованої моделі для прогнозування майбутньої поведінки часового ряду.
5. Побудова стратегій управління: на основі результатів аналізу часового ряду можна побудувати стратегії управління, що дозволять зменшити ризики та оптимізувати процеси управління.
6. Пошук аномалій та виявлення змін: аналіз часових рядів може допомогти виявити аномальні значення та зміни в динаміці ряду, що дозволить вчасно реагувати на події та уникнути можливих проблем.

Аналіз часових рядів зазвичай складається з наступних кроків:

- Первинний огляд даних: цей етап включає у себе вивчення структури даних, ознайомлення з рядом та його властивостями. На цьому етапі можна виявити аномальні значення, пропуски даних та інші неточності, що потребують подальшої обробки.

- Визначення тренду: тренд – це довгострокова тенденція, що відображає загальний напрямок розвитку ряду. На цьому етапі визначається чи є в ряду тренд, які його характеристики та як він впливає на ряд.
- Виявлення циклів та сезонності: циклічність та сезонність є властивими більшості часових рядів. На цьому етапі визначається частота та амплітуда циклів та сезонних ефектів.
- Аналіз шуму: шум – це випадковий процес, який не має відношення до тренду, циклів та сезонності. На цьому етапі вивчається характер шуму та вплив на ряд.
- Моделювання та прогнозування: на цьому етапі розроблюються математичні моделі ряду, що дозволяють краще розуміти його характеристики та здійснювати прогнозування майбутніх значень ряду [3].

Найпоширенішими моделями для прогнозування та аналізу часових рядів є наступні:

- ARIMA (інтегроване ковзне середнє з авторегресією): популярна модель, яка поєднує компоненти авторегресії (AR) і ковзного середнього (MA) для врахування тенденцій, сезонності та стаціонарності часового ряду.
- ARIMAX (авторегресійне інтегроване ковзне середнє з пояснювальними змінними): розширення моделі ARIMA, що включає додаткові пояснювальні змінні, які можуть впливати на часовий ряд.
- SARIMA (Seasonal Autoregressive Integrated Moving Average): варіант моделі ARIMA, яка враховує сезонні тенденції та паттерни в часовому ряді.
- SARIMAX (інтегроване ковзне середнє з сезонною авторегресією та пояснювальними змінними): розширення моделі SARIMA, що включає пояснювальні змінні та сезонні фактори.
- VAR (векторна авторегресія): модель, яка включає кілька змінних часових рядів та їхні взаємозалежності для прогнозування майбутніх значень змінних.

- LSTM (Long Short-Term Memory): тип рекурентної нейронної мережі, яка може вивчати минулі дані, щоб робити прогнози щодо майбутніх значень.
- Prophet: модель прогнозування, розроблена Facebook, яка використовує розкладну модель часового ряду з такими компонентами, як тренд, сезонність і вихідні (святкові) дні.

Розглянемо детальніше види моделей ARIMA, починаючи з авторегресійної моделі, поступово ускладнюючи її.

2.2. Модель ARIMA

Y_t називається авторегресійним рядом порядку p , AR(p), якщо він задовольняє умову

$$Y_t = \varphi_1 Y_{t-1} + \dots + \varphi_p Y_{t-p} + \varepsilon_t, \quad (1)$$

де ε_t – білий шум і φ_p – коефіцієнти параметрів. Наступне значення, яке спостерігається в ряді, є невеликим збуренням простої функції останніх спостережень.

У випадку k -го порядку кореляція між Y_t і Y_{t-k} частково може бути зумовлена кореляцією цих спостережень з проміжними лагами $Y_{t-1}, Y_{t-2}, \dots, Y_{t-k+1}$. Щоб коригувати цю кореляцію, розраховуються часткові автокореляції, PACF.

Для AR(1) моделі:

- якщо $\varphi_1 = 0$, Y_t – це білий шум;
- якщо $\varphi_1 = 1$ і $c = 0$, Y_t – це випадкове блукання;
- якщо $\varphi_1 = 1$ і $c \neq 0$, Y_t – це випадкове блукання зі зміщенням;
- якщо $\varphi_1 < 0$ і $c = 0$, Y_t має тенденцію коливатися навколо середнього.

Зазвичай авторегресійні моделі обмежуються стаціонарними даними, і в цьому випадку потрібні деякі обмеження на значення параметрів:

- для AR(1) моделі: $-1 < \varphi_1 < 1$;
- для AR(2) моделі: $-1 < \varphi_2 < 1$, $\varphi_1 + \varphi_2 < 1$, $\varphi_1 - \varphi_2 < 1$ [25].

Y_t називається процесом ковзного середнього порядку q , MA(q), якщо він задовольняє умову

$$Y_t = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q}, \quad (2)$$

де ε_t – білий шум і θ_u – коефіцієнти параметрів [25].

МА модель зазвичай використовують для моделювання випадкових помилок часових рядів. На практиці легко відрізнити серії МА та AR за поведінкою їхніх загальних кореляційних діаграм або Auto Correlation Function (ACF). МА ACF різко обривається, тоді як AR ACF спадає експоненціально. Також варто зазначити, що кінцева модель AR еквівалентна нескінченній моделі МА, а кінцева модель МА еквівалентна моделі нескінченної AR.

Y_t називається авторегресивним процесом ковзного середнього порядку (p,q), ARMA(p,q), якщо він задовольняє умову:

$$Y_t = \varphi_1 Y_{t-1} + \dots + \varphi_p Y_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q}. \quad (3)$$

ARMA модель узагальнює дві вищеописані моделі. Вона використовується для прогнозування стаціонарних часових рядів. Простими словами, часовий ряд є стаціонарним, якщо його поведінка не змінюється з часом. Це означає, наприклад, що значення завжди мають тенденцію змінюватися приблизно на одному рівні і що їхня мінливість є постійною в часі. Проте більшість часових рядів, які ми спостерігаємо, є нестаціонарними.

При обробці інформації про поведінку часових рядів необхідно враховувати, що методи аналізу нестаціонарних випадкових процесів суттєво відрізняються від прийомів роботи зі стаціонарними випадковими рядами часу. До методів, які можуть використовуватися при такому аналізі, відноситься застосування лінійної стохастичної моделі авторегресії і проінтегрованого ковзного середнього (AutoRegressive Integrated Moving Average, ARIMA).

Основна ідея моделі ARIMA – обробляти послідовність даних, сформовану передбаченим об'єктом з часом, як випадкову послідовність, і використовувати певну математичну модель, щоб описати цю послідовність. Як тільки ця модель ідентифікована, майбутні значення можуть бути передбачені з минулих та поточних значень часового ряду. Тобто, для побудови моделі ARIMA достатньо використовувати інформацію, що міститься в самих аналізованих даних часового ряду.

ARIMA(p,d,q) є розширенням моделі ARMA(p,q) для нестационарних часових рядів, що зводяться до стаціонарним взяттям різниць d-го порядку. Якщо $W_t = \Delta^d Y_t$ – це ряд ARMA(p,q), тоді Y'_t називають інтегрованим рядом авторегресивного ковзного середнього ARIMA(p,d,q). Повну модель можна записати так:

$$Y'_t = \varphi_1 Y'_{t-1} + \dots + \varphi_p Y'_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q}. \quad (4)$$

ARIMA – це абревіатура, що розшифровується як AutoRegressive Integrated Moving Average. Ця абревіатура є описовою, фіксуючи ключові аспекти самої моделі:

- AR – авторегресія (Autoregression). Модель, яка використовує залежний зв'язок між спостереженням і деяким числом спостережень із затримкою.
- I – інтегрований (Integrated). Використання диференціації вихідних спостережень, щоб зробити часовий ряд стаціонарним.
- MA – ковзне середнє. Модель, яка використовує залежність між спостереженням та залишковими похибками від моделі ковзного середнього, що застосовується до спостережень із запізненням.

Кожен з цих компонентів явно вказується в моделі як параметр. Використовується стандартне позначення ARIMA(p,d,q), де параметри замінюються цілочисельними значеннями, щоб швидко вказати конкретну модель ARIMA, що використовується.

Параметри моделі ARIMA визначаються таким чином:

- p – кількість спостережень у моделі із затримкою, також називається порядком затримки;
- q – розмір вікна ковзного середнього, який також називають порядком ковзного середнього;
- d – порядок диференціації, необхідний для того, щоб часовий ряд був стаціонарним [15].

Методологія побудови ARIMA-моделі для досліджуваного часового ряду за підходом Бокса-Дженкінса включає наступні основні етапи:

- ідентифікація пробної моделі;
- оцінювання параметрів моделі

- діагностична перевірка адекватності моделі [13].

Першим кроком є ідентифікація моделі ARIMA для даного часового ряду. На цьому етапі проводиться аналіз автокореляційної функції (ACF – autocorrelation coefficient function) і часткової автокореляційної функції (PACF – partial autocorrelation coefficient function). Ці дві функції дають цінну інформацію щодо часового ряду. Функція коефіцієнта автокореляції визначає, як точки даних у часовому ряді пов'язані з попередніми точками даних. Функція часткового коефіцієнта автокореляції, як і функція автокореляції, передає важливу інформацію щодо структури залежності стаціонарного процесу. Наприклад, швидке затухання значень ACF свідчить про стаціонарність часового ряду. В іншому випадку часовий ряд не є стаціонарним.

На цьому етапі використовуються також статистичні тести для перевірки нульової гіпотези про наявність одиничного кореня в авторегресійній моделі (тест Дікі-Фуллера або ADF-тест). Гіпотези ADF-тесту:

- Нульова гіпотеза для цього тесту полягає в тому, що існує одиничний корінь.
- Альтернативна гіпотеза полягає в тому, що часовий ряд є стаціонарним (або стаціонарним за трендом) [27].

Якщо р-значення з тесту менше деякого рівня значущості (наприклад, $\alpha = 0,05$), то ми можемо відхилити нульову гіпотезу і зробити висновок, що часовий ряд є стаціонарним.

Після ідентифікації відповідних параметрів p , d та q , модель ARIMA може бути оцінена на основі історичних даних. Для цього використовується метод максимальної правдоподібності, який дозволяє знайти оптимальні значення параметрів моделі ARIMA для максимізації шансу того, що модель підійде до ряду.

Останнім етапом є діагностика моделі, щоб переконатися, що модель є адекватною для прогнозування. Це включає перевірку залишкової автокореляції, перевірку на гетероскедастичність та на нормальність залишків. Якщо залишки

моделі є нормально розподіленими, мають низький рівень автокореляції та не є гетероскедастичними, то модель можна вважати адекватною.

Після побудови ARIMA-моделі за допомогою підходу Бокса-Дженкінса можна використовувати цю модель для прогнозування майбутніх значень часового ряду. Для цього необхідно використовувати отримані параметри моделі та додаткові дані про майбутній період.

Крім прогнозування, з ARIMA-моделлю можна проводити аналіз часового ряду, щоб отримати додаткову інформацію про його характеристики та особливості. Наприклад, можна аналізувати тенденції, сезонність, циклічність, стаціонарність та автокореляцію часового ряду. Це дозволяє отримати більш глибоке розуміння процесу, який моделюється, та зробити більш обґрунтовані прогнози.

2.3. Розширена модель ARIMA (ARIMAX)

ARIMAX (Autoregressive Integrated Moving Average with Explanatory Variables) – це модель для аналізу часових рядів, яка розширює ARIMA-модель, дозволяючи включати додаткові пояснювальні змінні, які можуть впливати на часовий ряд. Модель ARIMAX була вперше розглянута Боксом і Тіао в 1975 році, як здатність ідентифікувати основні закономірності в даних часових рядів і кількісно оцінити вплив зовнішніх факторів [14].

ARIMAX дозволяє враховувати вплив змінних, що не входять до самого ряду, на його поведінку. Це дозволяє збільшити точність прогнозування і зробити більш точні висновки про те, що впливає на поведінку ряду. Рівняння ARIMAX(p,q,d) виглядає наступним чином:

$$Y'_t = \beta x_t + \varphi_1 Y'_{t-1} + \dots + \varphi_p Y'_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q}. \quad (5)$$

Два нових доповнення до моделі – x_t і його коефіцієнт β . Так само, як φ , β є коефіцієнтом, який буде оцінено на основі вибору моделі та даних [30]. x_t – це екзогенна змінна, і це може бути будь-яка змінна, яка цікава для дослідження. У моделі ARIMAX, екзогенні змінні додаються до моделі ARIMA як додаткові вхідні параметри, що допомагають краще моделювати основний часовий ряд, враховуючи залежності від зовнішніх факторів. Це може бути вимірювання, що змінюється в часі,

наприклад рівень інфляції або ціна іншого індексу. Або категоріальна змінна, що розділяє різні дні тижня. Нарешті, це може означати поєднання кількох різних зовнішніх факторів.

Ідея полягає в тому, що це може бути будь-яка інша змінна або змінні, які можуть впливати на ціни, якщо у нас є доступні дані. Такі зовнішні фактори відомі як екзогенні змінні.

2.4. Пакет Prophet

Особливістю моделі лінійної регресії є те, що прогноз моделюється як зважена сума ознак. Крім того, лінійна модель має багато інших припущень. Всі ці припущення часто порушуються в реальності.

GAM – це модель, яка дозволяє лінійній моделі вивчати нелінійні зв'язки. Вона передбачає, що замість використання простих зважених сум можна використовувати суму довільних функцій. GAM була розроблена Тревором Хасті та Робертом Тібшірані для об'єднання якостей лінійних моделей з адитивними, тобто GAM є розширенням лінійних моделей. Як відомо, формула лінійної регресії:

$$y = \beta_0 + \beta_1 x_1 + \dots + \beta_n x_n + \epsilon$$

Це передбачає, що зважена сума з деякою помилкою ϵ виражає результат y .

У GAM відкидається припущення, що ціль може бути розрахована за допомогою лінійної комбінації змінних, просто кажучи, що можна використовувати нелінійну комбінацію змінних, позначену s , для «гладкої функції»:

$$y = s_0 + s_1 x_1 + \dots + s_n x_n + \epsilon, \text{ де} \\ s(x) = \sum_{k=1}^k \beta_k b_k(x) \quad (5)$$

У формулі (5) бачимо, що β позначає ті ж ваги, що й у лінійній регресії, а b – розширення базису. Тобто можна мати k вагів і функції для кожної змінної в рівнянні. Це набагато гнучкіше і набагато менш лінійно, ніж у лінійній регресії [37].

Тобто GAM – це метод адитивного моделювання, в якому вплив прогностичних змінних фіксується за допомогою гладких функцій, які залежно від закономірностей, що лежать в основі, в даних можуть бути нелінійними [28].

Основними перевагами GAM є те, що цю модель легко інтерпретувати, гнучкі функції провісників можуть виявити приховані закономірності в даних, регуляція функцій предиктора допомагає уникнути переобладнання. Отже, коли модель містить нелінійні ефекти, GAM надає регуляризоване та інтерпретоване рішення, у той час як в інших методах зазвичай немає принаймні однієї з цих трьох функцій. Іншими словами, GAM забезпечують хороший баланс між інтерпретованою, але упередженою лінійною моделлю та надзвичайно гнучкими алгоритмами навчання «чорної скриньки».

У R існує велика кількість пакетів для аналізу часових рядів. Наприклад, одним з найпопулярніших є пакет `forecast`, який реалізується як класичні моделі (експоненціальне згладжування, модель Хольта-Вінтерса, ARIMA тощо), і нещодавно розроблені методи прогнозування (моделі для згрупованих часових рядів, рядів з кількома сезонними компонентами та інші). Такі різноманітні методи є і перевагою, і недоліком пакету прогнозів.

У 2017 році фахівці Facebook оголосили про новий пакет, який вони розробили для прогнозування часових рядів – Prophet. Пакет Prophet дозволяє створювати точні моделі прогнозів у напівавтоматичному режимі.

Методологія Prophet заснована на процедурі коригування моделей регресії адитивних регресій (Generalized Additive Models, GAM) наступного типу:

$$y(t) = g(t) + s(t) + h(t) + \varepsilon_t, \quad (6)$$

де $g(t)$ – це функція тренду, яка моделює неперіодичні зміни у значеннях часових рядів, $s(t)$ являє собою сезонні зміни (наприклад, щотижнева та щорічна сезонність), $h(t)$ – це наслідки свят та інших важливих подій, які відбуваються нерегулярно протягом одного або декількох днів, а ε_t – нормально розподілені випадкові збурення [38]. Для апроксимації перелічених функцій використовуються такі методи:

- Тренд – кусково-лінійна регресія;
- Річна сезонність – часткова сума ряду Фур'є, кількість членів яких (порядок) визначає плавність (гладкість) функції;
- Тижнева сезонність – представлена у вигляді фіктивної змінної;

- «Свята» (наприклад, офіційні свята та вихідні – Новий рік, Різдво тощо, а також інші дні, протягом яких властивості тимчасової серії можуть значно змінюватися – спортивні чи культурні заходи, природні явища тощо): представлені у вигляді фіктивних змінних [6].

Оцінка параметрів скоригованої моделі проводиться за допомогою принципів байєсівської статистики. Пакет Prophet – це не що інше, як зручний інтерфейс для роботи в середовищі R. Окрім того, у пакету Prophet є ряд практичних переваг над моделями ARIMA та ARIMAX:

- Гнучкість: можна легко вмістити сезонність з декількома періодами, це дозволяє аналітикам робити різні припущення щодо тенденцій.
- На відміну від моделей ARIMA (ARIMAX), Prophet є нечутливим до пропущених значень, зміщень у тренді і значних викидів.
- Пристосування дуже швидке, що дозволяє аналітику інтерактивно досліджувати багато специфікацій моделі.
- Модель прогнозування легко інтерпретується параметри, які аналітиком може бути змінено, щоб накласти припущення на прогноз. Більше того, аналітики, здебільшого, мають досвід регресії і легко можуть розширити модель, щоб включити нові компоненти.

Отже, для аналізу часових рядів використовується багато різних моделей та пакетів, основними з них, які були розглянуті у цьому розділі, є моделі ARIMA, ARIMAX а також пакет Prophet. Далі перейдемо безпосередньо до аналізу впливу соціальних мереж на курс криптовалют з використанням розглянутих моделей.

РОЗДІЛ III

АНАЛІЗ ВПЛИВУ СОЦІАЛЬНОЇ МЕРЕЖІ TWITTER НА BITCOIN

3.1. Підготовка даних

У роботі проведено збір та підготовку даних зі спеціалізованого джерела – сайту Kaggle [8]. Набір даних містить наступну інформацію:

- Date: Часовий проміжок, до якого відноситься набір даних.
- Compound_Score: Середнє значення усіх оцінок настрою. Це значення досить мале, оскільки кількість нейтральних оцінок зменшує вплив настрою.
- n: Загальний обсяг твітів, пов'язаних з Bitcoin.
- Count_Negatives: Загальний обсяг негативних твітів.
- Count_Positives: Загальний обсяг позитивних твітів.
- Count_Neutrals: Загальний обсяг нейтральних твітів.
- Sent_Negatives: Середнє значення настрою негативних твітів.
- Sent_Positives: Середнє значення настрою позитивних твітів.
- Open: Початкова ціна Bitcoin.
- High: Найвища ціна Bitcoin протягом періоду (1 година).
- Low: Найнижча ціна Bitcoin протягом періоду (1 година).
- Close: Остання ціна Bitcoin протягом періоду (1 година).
- Volume (BTC): Обсяг Bitcoin.
- Volume (Currency): Обсяг Bitcoin у доларах США.

Ці дані надають можливість аналізувати залежність між кількістю та настроєм твітів, що стосуються Bitcoin, та його ціною на ринку. Цей набір даних є корисним для вивчення впливу соціальних мереж на курс Bitcoin та розробки прогностичних моделей.

На початковому етапі, отримані дані було відображено для огляду та аналізу за допомогою функції View() у середовищі програмування R. Цей крок дозволив здійснити перші висновки та попередню оцінку структури та змісту даних.

Після огляду даних, було виконано процес підготовки для подальшого аналізу. У цьому контексті, важливо було вирішити проблеми відсутності значень та некоректних даних. Для цього було використано функцію `mutate()` з пакету `dplyr`, що дозволило замінити відсутні значення (NA) на відповідні значення з попередніх спостережень (функція `na.locf()`). Крім того, для покращення обробки даних, було виконано перейменування стовпців з метою покращення зрозумілості та зручності подальшого використання.

Наступним кроком було перетворення стовпця дати з текстового формату в формат `POSIXct` (часовий тип даних) для забезпечення правильного подальшого аналізу та візуалізації часових рядів. Це було досягнуто за допомогою функції `as.POSIXct()`, в якій також було встановлено часовий пояс «UTC» для забезпечення однорідності даних.

Інтелектуальний аналіз даних також включав вибір необхідних стовпців з підготованих даних. З метою використання лише необхідних змінних для подальшого аналізу, було створено новий набір даних з відповідними стовпцями. Використання функції `select()` з пакету `dplyr` дозволило вибрати стовпці, що містять кількість твітів (`tweets`), кількість негативних твітів (`neg_tweets`), кількість позитивних твітів (`pos_tweets`), кількість нейтральних твітів (`neut_tweets`), ціну (`price`) та дату (`date`) для подальшого використання.

З метою забезпечення зручності та зрозумілості, було також перейменовано стовпці нового набору даних. Це було зроблено за допомогою функції `colnames()`, де кожному стовпцю було присвоєно нове інтуїтивно зрозуміле ім'я.

Отже, на даному етапі інтелектуального аналізу даних було здійснено збір даних з вказаного джерела та їх підготовку для подальшого використання. Застосування відповідних функцій та методів дозволило розпочати наступні етапи аналізу, такі як прогнозування та моделювання залежностей у досліджуваних даних.

3.2. Візуальний аналіз

Для подальшого вивчення зв'язку між твітами та ціною Bitcoin було здійснено візуалізацію даних за допомогою графіків. На відображених графіках (рис. 3.1-3.3) представлено залежність між роками та кількістю твітів, що згадують Bitcoin. Кожен

графік відображає певний тип твітів, надаючи уявлення про загальну кількість, негативність, позитивність та нейтральність твітів протягом років. Графіки використовуються для аналізу динаміки цих показників з плином часу.

Зокрема, графік «Загальна кількість твітів» на рис. 3.1 відображає загальну активність користувачів у спілкуванні про Bitcoin протягом років. Цей графік дозволяє оцінити загальний обсяг обговорень Bitcoin та спостерігати зміни цієї активності протягом часу.

Найбільша активність обговорень про Bitcoin спостерігалася в кінці 2017 року та на початку 2018 року. Цей період був помітним для криптовалюти Bitcoin, оскільки саме тоді відбулось значне зростання ціни Bitcoin. Проте згодом спостерігався спад кількості твітів про Bitcoin. З рис. 3.1 можна побачити, що після піку в кінці 2017 – на початку 2018 року, кількість твітів поступово зменшувалася. Цей тренд вказує на зниження інтересу та обговорень про Bitcoin в мережі соціальних медіа.

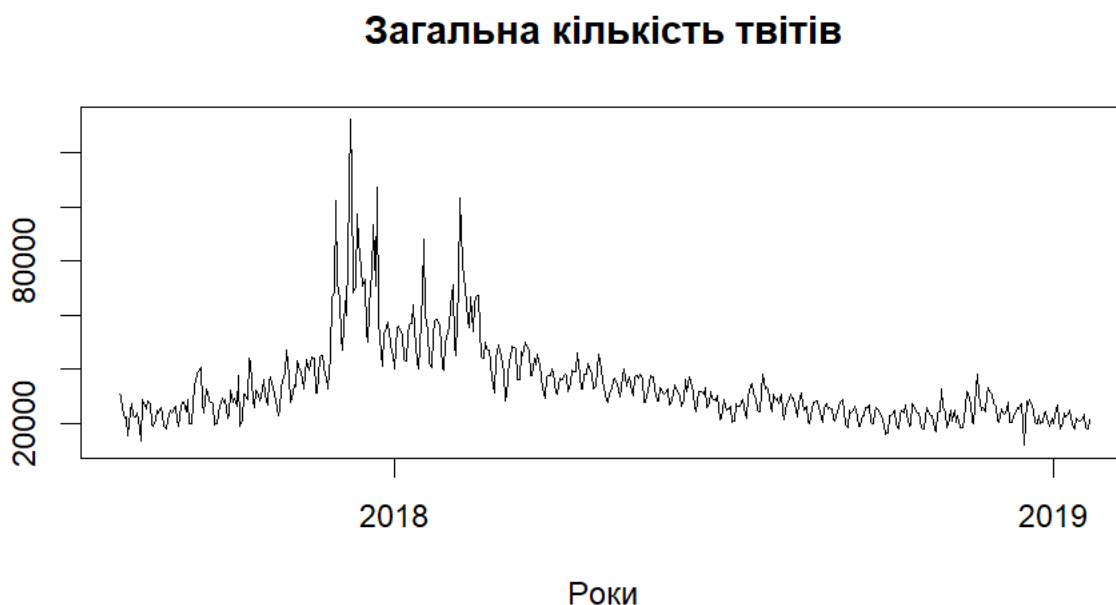


Рис. 3.1. Загальна кількість твітів за період 2017-2019 років. Джерело: створено автором на основі [8].

Додатково, графіки «Негативні твіти», «Позитивні твіти» та «Нейтральні твіти» показують розподіл твітів, що згадують Bitcoin, за їхнім відношенням до негативних, позитивних та нейтральних відтінків. Ці графіки надають можливість оцінити

емоційний спектр обговорень, пов'язаних з Bitcoin, та зрозуміти, які емоції переважають у спілкуванні користувачів.

На рис. 3.2. відображено зміну кількості позитивних обговорень про Bitcoin протягом досліджуваного періоду. Виявлено, що пік активності позитивних твітів також спостерігався в кінці 2017 року та на початку 2018 року, разом із загальним зростанням обговорень про Bitcoin.

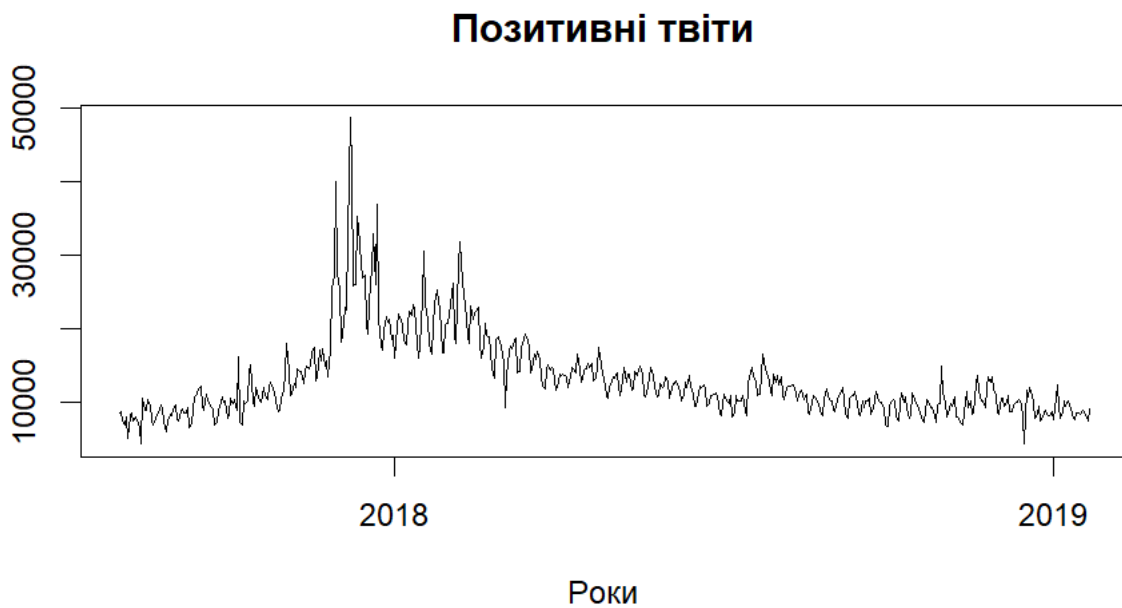


Рис. 3.2. Загальна кількість позитивних твітів за період 2017-2019 років. Джерело: створено автором на основі [8].

На рис. 3.3 відображено зміну кількості негативних обговорень про Bitcoin протягом досліджуваного періоду. Спостерігається, що негативні твіти про Bitcoin були присутні протягом всього аналізованого періоду, але їх інтенсивність коливалася.

На піку активності, який спостерігався в кінці 2017 року та на початку 2018 року, кількість негативних твітів про Bitcoin була значно вищою. Це може бути пов'язано зі зростанням ціни Bitcoin та виникненням ейфорії, а також зі збільшеним інтересом до криптовалюти в цей період. Однак, разом з подальшим зниженням ціни та змінами на ринку, кількість негативних твітів про Bitcoin почала зменшуватися. Це свідчить

про те, що негативні настрої та критика щодо Bitcoin можуть залежати від його цінової динаміки та ринкової ситуації.

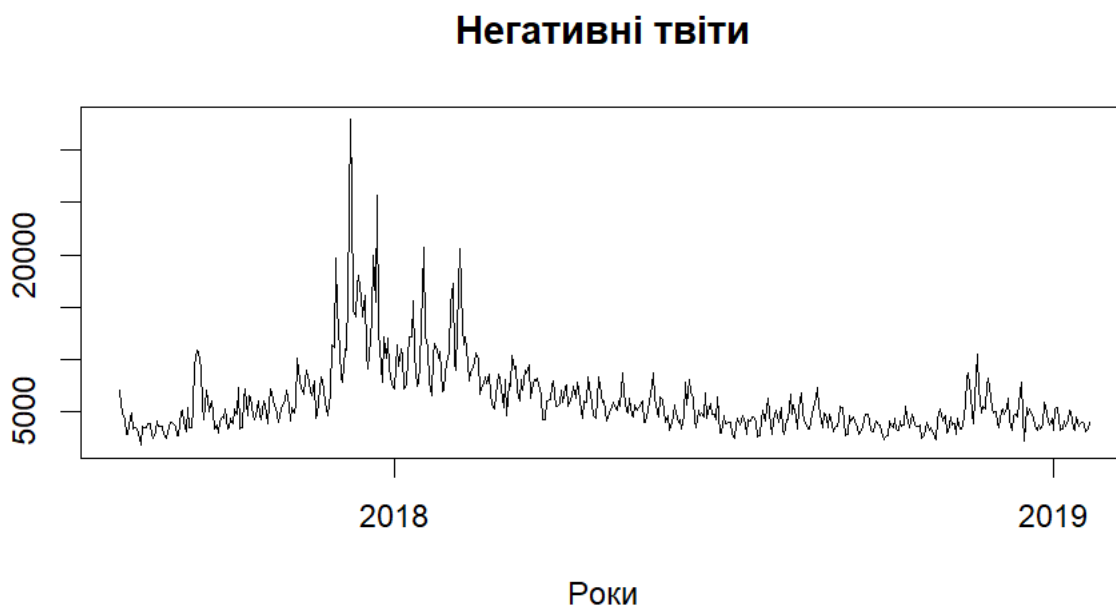


Рис. 3.3. Загальна кількість негативних твітів за період 2017-2019 років. Джерело: створено автором на основі [8].

На рис. 3.4 представлена динаміка кількості нейтральних обговорень про Bitcoin протягом досліджуваного періоду. Нейтральні твіти відображають загальний обсяг повідомлень, які не виражають чіткого позитивного або негативного ставлення до Bitcoin. На графіку можна побачити, що кількість нейтральних твітів була значною протягом великої частини аналізованого періоду. Це свідчить про те, що багато обговорень про Bitcoin були нейтрального характеру і не містили явного позитивного або негативного відношення до цієї криптовалюти. Нейтральні твіти можуть включати різноманітну інформацію, новини, статистику, аналітику та загальні спостереження про Bitcoin без явно вираженого ставлення до його ціни або перспектив.

У результаті аналізу динаміки цін та кількості твітів, було виявлено, що існує співпадіння між цими двома змінними. Графіки показують, що під час підвищення цін на Bitcoin в кінці 2017 року та на початку 2018 року, спостерігалось значне зростання кількості твітів, що стосуються цього криптовалютного активу. Подальше

зниження цін Bitcoin супроводжувалося зменшенням кількості твітів. Це свідчить про тісну залежність між ціною Bitcoin і інтересом користувачів в соціальних медіа.

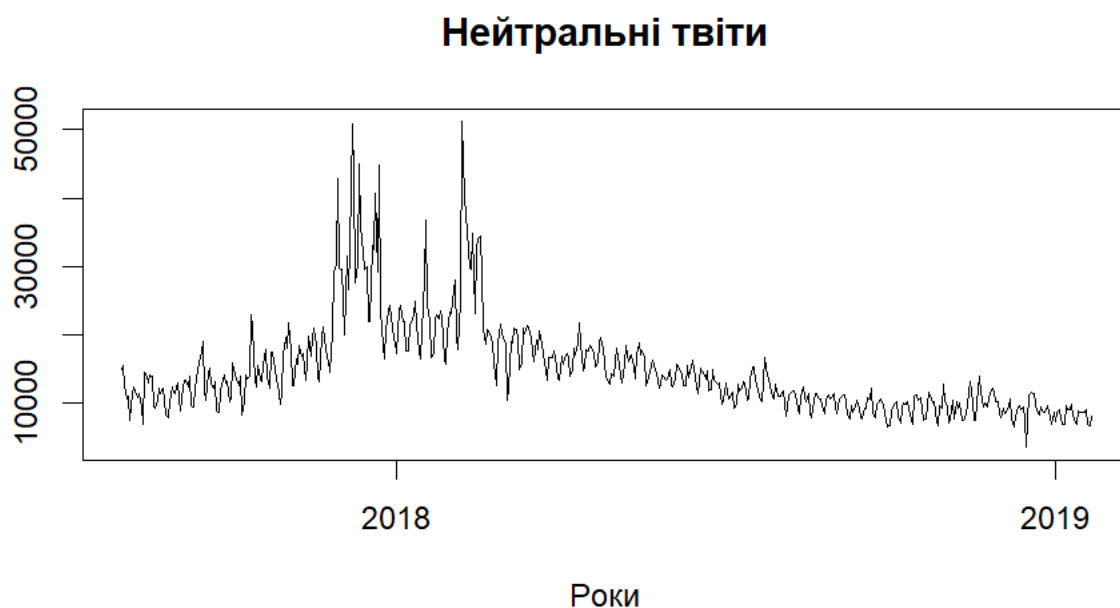


Рис. 3.4. Загальна кількість нейтральних твітів за період 2017-2019 років. Джерело: створено автором на основі [8].

3.3. Аналіз за допомогою ARIMA

Для побудови моделі ARIMAX, як зазначалося в розділі II, необхідно протестувати ряд на стаціонарність. Для цього скористаємося ACF і PACF тестами, а також тестом Дікі-Фуллера.

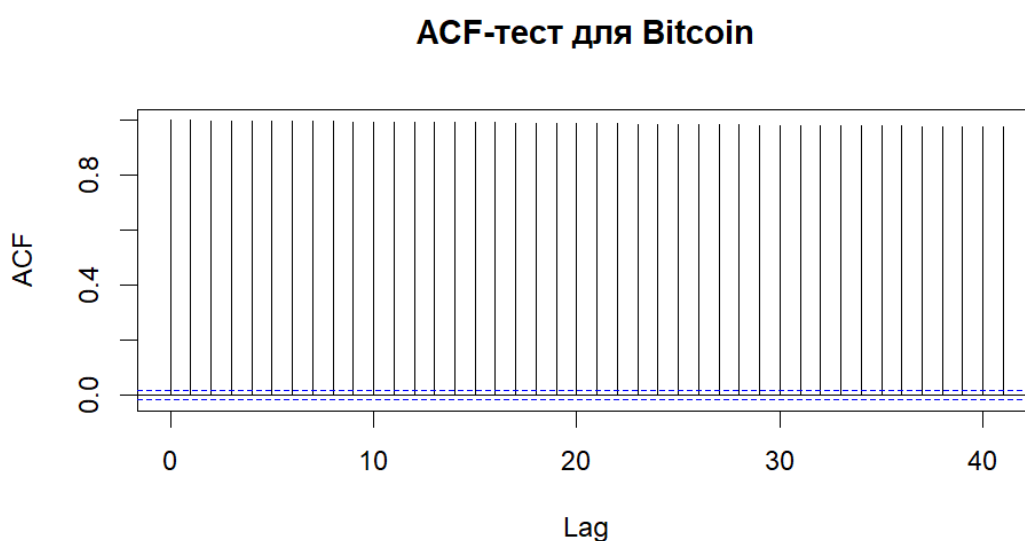


Рис. 3.5. Результати ACF-тесту для Bitcoin. Джерело: розрахунки автора.

На рис. 3.5 та рис. 3.6 зображено результати тестів. Оскільки значення ACF на графіку практично не затухають, це свідчить про наявність сильної автокореляції у ряді. Автокореляція вказує на залежність між значеннями ряду в різних часових періодах.

Оскільки значення ACF не затухають повільно, а нестационарний ряд часто супроводжується сильною автокореляцією, можна зробити висновок, що ряд вірогідно є нестационарним. Нестационарний ряд характеризується зміною статистичних властивостей з часом, таких як середнє значення, дисперсія або автокореляція.

Нестационарний ряд може мати тренд, що означає постійне зростання або спад значень з часом, або сезонність, коли певні патерни повторюються в ряді на регулярних інтервалах. Для аналізу такого ряду можуть бути застосовані методи диференціювання, щоб зробити його стаціонарним, або моделі, які враховують нестационарність, наприклад, моделі ARIMAX з інтегрованим компонентом. Такий ж висновок можна зробити, проаналізувавши результат тесту ADF на рис. 3.7, оскільки p-value становить 0,3898, що є більшим за 0,05.

PACF-тест для Bitcoin

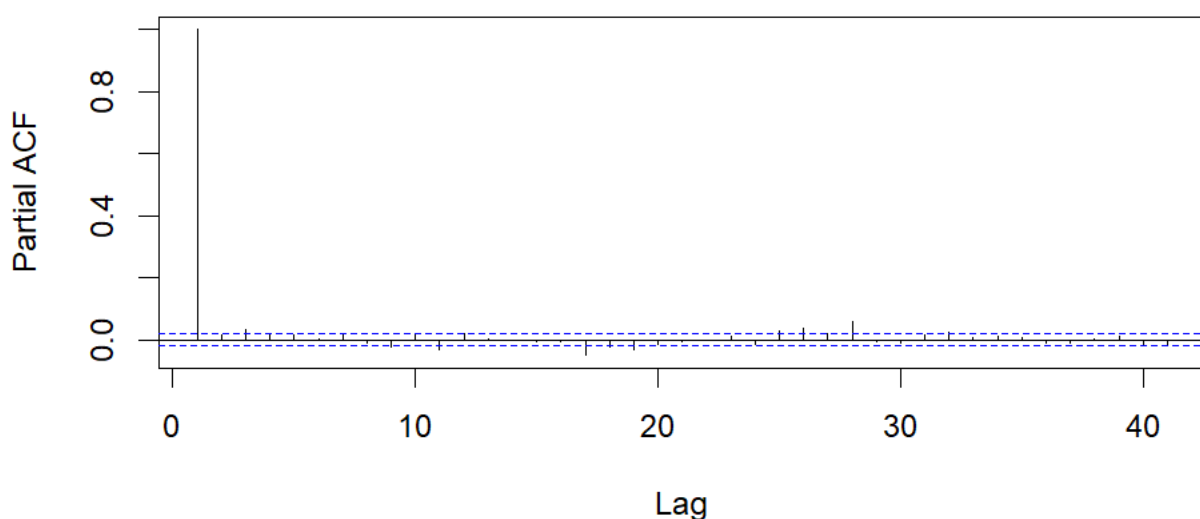


Рис. 3.6. Результати PACF-тесту для Bitcoin. Джерело: розрахунки автора.

Augmented Dickey-Fuller Test

```
data: data$price
Dickey-Fuller = -2.4451, Lag order = 23, p-value = 0.3898
alternative hypothesis: stationary
```

Рис. 3.7. Результати тесту ADF для Bitcoin. *Джерело: розрахунки автора.*

Для проведення аналізу було вирішено розділити дані на тренувальну і тестову множини. Тренувальна множина використовується для побудови моделі ARIMAX, що дозволяє аналізувати та прогнозувати часовий ряд залежно від інших змінних, у нашому випадку – кількості твітів.

Розділення на тренувальну і тестову множини має на меті перевірити адекватність та точність побудованої моделі. Тренувальна множина використовується для навчання моделі на доступних даних, в той час як тестова множина використовується для перевірки, наскільки добре модель прогнозує значення на нових, не використаних раніше даних.

У контексті ARIMAX, розділення на тренувальну і тестову множини дозволяє перевірити, наскільки добре модель враховує залежність між цільовим часовим рядом (цінами Bitcoin) та екзогенними змінними (кількістю твітів з різними настройками). Тестування на незалежних даних допомагає оцінити точність та ефективність моделі перед її застосуванням для прогнозування майбутніх значень.

Для аналізу впливу кількості твітів на динаміку ціни Bitcoin спочатку було побудовано автоматичну модель ARIMA за допомогою функції `auto.arima()`, в яку були включені регресійні змінні `regressors_train` у вигляді матриці. Отримані результати та характеристики моделі були детально проаналізовані за допомогою функції `summary()`, яка надає статистичну інформацію про показники моделі, включаючи коефіцієнти, стандартні помилки, критерії інформаційної втрати (AIC, AICc, BIC) та інші. Такий підхід дозволяє здійснити аналіз якісних та кількісних характеристик моделі та оцінити її адекватність для прогнозування динаміки ціни Bitcoin.

На рис. 3.8 представлені результати аналізу за допомогою ARIMA моделі з регресійними факторами. Модель ARIMA(1,1,2) з показниками авторегресії (AR)

рівня 1, рівня диференціювання (I) рівня 1 та ковзного середнього (MA) рівня 2 була використана для оцінки ціни Bitcoin.

Коефіцієнти моделі вказують на вплив різних регресорів на ціну Bitcoin. За результатами аналізу, коефіцієнти показують, що зростання кількості твітів (tweets) має позитивний вплив на ціну Bitcoin, тоді як зростання кількості негативних твітів (negatives) та позитивних твітів (positives) мають негативний вплив. Коефіцієнт для регресора нейтральних твітів (neutrals) також показує негативний вплив на ціну Bitcoin.

```
Regression with ARIMA(1,1,2) errors

Coefficients:
      ar1      ma1      ma2 tweets negatives positives
0.5032 -0.5258 -0.0301 0.5287  -0.5704  -0.5068
s.e. 0.1165 0.1167 0.0117 0.3393 0.3393 0.3380
      neutrals
      -0.5345
s.e. 0.3389

sigma^2 = 11884: log likelihood = -63226.86
AIC=126469.7 AICc=126469.7 BIC=126527.7

Training set error measures:
              ME      RMSE      MAE      MPE      MAPE
Training set 0.4088518 108.9696 58.78229 0.002712816 0.6768704
              MASE      ACF1
Training set 0.9967611 0.0001636589
```

Рис. 3.8. Результат аналізу за моделлю ARIMAX. Джерело: розрахунки автора.

Загальна адекватність моделі підтверджується низькими значеннями помилок тренувальної вибірки, такими як середня помилка (ME), середньоквадратичне відхилення (RMSE) та середня абсолютна помилка (MAE). Коефіцієнт детермінації (ACF1), наближаючись до нуля, свідчить про відсутність автокореляції в помилках моделі, що також підтверджує адекватність моделі.

Також були обчислені прогнози значення на основі побудованої ARIMAX моделі за допомогою функції `forecast`. Прогноз був зроблений з використанням регресійних змінних `regressors_test` у вигляді матриці. Результати відображені на рис. 3.9. Важливим спостереженням є те, що реальні дані про ціну Bitcoin на майбутній період відповідають негативній частині прогнозу ARIMAX моделі.

Коефіцієнт детермінації (R^2) між прогнозованими значеннями ціни Bitcoin і фактичними значеннями ціни Bitcoin становить 0,9988. Це високе значення R^2 , що вказує на дуже сильну залежність між прогнозованими і спостережуваними значеннями ціни Bitcoin, а також свідчить про високу точність прогнозу моделі ARIMA(1,1,2) з використанням регресорів.

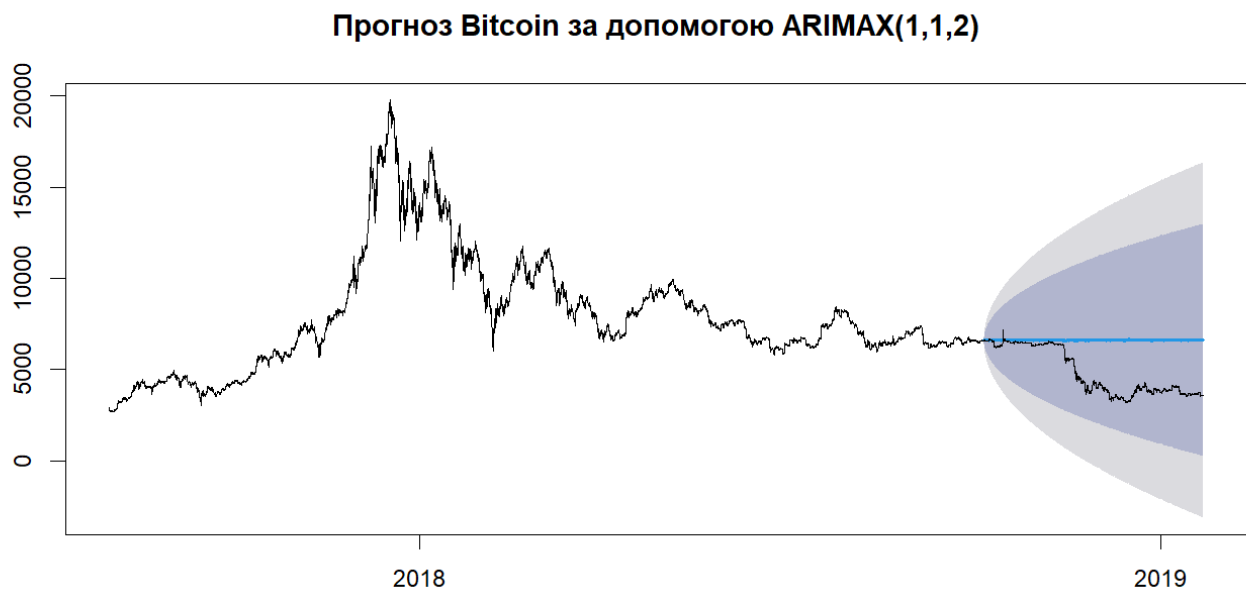


Рис. 3.9. Прогноз курсу Bitcoin з використанням моделі ARIMAX(1,1,2).
Джерело: розрахунки автора.

Для ARIMAX моделі був проведений тест Вох-Ljung для перевірки автокореляції залишків моделі (рис. 3.10). Результати тесту показали значення статистики X-squared рівним 0,00027727 з 1 ступенем свободи та p-value рівним 0,9867.

Високе значення p-value свідчить про те, що нульова гіпотеза про відсутність автокореляції залишків не відхиляється на рівні значущості 0,05. Це означає, що залишки моделі ARIMAX є незалежними та неавтокорельованими. Такий результат свідчить про адекватність моделі ARIMAX, оскільки відсутність автокореляції залишків вказує на те, що модель коректно врахувала динаміку цінового ряду та використовувала наявну інформацію для прогнозування.

Box-Ljung test

```
data: fcst$residuals
X-squared = 0.00027727, df = 1, p-value = 0.9867
```

Рис. 3.10. Результат тесту Box-Ljung для перевірки автокореляції залишків моделі ARIMAX(1,1,2). Джерело: розрахунки автора.

Це підтверджує, що розроблена модель ARIMA(1,1,2) з регресійними факторами є придатною для прогнозування ціни Bitcoin. Залежність між кількістю твітів, негативних твітів, позитивних твітів і нейтральних твітів та ціною Bitcoin була успішно врахована в моделі, що дозволяє отримати точний і надійний прогноз для майбутніх значень ціни Bitcoin.

3.4. Аналіз за допомогою Prophet

Далі розглянемо альтернативний підхід до аналізу з використанням моделі Prophet. Модель Prophet є потужним інструментом прогнозування, який дозволяє враховувати явні сезонність, тренди і зміну трендів у часових рядах. На відміну від ARIMAX моделі, яка базується на статистичних методах, модель Prophet заснована на підході машинного навчання і дозволяє автоматично виявляти та моделювати складні залежності в даних.

Спочатку був створений датафрейм `btc_df`, що містить необхідні змінні, включаючи дату (`ds`), ціну Bitcoin (`y`) та регресійні фактори. Наступним кроком було створено модель Prophet (`model_pr`) та додано регресійні фактори за допомогою функції `add_regressor()`. Після цього модель було підігнано за допомогою функції `fit.Prophet()`.

Для отримання прогнозів було створено майбутній період (`future`) за допомогою функції `make_future_dataframe()`. У цьому періоді були враховані значення регресійних факторів (`n`, `negative_tweets`, `neutral_tweets`, `positive_tweets`) з вихідного датасету (`btc_df`).

Завершальним кроком було отримання прогнозів для ціни Bitcoin за допомогою функції `predict()`. Результати аналізу моделі Prophet представлені на рис. 3.11, , з чого можна зробити висновок, що загальна тенденція курсу є більш низхідною.

На рис. 3.12 видно, що модель добре передає тренд, що є в даних. Видно також, що в цьому часовому ряду є виражені внутрішньорічні коливання та коливання в межах дня і тижня.

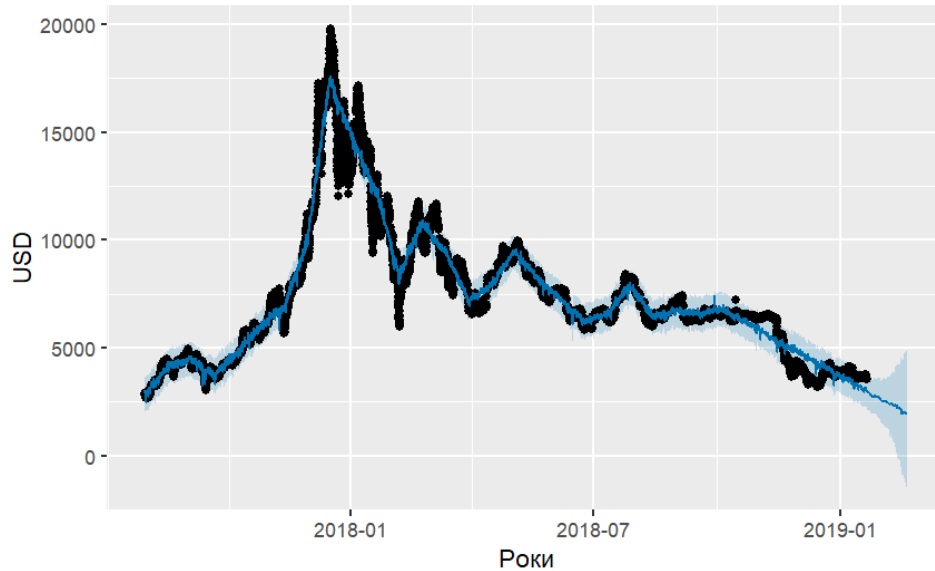


Рис. 3.11. Прогноз Bitcoin за допомогою Prophet. Джерело: розрахунки автора.

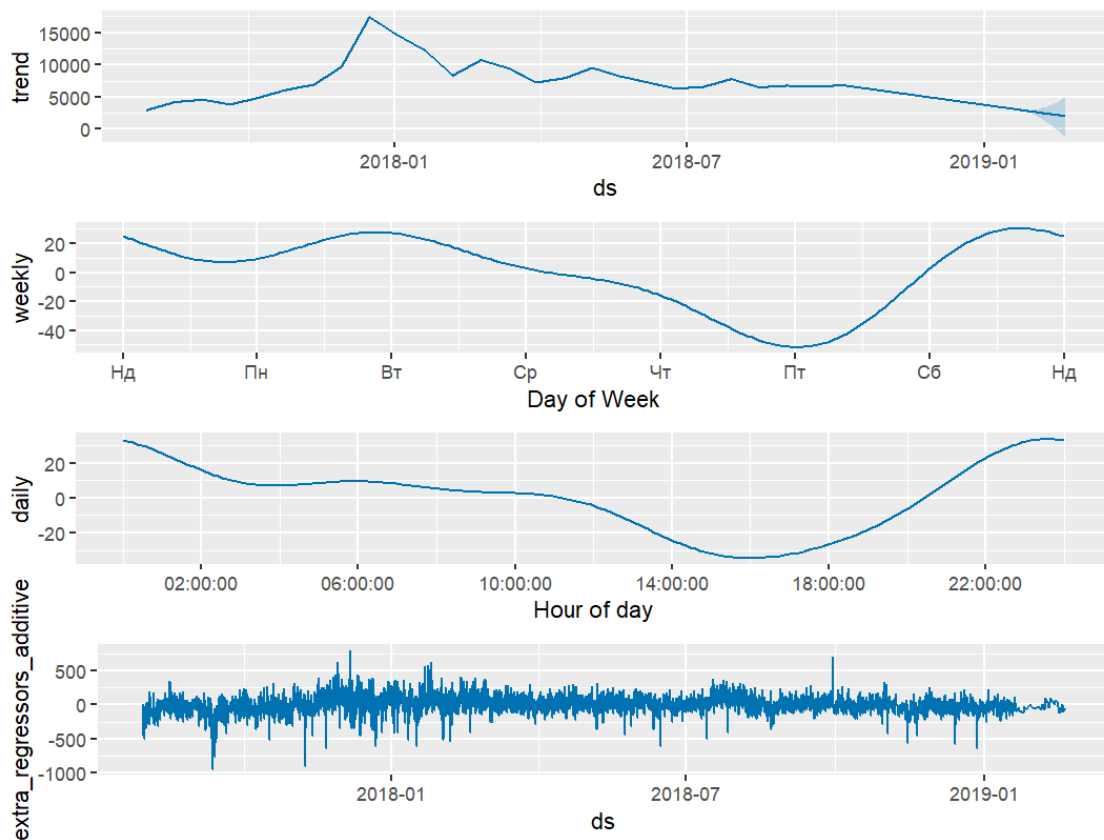


Рис. 3.12. Компоненти моделі Prophet. Джерело: розрахунки автора.

Коефіцієнти моделі Prophet показують вплив кожного регресора на прогнозоване значення цільової змінної. Основна модель прогнозування у моделі Prophet є

адитивною, тому кожен регресор додає свою власну складову до прогнозованого значення. Коефіцієнти моделі Prophet для регресорів відображені на рис. 3.13 та виглядають наступним чином:

- Загальна кількість твітів. Центральне значення регресора становить 1428,8384. Коефіцієнт показує, що кожне збільшення величини регресора на одиницю відповідає зниженню прогнозованого значення цільової змінної на 0,0333 одиниць.
- Кількість негативних твітів. Центральне значення регресора становить 255,7542. Кожне збільшення цього регресора на одиницю призводить до зниження прогнозованого значення цільової змінної на 0,0018 одиниць.
- Кількість позитивних твітів. Центральне значення регресора становить 549,7411. Кожне збільшення цього регресора на одиницю відповідає збільшенню прогнозованого значення цільової змінної на 0,0243 одиниць.
- Кількість нейтральних твітів. Центральне значення регресора становить 621,7476. Збільшення цього регресора на одиницю призводить до збільшення прогнозованого значення цільової змінної на 0,0129 одиниць.

	regressor	regressor_mode	center	coef_lower	coef	coef_upper
1	n	additive	1428.8384	-0.033290139	-0.033290139	-0.033290139
2	negative_tweets	additive	255.7542	-0.001821371	-0.001821371	-0.001821371
3	positive_tweets	additive	549.7411	0.024309214	0.024309214	0.024309214
4	neutral_tweets	additive	621.7476	0.012913667	0.012913667	0.012913667

Рис. 3.13. Коефіцієнти моделі Prophet. Джерело: розрахунки автора.

Модель Prophet демонструє низькі значення RMSE (616,4552), MAE (434,8097) та MAPE (0,0625), що свідчить про невеликі відхилення прогнозованих значень від фактичних значень (табл. 3.1). Це говорить про те, що модель добре адаптується до тренувальних даних та здатна точно передбачати майбутні значення цільової змінної. Таким чином, результати аналізу помилок підтверджують адекватність моделі Prophet у прогнозуванні цільової змінної на основі регресорів.

Для моделі Prophet був проведений тест Вох-Ljung для перевірки автокореляції залишків моделі (рис. 3.14). Результати тесту показали значення статистики X-squared рівним 12345 з 1 ступенем свободи та p-value менше за $2.2e-16$. Низьке значення p-

value свідчить про відхилення нульової гіпотези про відсутність автокореляції залишків на рівні значущості 0,05. Це означає, що залишки моделі Prophet є автокорельованими, і можливо існує додаткова інформація або структура, яку модель не врахувала.

Таблиця 3.1

Значення похибок моделі з використанням пакету Prophet

Похибка	Значення
RMSE	616,4552
MAE	434,8097
MAPE	0,0625
MASE	8,5575

Джерело: розрахунки автора.

Box-Ljung test

```
data: btc_df$y - forecast$yhat
X-squared = 12345, df = 1, p-value < 2.2e-16
```

Рис. 3.14. Результат тесту Box-Ljung для перевірки автокореляції залишків моделі Prophet. *Джерело: розрахунки автора.*

Існує деяка різниця між двома моделями, розглянутих вище. Різниця в результатах моделей може пояснюватися кількома факторами. По-перше, ARIMA модель базується на стаціонарній природі ряду, тоді як Prophet може враховувати нестаціонарність та залежність від зовнішніх факторів. Крім того, різниця в результатах може бути викликана різними методами оцінки та підходами до моделювання, які використовуються в ARIMA та Prophet моделях.

ARIMA модель оцінюється за допомогою методу максимальної правдоподібності або методу найменших квадратів, залежно від конкретного випадку. Вона намагається знайти найкращі параметри моделі, які максимізують правдоподібність спостережень у тренувальній вибірці або мінімізують суму квадратів помилок. З іншого боку, модель Prophet використовує метод, який базується на генерації випадкових вибірок, які дозволяють знайти розподіли параметрів моделі та виконувати статистичні оцінки.

У даному розділі були досліджені та порівняні моделі ARIMAX і Prophet для прогнозування цінового ряду Bitcoin. Обидві моделі показали свою ефективність у прогнозуванні. Отримані результати свідчать про значимість соціальних мереж як важливого фактора, що впливає на курс Bitcoin. Аналіз цінового ряду Bitcoin разом з даними з соціальних мереж, такими як кількість позитивних і негативних твітів, а також загальна активність користувачів, дозволив виявити значиму кореляцію між цими факторами. Знання про активність користувачів у соціальних мережах може бути корисним для інвесторів та трейдерів, які мають інтерес до цінової динаміки Bitcoin. Подальше дослідження в цьому напрямку може допомогти розкрити більш детальний вплив соціальних мереж на курс Bitcoin та розробити ефективні інвестиційні стратегії.

ВИСНОВКИ

Нині на фінансовому ринку продовжується розвиток різноманітних фінансових інструментів, удосконалюється механізм фінансового регулювання, вводяться фінансові інновації. До однієї із відносно нещодавніх фінансових інновацій можна віднести цифрову (електронну) валюту, різновидом якої є криптовалюта.

Криптовалюта – набір концептів та технологій, які разом утворюють основу екосистеми цифрових грошей. Вона має власні характеристики, які відрізняють її як від традиційної валюти, так і від фінансового активу, і водночас роблять її схожою на них. На сьогодні криптовалюта набула великої популярності. Усьому цьому активно сприяє низка факторів. Сюди можна зарахувати зручність у користуванні, високу швидкість проведення різних транзакцій, використання сучасних технологій безпосередньо для забезпечення високого рівня безпеки угод. У роботі було розглянуто технологію блокчейн, її реалізацію. Також було описано процес купівлі-продажу криптовалюти.

Соціальні мережі стали одним з найважливіших джерел інформації та впливу на сучасне суспільство. Інформація, що поширюється через соціальні мережі, може швидко поширюватись та впливати на настрої та рішення учасників ринку. Таким чином, розуміння впливу соціальних мереж на курс криптовалют стає надзвичайно важливим для трейдерів, інвесторів та зацікавлених осіб, які прагнуть приймати обґрунтовані рішення та розробляти більш ефективні стратегії торгівлі та управління ризиками.

У кваліфікаційній роботі було досліджено вплив соціальних мереж, зокрема Twitter, на курс Bitcoin, одну з найбільш відомих криптовалют. У ході роботи було розглянуто моделі для моделювання часових рядів, а саме ARIMAX та пакет Prophet.

Загальною характеристикою обох моделей є їхня здатність прогнозувати майбутній курс Bitcoin на основі історичних даних та інформації з Twitter, а також вказувати вплив кожної екзогенної змінної. Вони дозволяють трейдерам та інвесторам отримувати важливі прогнози, які можуть бути використані для прийняття усвідомлених торговельних рішень.

Проте, варто враховувати, що прогнозування цінової динаміки криптовалюти є складним завданням через велику волатильність та непередбачуваність ринку. Жодна модель не може забезпечити абсолютну точність у прогнозуванні, оскільки це залежить від багатьох невизначених факторів, таких як геополітичні події, регуляторні рішення або технічні проблеми.

Також важливо враховувати, що соціальні мережі, зокрема Twitter, можуть бути суб'єктивними джерелами інформації, де поширюються як об'єктивні аналізи, так і особисті думки та спекуляції. Врахування цих особливостей може бути викликом при використанні цих даних у моделях прогнозування.

Незважаючи на ці обмеження, ARIMAX і Prophet моделі є цінними інструментами у прогнозуванні курсу Bitcoin з використанням даних з Twitter. Вони надають підтримку трейдерам та інвесторам у процесі прийняття рішень та допомагають зрозуміти вплив соціальних мереж на криптовалютний ринок.

Продовжуючи дослідження впливу соціальних мереж, зокрема Twitter, на курс Bitcoin, можливо розширити область дослідження і використовувати інші криптовалюти або фінансові активи. Це дозволить отримати більш широке розуміння взаємозв'язку між соціальними медіа та фінансовими ринками загалом.

Також варто розглянути можливість використання інших моделей та алгоритмів, які можуть бути більш адаптивними до змін в ринкових умовах та соціальній динаміці.

СПИСОК ДЖЕРЕЛ ТА ЛІТЕРАТУРИ

1. Пасик С. Війна в Україні та її вплив на курс криптовалют: аналіз та прогнозування. *ШЕВЧЕНКІВСЬКА ВЕСНА 2023. Повоєнне відновлення економіки України: проблеми та перспективи* : Матеріали XXI Міжнар. науково-практ. конф. студентів, аспірантів та молодих вчен., м. Київ. 2023. С. 70. URL: <https://econom.knu.ua/wp-content/uploads/2023/04/Збірник-Шевченківська-весна-2023.docx.pdf>.
2. Савченко М., Завидовська А., Краєвська В. Кореляційний аналіз волатильності криптовалют світу. *Галицький економічний вісник*. 2022. Т. 75. С. 122–133. URL: https://doi.org/10.33108/galicianvisnyk_tntu2022.02.122.
3. Ставицький А. Теорія часових рядів. *andriystav*. URL: http://andriystav.cc.ua/Downloads/AppliedEco/02_Time_Series.pdf.
4. 10 Elon Musk Tweets That Created Waves In Crypto World. *Outlookindia*. URL: <https://www.outlookindia.com/business/10-elon-musk-tweets-that-created-waves-in-crypto-world-news-233190>.
5. After the Crash: Cryptocurrency Predictions for 2022 and Beyond. *Binance Blog*. URL: <https://www.binance.com/en/blog/markets/after-the-crash-cryptocurrency-predictions-for-2022-and-beyond-421499824684903402>.
6. Alonso-Cortés M., Arribas V. Time Series Analysis and Forecasting with FB Prophet Python. *Model Differently*. URL: https://www.modeldifferently.com/en/2022/04/analysis_prediccion_ts_prophet/.
7. Archie A. Dogecoin price spikes after Elon Musk changes Twitter logo to the Shiba Inu dog. *NPR*. URL: <https://www.npr.org/2023/04/04/1167877216/dogecoin-elon-musk-twitter-logo>.
8. Bitcoin 17,7 million Tweets and price. *Kaggle: Your Machine Learning and Data Science Community*. URL: <https://www.kaggle.com/datasets/jaimebadiola/bitcoin-tweets-and-price?resource=download>.
9. Bitcoin Pizza Day. *National Today*. URL: <https://nationaltoday.com/bitcoin-pizza-day/>.

10. Bitcoin Price Forecasting Using Web Search And Social Media Data / R. Jain et al. 2018. URL: <https://support.sas.com/resources/papers/proceedings18/3601-2018.pdf>.
11. Bollen J., Mao H., Zeng X. Twitter mood predicts the stock market. *Journal of Computational Science*. 2011. Vol. 2, no. 1. P. 1–8. URL: <https://doi.org/10.1016/j.jocs.2010.12.007>.
12. Boughton N. Eight key factors that affect cryptocurrency value. *Canadian Investment News. Wealth Professional*. URL: <https://www.wealthprofessional.ca/news/industry-news/eight-key-factors-that-affect-cryptocurrency-value/367834>.
13. Box G. E. P. Time series analysis: Forecasting and control. San Francisco : Holden-Day, 1970. 553 p. URL: <https://archive.org/details/timeseriesanalys0000boxg/page/n6/mode/1up>.
14. Box G. E. P., Tiao G. C. Intervention Analysis with Applications to Economic and Environmental Problems. *Journal of the American Statistical Association*. 1975. Vol. 70, no. 349. P. 70-79. URL: <https://doi.org/10.1080/01621459.1975.10480264>.
15. Brownlee J. How to Create an ARIMA Model for Time Series Forecasting in Python. *Machine Learning Mastery*. URL: <https://machinelearningmastery.com/arima-for-time-series-forecasting-with-python/>.
16. Crypto Market Cap. *Investing.com*. URL: <https://www.investing.com/crypto/charts>.
17. Cryptocurrencies: A Guide to Getting Started. World economic forum, 2021. URL: https://www3.weforum.org/docs/WEF_Getting_Started_Cryptocurrency_2021.pdf.
18. Cryptocurrency Prices, Charts And Market Capitalizations. *CoinMarketCap*. URL: <https://coinmarketcap.com/>.
19. Dai W. B-money. *Wei Dai's Home Page*. URL: <http://www.weidai.com/bmoney.txt>.
20. Effect of Twitter investor engagement on cryptocurrencies during the COVID-19 pandemic / A. Bouteska et al. *Research in International Business and Finance*. 2023. Vol. 64. URL: <https://doi.org/10.1016/j.ribaf.2022.101850>.
21. Ethereum price: ETH to USD chart. *Ledger*. URL: <https://www.ledger.com/coin/price/ethereum>.

22. Frankenfield J. Initial Coin Offering (ICO): Coin Launch Defined, with Examples. *Investopedia*. URL: <https://www.investopedia.com/terms/i/initial-coin-offering-ico.asp>.
23. H.R.6154 – Crypto-Currency Act of 2020: bill. 2020. 7 p. URL: <https://www.congress.gov/bill/116th-congress/house-bill/6154/text?r=1&am;s=1>.
24. Howarth J. How Many Cryptocurrencies are There In 2023? *Exploding Topics*. URL: <https://explodingtopics.com/blog/number-of-cryptocurrencies>.
25. Hyndman R. J., Athanasopoulos G. *Forecasting: Principles and Practice*. 2nd ed. Monash University, 2018. 382 p. URL: <https://otexts.com/fpp2/>.
26. Koblitz N., Menezes A. J. Cryptocash, cryptocurrencies, and cryptocontracts. *Designs, Codes and Cryptography*. 2015. Vol. 78, no. 1. P. 87-102. URL: <https://doi.org/10.1007/s10623-015-0148-5>.
27. Kumar V. G. Statistical Tests to Check Stationarity in Time Series. *Analytics Vidhya*. URL: <https://www.analyticsvidhya.com/blog/2021/06/statistical-tests-to-check-stationarity-in-time-series-part-1/>.
28. Larsen K. GAM: The Predictive Modeling Silver Bullet. *Stitch Fix Technology*. URL: <https://multithreaded.stitchfix.com/blog/2015/07/30/gam/>.
29. Maloney T., Yang Y., Bartenstein B. World's Biggest Crypto Fortune Began With a Friendly Poker Game. *Bloomberg*. URL: <https://www.bloomberg.com/news/features/2022-01-09/binance-ceo-cz-s-net-worth-billionaire-holds-world-s-biggest-crypto-fortune>.
30. Mehandzhiyski V. What Is an ARIMAX Model?. 365 *Data Science*. URL: <https://365datascience.com/tutorials/python-tutorials/arimax/>.
31. Mt.Gox: How many Bitcoins were stolen. *The Cryptonomist*. URL: <https://en.cryptonomist.ch/2022/08/15/mt-gox-how-bitcoin-stolen/>.
32. Nakamoto S. Bitcoin: A Peer-to-Peer Electronic Cash System. 2008. 9 p. URL: <https://bitcoin.org/bitcoin.pdf>.
33. Nick Szabo. *CoinMarketCap*. URL: <https://coinmarketcap.com/alexandria/people/nick-szabo>.
34. Oxford Learner's Dictionaries. *Oxford Learner's Dictionaries*. URL: <https://www.oxfordlearnersdictionaries.com>.

35. Phillips G. How Does a Twitter Crypto Scam Work?. *MAKEUSEOF*. URL: <https://www.makeuseof.com/how-does-twitter-crypto-scam-work/>.
36. Rao T., Srivastava S. Analyzing stock market movements using Twitter sentiment analysis. *ASONAM 2012*. 2012. URL: <https://dl.acm.org/doi/pdf/10.5555/2456719.2456923>.
37. Shafi A. What is a Generalized Additive Model? *towardsdatascience.com*. URL: <https://towardsdatascience.com/generalised-additive-models-6dfbedf1350a#fb90>.
38. Taylor S. J., Letham B. Forecasting at Scale. *The American Statistician*. 2018. Vol. 72, no. 1. P. 37-45. URL: <https://doi.org/10.1080/00031305.2017.1380080>.
39. The idea and a brief history of cryptocurrencies. *The Guardian*. URL: <https://guardian.ng/technology/tech/the-idea-and-a-brief-history-of-cryptocurrencies/>.
40. Time Series Analysis: The Basics. *Australian Bureau of Statistics*. URL: <https://www.abs.gov.au/websitedbs/d3310114.nsf/home/time+series+analysis:+the+basics>.
41. What is cryptocurrency and how does it work?. *www.kaspersky.com*. URL: <https://www.kaspersky.com/resource-center/definitions/what-is-cryptocurrency>.
42. Yahoo Finance – Stock Market Live, Quotes, Business & Finance News. *Yahoo Finance*. URL: <https://finance.yahoo.com/>.
43. Yustindra A. F., Claster W. B. Forecasting cryptocurrency price using Twitter sentiment variable. *International Journal of Computer Techniques*. URL: <http://www.ijctjournal.org/Volume6/Issue1/IJCT-V6I1P9.pdf>.

ДОДАТОК А

Програмний код

```
library(readxl)
library(zoo)
library(lubridate)
library(forecast)
library(lmtest)
library(dplyr)
library(tseries)
library(prophet)
library(«Metrics»)

tw_data = read_excel(«1 database.xlsx»)

tw_data = tw_data %>%
  mutate(across(c(Count_Negatives, Count_Positives, Count_Neutrals),
    ~if_else(is.na(Compound_Score), replace(., . == 0, NA), .)),
    n = if_else(is.na(Compound_Score), replace(n, n %in% c(60, 120), NA), n)) %>%
  na.locf()

View(tw_data)

data = tw_data %>% select(n, Count_Negatives, Count_Positives, Count_Neutrals, Close,
Date)
colnames(data) = c(«tweets», «neg_tweets», «pos_tweets», «neut_tweets», «price», «date»)
data$date = as.POSIXct(data$date, tz = «UTC»)

btc_ts = zoo(data$price, data$date)
plot(btc_ts, type = 'l', ylab = «USD», xlab = «Роки», main = «Курс BTC/USD»)
```

```

tw_data$date_day = as.Date(tw_data$Date)
daily1 = aggregate(n ~ date_day, data = tw_data, FUN = sum)
daily2 = aggregate(Count_Positives ~ date_day, data = tw_data, FUN = sum)
daily3 = aggregate(Count_Negatives ~ date_day, data = tw_data, FUN = sum)
daily4 = aggregate(Count_Neutrals ~ date_day, data = tw_data, FUN = sum)

plot(daily1, main = «Загальна кількість твітів», ylab = «», xlab = «Роки», type = 'l')
plot(daily3, main = «Негативні твіти», ylab = «», xlab = «Роки», type = 'l')
plot(daily2, main = «Позитивні твіти», ylab = «», xlab = «Роки», type = 'l')
plot(daily4, main = «Нейтральні твіти», ylab = «», xlab = «Роки», type = 'l')

acf(data$price, main = «ACF-тест для Bitcoin»)
pacf(data$price, main = «PACF-тест для Bitcoin»)
adf.test(data$price)

train_data = data[1:round(0.8*nrow(data)), ]
test_data = data[(round(0.8*nrow(data))+1):nrow(data), ]

regressors_train = data.frame(
  tweets = train_data$tweets,
  negatives = train_data$neg_tweets,
  positives = train_data$pos_tweets,
  neutrals = train_data$neut_tweets
)

regressors_test = data.frame(
  tweets = test_data$tweets,
  negatives = test_data$neg_tweets,
  positives = test_data$pos_tweets,
  neutrals = test_data$neut_tweets
)

```

```

arima = auto.arima(train_data$price, xreg = as.matrix(regressors_train))
summary(arima)

fcst = forecast(arima, h = 10, xreg = as.matrix(regressors_test))
plot(fcst, main = «Прогноз Bitcoin за допомогою ARIMAX(1,1,2)», xaxt = «n»)
axis(side = 1, at = c(3673, 12433), labels = c(2018,2019))
lines(data$price)

rsq_btc = cor(fcst$x, fcst$fitted)^2
rsq_btc
Box.test(fcst$residuals, type = «Ljung»)

#####

btc_df = data.frame(tw_data$Date, tw_data$Close, tw_data$Count_Negatives,
tw_data$Count_Positives, tw_data$Count_Neutrals,
pos_tweets = tw_data$Count_Positives, neut_tweets =
tw_data$Count_Neutrals)
colnames(btc_df) = c(«ds», «y», «n», «negative_tweets», «positive_tweets»,
«neutral_tweets»)
btc_df$ds = as.POSIXct(btc_df$ds, tz = «UTC»)
View(btc_df)

train_df = btc_df[1:(12936-30), ]

model_pr = prophet()
model_pr = add_regressor(model_pr, «n»)
model_pr = add_regressor(model_pr, «negative_tweets»)
model_pr = add_regressor(model_pr, «positive_tweets»)
model_pr = add_regressor(model_pr, «neutral_tweets»)

```

```
model_pr = fit.prophet(model_pr, train_df)

future = make_future_dataframe(model_pr, periods = 30)
future$n = btc_df$n
future$negative_tweets = btc_df$negative_tweets
future$neutral_tweets = btc_df$neutral_tweets
future$positive_tweets = btc_df$positive_tweets

forecast = predict(model_pr, future)

plot(model_pr, forecast, main = «Прогноз Bitcoin за допомогою prophet», ylab = «USD»,
xlab = «Роки»)
dyplot.prophet(model_pr, forecast)

prophet_plot_components(model_pr, forecast)

regressor_coefficients(model_pr)

rmse(btc_df$y, forecast$yhat)
mae(btc_df$y, forecast$yhat)
mape(btc_df$y, forecast$yhat)
mase(btc_df$y, forecast$yhat)

rsq_btc_pr = cor(btc_df$y, forecast$yhat)^2
rsq_btc_pr
Box.test(btc_df$y-forecast$yhat)
```