

**МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
КИЇВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ
ІМЕНІ ТАРАСА ШЕВЧЕНКА**

Економічний факультет

Кафедра статистики, інформаційно-аналітичних систем і демографії

КВАЛІФІКАЦІЙНА РОБОТА МАГІСТРА

**ПОРІВНЯЛЬНИЙ АНАЛІЗ МЕТОДІВ ПРОГНОЗУВАННЯ КУРСУ
КРИПТОВАЛЮТ**

Студента II курсу
спеціальності 051 Економіка
за освітньо-науковою
програмою
«Економічна аналітика та
статистика»
денної форми навчання
Юрія СТАЩУКА

(підпис)

Науковий керівник:
канд. екон. наук, доцент
Тетяна ТРУБНІК

(підпис)

Роботу допущено до захисту на засіданні ЕК рішенням кафедри
статистики, інформаційно-аналітичних систем і демографії, протокол № 12 від
«12» травня 2022 року

Завідувач кафедри

д-р екон. наук Наталія КОВТУН

РЕФЕРАТ

Кваліфікаційна робота магістра містить 50 с., 15 рис., 18 табл., 30 джерел, 1 додаток.

Ключові слова: криптовалюта, ринок криптовалют, прогнозування часових рядів, курс криптовалюти, волатильність, моделі часових рядів.

Предмет дослідження: теоретико-методологічні та практичні засади статистичного прогнозування ціни криптовалют.

Об'єкт дослідження: ринок криптовалют.

Мета дослідження: побудова прогнозів курсу криптовалют, порівняння різних моделей та оцінювання їх якості.

Методи дослідження: кореляційний та регресійний аналіз, моделі часових рядів, крос перевірка часових рядів.

Наукова новизна, практична значимість дослідження: обґрунтування методів добору та побудова моделей для прогнозування часових рядів ціни криптовалют, порівняння методів прогнозу ціни на ринку криптовалют.

Практична цінність: оцінювання якості та точності прогнозу ціни криптовалюти та співставлення результатів з реальними значеннями, розробка рекомендації щодо використання апробованих алгоритмів для моделювання високоволатильних часових рядів.

Основні положення роботи апробовані на студентській конференції «Шевченківська весна» та опубліковані у збірнику «Шевченківська весна 2022. Цифрова трансформація економіки в умовах пандемії COVID-19».

RESUME

Taras Shevchenko National University of Kyiv
Faculty of Economics, Department of Statistics, Information and Analytical
Systems and Demography

Key words: cryptocurrency, cryptocurrency market, time series forecasting, cryptocurrency exchange rate, volatility, time series models.

The purpose of graduation research of 2-year master student Yurii Stashchuk's "Comparative analysis of cryptocurrency exchange rate forecasting methods" is build cryptocurrency exchange rate forecasts, compare different models and evaluate their quality.

The cryptocurrency market is a fairly new phenomenon, but contains quite similar to the classic stock market. Cryptocurrency market has gained momentum recently, but has become quite popular, which has aroused the interest of scientists in this matter.

In this research were build 8 different models such as ARIMA, SARIMA, ARIMAX, Holt-Winter's model, ETS model, BATS, THETA and neural network models for cryptocurrency exchange rate time series. Among fitted models using time series cross validation was chosen the best model comparing RMSE and MAE.

Research contains: 50 pages, 15 fig., 18 tables, 30 bibliog., 1 append.

ЗМІСТ

ВСТУП	3
РОЗДІЛ I. ТЕОРЕТИКО-МЕТОДОЛОГІЧНІ ЗАСАДИ СТАТИСТИЧНОГО ВИВЧЕННЯ РИНКУ КРИПТОВАЛЮТ	6
1.1. Поняття ринку криптовалют, особливості та перспективи його функціонування.....	6
1.2. Волатильність ринку криптовалют як об'єкт статистичного дослідження	11
РОЗДІЛ II. МЕТОДОЛОГІЧНІ ЗАСАДИ МОДЕЛЮВАННЯ ДИНАМІКИ	14
2.1. Особливості моделей прогнозування часових рядів.	14
2.2. Специфіка перевірки адекватності моделей динаміки реальним процесам.	22
2.3. Інформаційне забезпечення моделювання курсу Біткоїна, Ефіріума та Сенд.	25
РОЗДІЛ III. СТАТИСТИЧНЕ МОДЕЛЮВАННЯ ТА ПРОГНОЗУВАННЯ КУРСУ КРИПТОВАЛЮТ	28
3.1. Порівняльний аналіз моделей прогнозування цін на ринку криптовалют. ..	28
3.2. Оцінювання точності моделей за допомогою крос перевірки часових рядів.	35
ВИСНОВКИ	40
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ	42
ДОДАТКИ	46

ВСТУП

За останні 20 років спостерігається стрімкий розвиток технологій майже у всіх сферах економіки. Найголовнішим поштовхом змін до стрімкого розвитку цифрових технологій стала поява інтернету. З'явилась можливість обміну інформацією між різними користувачами на великих відстанях та як наслідок стало можливим придбання товарів та послуг, здійснення банківських переказів, спілкування на відстані, робота та навчання з дому. Вказана перевага приваблює більшість користувачів мережі, споживачів та продавців, працівників та роботодавців. Державні установи також поступово переходять у режим онлайн для спрощення доступу до державних послуг та зменшення часу обслуговування громадян. Особливо це стало актуальним під час епідемії коронавірусу для зменшення особистих контактів між людьми.

На особливу увагу заслуговують нові технології у сфері фінансів, завдяки яким пришвидшується розвиток економіки та забезпечується зростання благополуччя середнього класу. Звичайним та загальноприйнятим стає застосування технологій у сфері фінансів через онлайн платежі з використанням «банку у смартфоні», переказ коштів у будь-яку точку світу, розрахунок безконтактним способом та онлайн-покупки, що є надзвичайно зручним та приваблює багатьох людей. Проте, не завжди є можливість здійснити ті ж платежі за кордон, особливо це стосується країн з низьким рівнем довіри до банківської системи та законодавством країни в якій функціонує фінансова організація. Також, оскільки перекази регулюють банки, вони отримують значну комісію за свою роботу та можуть взагалі відмовити здійснювати будь-які операції без пояснень.

На зміну цьому приходять криптовалюти та технології які за ними стоять. Криптовалюти на даний момент займають особливе місце, адже це новий інструмент розрахунку, який забезпечує швидкість, прозорість та низьку комісію при транзакціях, а також є децентралізованою системою, що живе за рахунок власних користувачів. Багато високорозвинених країн замислюються над тим, щоб визнати криптовалюти фінансовим інструментом та застосувати технології

криптографії для розробки державних криптовалют. Проте, основною перешкодою для масового використання криптовалют є саме невизначеність у правовому полі, відносна складність використання та висока волатильність ціни криптовалют.

Об'єктом дослідження є ринок криптовалют.

Предметом дослідження є теоретико-методологічні та практичні засади статистичного прогнозування ціни криптовалют.

Метою роботи є побудова прогнозів курсу криптовалют, порівняння різних моделей, оцінювання їх якості.

Для досягнення поставленої мети, визначені основні завдання дослідження:

- визначити теоретичні особливості побудови моделей часових рядів;
- обґрунтувати методологію дослідження ринку криптовалют;
- охарактеризувати інформаційне забезпечення щодо історичних даних про курс криптовалют;
- дослідити особливості часових рядів та провести їх графічний аналіз;
- побудувати моделі часових рядів цін Біткоїна, Ефіріума та Сенд;
- виконати порівняльний аналіз моделей прогнозування;
- оцінити адекватність моделей реальним процесам та точність прогнозів ціни криптовалют;
- розробити рекомендації щодо використання напрацьованих (апробованих) алгоритмів для моделювання високоволатильних часових рядів.

Інформаційною базою проведеного дослідження є історичні дані щодо курсу криптовалют, які були отриманні за допомогою API [22].

Дана робота базується на загальнонаукових і фундаментальних положеннях статистичної науки та економіко-математичного моделювання. У дослідженні застосовувалися такі статистичні методи, як ARIMA, SARIMA, ARIMAX, модель Хотта-Вінтерса, модель експоненціального згладжування, BATS, THETA та модель нейронних мереж. Вибір кращої моделі здійснювався за допомогою процедури крос перевірки часових рядів. Практична частина виконувалась у середовищі програмування R.

Магістерська робота складається з трьох розділів. У першому розглядаються теоретичні засади статистичного оцінювання ринку криптовалют. У другому розділі розкривається методика побудови та оцінювання якості моделей прогнозування курсу криптовалют, здійснена підготовка первинних даних, у третьому подано результати побудови прогнозу ціни криптовалюти різними методами та визначені найкращі з них.

РОЗДІЛ I. ТЕОРЕТИКО-МЕТОДОЛОГІЧНІ ЗАСАДИ СТАТИСТИЧНОГО ВИВЧЕННЯ РИНКУ КРИПТОВАЛЮТ

1.1. Поняття ринку криптовалют, особливості та перспективи його функціонування.

Важливою складовою успішного функціонування економіки будь-якої країни та й світу в цілому є розвинутий та прозорий фінансовий сектор. Основне його призначення – це акумулювання вільних грошових коштів та направлення їх у різні галузі економіки для їх постійного розвитку та безперервного функціонування. За останні 20 років фінансовий сектор зазнає значних змін, особливо з появою інтернет технологій. Зокрема це стосується доступності фінансових інструментів різної складності починаючи від онлайн платежів, P2P кредитування, швидкого онлайн кредитування, закінчуючи доступом до світових фінансових ринків. Саме розвиток фінансової інфраструктури, доступність та прозорість надають можливість акумулювати значно більші грошові потоки на фінансових ринках, оскільки значна частина населення не просто заощаджує кошти, а інвестує їх у різні галузі економіки.

Основною ознакою функціонування сучасних фінансових ринків є довіра між його учасниками, або ж по-іншому дотримання взятих на себе зобов'язань. Основним гарантом прав та обов'язків учасників фінансового ринку є закони держави, на території якої проводиться фінансова діяльність та міжнародні закони. Проте закони не завжди працюють досконало, особливо якщо це стосується «офшорних» юрисдикцій. На заміну централізованим фінансовим системам варто розглядати децентралізовані, які працюють на потужностях своїх користувачів та не мають центрального регулятора. Одною з таких децентралізованих систем є технологія блокчейн.

Блокчейн дозволяє суттєво змінити управління системою через розподіл між людиною та закладеними алгоритмами процес прийняття рішень. З'являється новий механізм контролю за виконанням рішень – це технічне рішення для поняття «довіра». Іншими словами, технологія усуває вплив

людського фактора з механізму виконання договорів на будь-якому рівні [1, с. 23].

На думку фахівців блокчейн – це багатофункціональна та багаторівнева інформаційна технологія, призначена для надійного обліку різних активів, технологія надійного розподіленого зберігання записів про всі коли-небудь зроблені транзакції [17, с. 1].

На основі технологію блокчейн було розроблено криптовалюти. Проте у світі не існує єдиного підходу до визначення криптовалют та їх правового та фінансового регулювання.

Криптовалюта – це різновид цифрової валюти, облік внутрішніх розрахункових одиниць якої забезпечує децентралізована платіжна система, що працює в повністю автоматичному режимі. Криптовалюта не має матеріальної форми – це певна кількість одиниць валюти, яка розміщена на електронному гаманці [12].

Слід зазначити основна ціль криптовалют – це створення альтернативної децентралізованої фінансової системи, яка може виправити недоліки наявної фінансової системи.

Звичайно, криптовалюти як і будь-яка функціонуюча система має свої переваги та недоліки.

До основних переваг використання криптовалют відносять:

1. Низька комісія при переказах (мінімальна вартість при транзакції в мережі Біткоїн становить 0,00001 Біткоїна – на даний момент близько 0,4\$, а мінімальна комісія у мережі XLM становить 1 XLM, а це близько 0,02\$ незалежно від розміру переказу);

2. Висока швидкість транзакції. Обробка однієї транзакції займає до 10 хвилин та не залежить від того, в яку точку світу направлена ця транзакція;

3. Відносна анонімність та прозорість. Усі користувачі мають змогу бачити абсолютно усі перекази між електронними гаманцями, але визначити кому належить гаманець є досить складним завданням;

4. Для проведення транзакції двом учасникам необхідно мати лише доступ до мережі інтернет та номер електронного гаманця, на який буде здійснено переказ;

5. Можливість будь-якої особи брати участь у емісії криптовалюти та отримувати винагороду від мережі.

До недоліків криптовалют варто віднести наступні:

1. Висока волатильність курсу. Всі криптовалюти не мають центрального регулятора та не підкріплені реальними ресурсами, що робить їх курс залежним лише від попиту та пропозиції, а також від маніпуляцій з боку учасників, які оперують великими об'ємами криптовалюти;

2. Якщо банківські установи повністю відповідають за зберігання та перекази коштів користувачів, то використовуючи криптовалюти, користувачі змушені самостійно забезпечити зберігання криптовалют та несуть повну відповідальність за свої перекази, адже повернути втрачені кошти при помилковому переказі неможливо;

3. Використання криптовалют у злочинних цілях для купівлі незаконних предметів та речовин, а також для легалізації коштів, отриманих незаконним шляхом;

4. Відсутність законодавчого регулювання та незрозумілість у визначенні криптовалюти як активу, що дає можливість інвесторам ухилятися від сплати податків.

Водночас, враховуючи всі переваги та недоліки криптовалют та відносну новизну, ринок криптовалют має досить великі перспективи. Перш за все, криптовалюти мають свій повноцінний ринок, який функціонує за усіма ринковими законами. Тобто особа може брати участь в емісії криптовалют, має право набувати криптовалюту через купівлю/обмін, зберігати та використовувати криптовалюти.

Незважаючи на високий ризик, при інвестуванні у ринок криптовалют, кількість учасників на ринку постійно зростає. Так, згідно з дослідженням Кембриджського університету, кількість користувачів криптовалют становить

101 млн осіб станом на 3 квартал 2020 року. При цьому у 2018 році ця цифра становила 35 млн осіб, що вказує на збільшення користувачів на 189%. [18, с. 44]. Зростання користувачів пов'язано з значною медійністю навколо криптовалют, запуску нових проектів, що обіцяють учасникам заробити швидкі та значні статки, розмови на високому рівні про легалізацію криптовалют та впровадження криптовалют великими компаніями (наприклад, PayPal). Навіть інституційні інвестори почали розглядати можливість інвестицій у Біткоїни. Так, у жовтні 2021 року з'явився перший біткоїн-ETF, який був схвалений Комісією з цінних паперів і бірж США та почав торгуватись на фондовому ринку [15]. Також варто згадати й про компанію Grayscale, яка почала активно купувати різні криптовалюти та створювати трасти. [23]

Перші, хто найбільше зазнають впливу від популяризації криптовалют є банки. Багато центральних та приватних банків слідкують за розвитком ринку криптовалют та намагаються адаптуватись до нього. Зокрема, банки активно розробляють власні криптовалюти, або використовують уже наявні. Наприклад, криптовалюту Ripple (третя за капіталізацією криптовалюта) використовують компанії «UniCredit», «UBS», «Santander». [26]

Також варто згадати й про Національний Банк України, який протягом 2018-2019 років реалізовував пілотний проект е-гривні, а перші кроки для здійснення проекту були закладені ще у 2016 році. Е-гривня розроблена на основі технології блокчейн, що дає можливість застосувати більшість переваг криптовалют для національної електронної валюти. Основна ж мета е-гривні – це зменшення вартості транзакцій, прозорість та законність усіх переказів. [7]

Окрім банків криптовалюту починають використовувати й приватні компанії. Зазвичай це компанії, діяльність яких пов'язана з високими технологіями або продажами через мережу інтернет. До них можна віднести Steam, Tesla, Wordpress, Virgin, Jeep [6, с. 4].

Підтвердженням цієї думки є постійне зростання самого ринку криптовалют. Капіталізація ринку станом на лютий 2022 року складає 1,8 трлн дол. США, а на початок 2021 року ця сума складала майже 1 трлн дол. США. Пік

капіталізації ринку криптовалют припав на кінець 2021 року та склав 3 трлн дол. США. [9]

Слід відмити той факт, що основна перспектива саме криптовалют як домінуючого засобу платежу – це використання «стейблкоїнів», або якщо говорити по-іншому криптовалют, які ціна токена якої становить 1 долар США. Вони мають усі переваги криптовалют, але не є такими волатильними, тому зручні для використання як засіб платежу та заощаджень.

Разом з тим, потрібно розкрити юридичний статус криптовалют в Україні. До недавнього часу криптовалюти в Україні не мали чіткого статусу, що відповідно ускладнювало процес обліку та оподаткування криптовалютних активів. Єдиним документом, розміщеним на сайті НБУ, який визначав правовий статус криптовалют – Спільна заява фінансових регуляторів щодо статусу криптовалют в Україні датований кінцем листопада 2017 року. У ній регулятори не визнавали криптовалюти як будь-який засіб платежу чи інший грошовий сурогат. [14]. Але треба зазначити, що наприкінці лютого 2022 р. Верховною Радою України було прийнято закон «Про віртуальні активи» та підписаний Президентом України. Відповідно до закону:

- встановлено правовий статус, класифікацію та права власності віртуальних активів;
- створено умови для подальшого формування правового поля на ринку віртуальних активів;
- визначено перелік постачальників послуг віртуальних активів та умови їх реєстрації;
- передбачено впровадження заходів фінансового моніторингу у сфері віртуальних активів.

Проте закон вступить у силу лише після змін у Податковому кодексі, які будуть передбачати особливості оподаткування операцій з криптовалютними активами, а регулятором цього процесу є Національна комісія з цінних паперів та фондового ринку та НБУ. [2]

1.2. Волатильність ринку криптовалют як об'єкт статистичного дослідження.

Волатильність – це показник, що показує ступінь коливання часових рядів, у фінансовій статистиці – це показник, що характеризує динаміку зміни ринкових цін та доходності впродовж певного часу. [3]

Іноді волатильність називають «розкидом» цін, курсів та інших кількісних показників і зазвичай вимірюється дисперсією або представленим у річному виразі стандартним відхиленням змін. [13]

Ринок криптовалют має досить значні відмінності від традиційних ринків, та характеризується високою волатильністю. Формуванням курсу криптовалют відбувається виключно під впливом попиту та пропозиції, що і спричиняє високу волатильність ціни криптовалюти. Тому при інвестуванні у ринок криптовалют варто розуміти сутність волатильності та відповідно оцінювати ризики, які з нею пов'язані. Зазвичай інвестори, які діють на ринку криптовалют є схильними до високих ризиків, що дає змогу заробити їм значні статки протягом короткого терміну, або ж так само швидко їх втратити. З цього можна дійти висновку до важливості постійного аналізу ринку.

До основних причин волатильності криптовалют відносять:

1. Відсутність регулювання з боку центральних органів влади. Ринок криптовалют є децентралізованим, тому діяльність криптовалютних бірж знаходиться у невизначеному правовому полі; учасники ринку залишаються незахищеними, оскільки можуть повністю та безповоротно втратити свій капітал, відповідно вони хочуть отримати бонус за ризик;

2. Спекуляція. Оскільки курс залежить виключно від попиту та пропозиції, тому учасники ринку з великим обсягом криптовалюти можуть безпосередньо впливати на ціну [4];

3. Відсутність забезпечення реальними активами. Криптовалюти функціонують лише за умови, якщо їх використовують, але як тільки великі учасники ринку виведуть свої статки у фіатні валюти, ринок криптовалют може рухнути;

4. Новини. Інформаційні приводи значно впливають на поведінку учасників ринку [5].

Для оцінки волатильності ринку криптовалют скористаємось даними денної відсоткової зміни цін найбільш популярних за капіталізацією криптовалют (Bitcoin, Ethereum), а також для порівняння акції компанії Apple. Дані взяті у період з 23.04.2021 – 23.04.2022 [11]. Графічно динаміка представлена на рисунку 1.1.

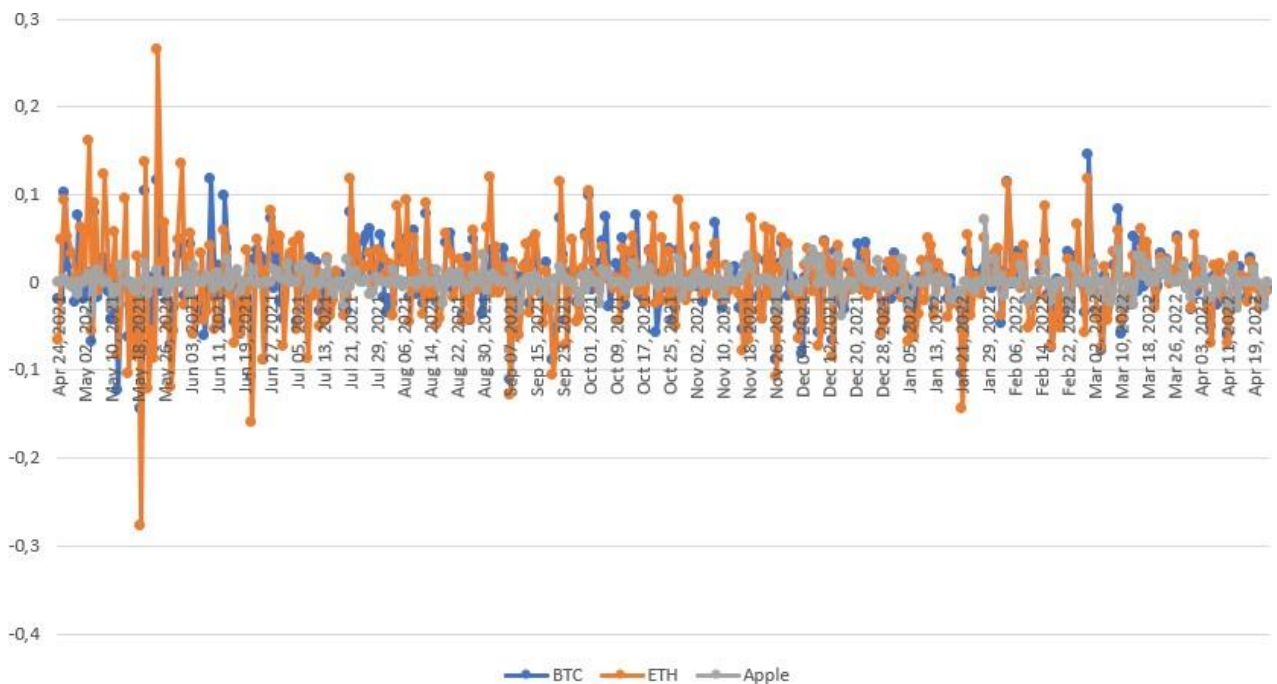


Рис. 1.1. – Динаміка волатильності Bitcoin, Ethereum, Apple

Джерело: Побудовано автором на основі даних [11]

Як видно з рис. 1.1., волатильність криптовалют має схожу тенденцію, що власне показує загальну ситуацію на ринку (кореляція волатильності становить 0,82). Зазвичай інші, менші за капіталізацією криптовалюти, реагують на зміни так само, і лише в окремих випадках ситуація є іншою. При цьому волатильність акцій Apple є набагато меншою за волатильність криптовалют. Це можна проілюструвати за допомогою таблиці 1.1.

Табл. 1.1. – Розкид максимальних та мінімальних значень волатильності

	BTC	ETH	Apple
Максимум	0,1449	0,2646	0,0698
Мінімум	-0,1439	-0,2786	-0,0393

Джерело: Побудовано автором на основі даних [11]

Для оцінювання доцільності інвестування у криптовалюту застосовують показник доходності від інвестування, який розраховується за формулою:

$$r_i = \frac{(I_i - I_0)}{I_0} * 100, \quad (1.1)$$

де I_i – значення відповідного активу на початок періоду інвестування;

I_0 – значення відповідного активу на кінець періоду інвестування.

Ризик у даному випадку розраховується як стандартне відхилення. [10]

Табл. 1.2. – Дохідність-ризик для Біткоїна та Apple

	Дохідність		Ризик	
	BTC	Apple	BTC	Apple
1 міс	32,68%	-0,03%	29,49%	3,44%
2 міс	102,63%	6,32%	31,50%	3,32%
3 міс	123,79%	0,25%	26,80%	5,45%
6 міс	303,72%	3,51%	28,56%	6,12%
9 міс	1175,77%	17,79%	26,18%	5,73%
1 рік	540,22%	16,76%	32,23%	5,66%
1 рік 6 міс	511,82%	57,13%	29,28%	6,81%
2 роки	285,28%	32,21%	28,22%	8,04%
2 роки 6 міс	670,19%	55,87%	27,58%	8,13%
3 роки	494,22%	76,98%	26,66%	8,23%
3 роки 6 міс	897,87%	222,41%	25,32%	9,11%
4 роки	5335,52%	240,06%	25,37%	8,93%
4 роки 6 міс	3951,96%	293,93%	25,01%	8,68%
5 рік	4109,58%	386,11%	24,65%	8,42%

Джерело: Побудовано автором на основі даних [11]

За початковий період інвестування було взято березень 2017 року. За даними табл. 1.2. можна побачити, якщо інвестувати у 2017 році у Біткоїн, то у березні 2022 року отримаємо дохідність 4109%, а за акції Apple інвестовані за той самий період ми отримаємо дохідність всього 386%.

РОЗДІЛ II. МЕТОДОЛОГІЧНІ ЗАСАДИ МОДЕЛЮВАННЯ ДИНАМІКИ

2.1. Особливості моделей прогнозування часових рядів.

Розглянемо теоретичні основи моделей часових рядів, а саме моделі експоненціального згладжування, модель Холта-Вінтерса, ARIMA, BATS, ТНІТА, моделі нейронних мереж.

Експоненціальне згладжування – правило техніки згладжування даних часових рядів за допомогою функції експоненціального згладжування [23]. Дана модель добре працює для прогнозування рядів без вираженої тенденції або сезонності. Також є різновиди моделей, які мають в основі експоненціальне згладжування, але уже враховують тенденцію та сезонність. Якщо порівнювати з найвними методами, то найвні методи беруть до уваги лише останнє спостереження, а всі попередні спостереження не дають ніякої інформації щодо майбутнього прогнозу. Прогнози експоненціального згладжування обчислюються з використанням середньозважених ваг, але ваги зменшуються експоненціально – найменші ваги пов'язані з найдавнішими спостереженнями:

$$y_{T+h|T} = \alpha y_T + \alpha(1-\alpha)y_{T-1} + \alpha(1-\alpha)^2 y_{T-2} + \dots, \quad (2.1.)$$

де $0 \leq \alpha \leq 1$ параметр згладжування. Прогноз на один крок вперед $T+h$ є середньозваженим значенням усіх спостережень у часовому ряді y_1, y_2, \dots, y_T , швидкість, з якою знижуються ваги контролюється параметром α .

Середньо зважена форма запису простого експоненціального згладжування:

Прогноз $T+h$ є середньозваженим останнього спостереження та попереднього прогнозу:

$$y_{T+h|T} = \alpha y_T + (1-\alpha)y_{T|T-1} \quad (2.2.), \text{ де } 0 \leq \alpha \leq 1 \text{ параметр згладжування [28].}$$

Метод Холта-Вінтерса – це сезонний метод моделювання часового ряду. Він включає три різні рівні згладжування – згладжування рівня, згладжування тренду та сезонний компонент s_t з відповідними параметрами згладжування α, β^*, γ . Розрізняють адитивний та мультиплікативний методи.

Аддитивний метод Холта-Вінтера реалізується за допомогою наступного алгоритму:

$$y_{t+h|t} = \ell_t + hb_t + s_{t+h-m(k+1)} \quad (2.3.)$$

$$\ell_t = \alpha(y_t - s_{t-m}) + (1-\alpha)(\ell_{t-1} + b_{t-1}) \quad (2.4.)$$

$$b_t = \beta^*(\ell_t - \ell_{t-1}) + (1-\beta^*)b_{t-1} \quad (2.5.)$$

$$s_t = \gamma(y_t - \ell_{t-1} - b_{t-1}) + (1-\gamma)s_{t-m}, \quad (2.6.)$$

де k ціла частина $(h-1)/m$, що гарантує, що оцінки сезонних індексів, які використовуються для прогнозування, походять з останнього року вибірки. Рівняння рівня показує середньозважену між сезонно скоригованим спостереженням $y_t - s_{t-m}$ і несезонним прогнозом $\ell_{t-1} + b_{t-1}$ за час t . Рівняння тренду ідентичне лінійному методу Холта. Сезонне рівняння показує середньозважену величину між поточним сезонним індексом, $y_t - \ell_{t-1} - b_{t-1}$ і сезонний показник того ж сезону минулого року (тобто, m періодів часу минулого року).

Рівняння для сезонної складової часто виражається як:

$$s_t = \gamma^*(y_t - \ell_{t-1}) + (1-\gamma^*)s_{t-m}, \quad (2.7.)$$

Якщо підставити ℓ_t з рівняння згладжування для рівня компонентної форми вище, то отримаємо

$$s_t = \gamma^*(1-\alpha)(y_t - \ell_{t-1} - b_{t-1}) + [1-\gamma^*(1-\alpha)]s_{t-m}, \quad (2.8.)$$

Алгоритм мультиплікативного методу Холта-Вінтера набуває наступного вигляду

$$y_{t+h|t} = (\ell_t + hb_t)s_{t+h-m(k+1)} \quad (2.9.)$$

$$\ell_t = \alpha \frac{y_t}{s_{t-m}} + (1-\alpha)(\ell_{t-1} + b_{t-1}) \quad (2.10.)$$

$$b_t = \beta^*(\ell_t - \ell_{t-1}) + (1-\beta^*)b_{t-1} \quad (2.11.)$$

$$s_t = \gamma \frac{y_t}{\ell_{t-1} - b_{t-1}} + (1-\gamma)s_{t-m}, \quad (2.12.)$$

Аддитивний метод є кращим, коли сезонні зміни приблизно є постійними у часовому ряді, а мультиплікативний метод – коли сезонні коливання змінюються пропорційно рівню часового ряду. За допомогою аддитивного методу сезонна складова виражається в абсолютному виразі в масштабі спостережуваного ряду, а в рівнянні рівня сезонно коригується шляхом віднімання сезонної складової. При мультиплікативному методі сезонна складова виражається у відносних показниках (відсотках), а часовий ряд сезонно коригується шляхом ділення на сезонну складову. [28]

ARIMA–моделі. ARIMA модель (autoregressive integrated moving average) – це інтегрована модель авторегресії ковзного середнього, модель і методологія аналізу часових рядів [19].

Несезонна ARIMA – це поєднання диференціювання з авторегресійною моделлю і моделлю ковзного середнього:

$$y'_t = c + \phi_1 y'_{t-1} + \dots + \phi_p y'_{t-p} + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q} + \varepsilon_t, \quad (2.13.)$$

ARIMA має наступні параметри p , d , q , де p – порядок авторегресійної частини, d – ступінь першої диференціації, q – порядок частини ковзного середнього. [8]

Табл. 2.1. – Деякі випадки ARIMA моделей

Білий шум	ARIMA(0,0,0)
Випадкове блукання	ARIMA(0,1,0) без константи
Випадкове блукання із затуханням	ARIMA(0,1,0) з константою
Авторегресія	ARIMA(p,0,0)
Середнє ковзне	ARIMA(0,0,q)

Сезонна ARIMA модель утворюється з додаванням до несезонної моделі сезонної частини. Її можна записати наступним чином:

$$ARIMA(p,d,q)(P,D,Q)_m$$

де p,d,q – несезонна частина моделі, а $(P,D,Q)_m$ – сезонна частина, P = кількість сезонної авторегресії, D = кількість сезонних диференціювань, Q = кількість сезонних ковзних середніх і m – кількість спостережень протягом року [8, с.193].

Моделі нейронних мереж. Це мережа нейронів, які організовані по шарах. Входи утворюють нижній шар, а прогнози або ж виходи утворюють верхній шар. Схематично це виглядає як:

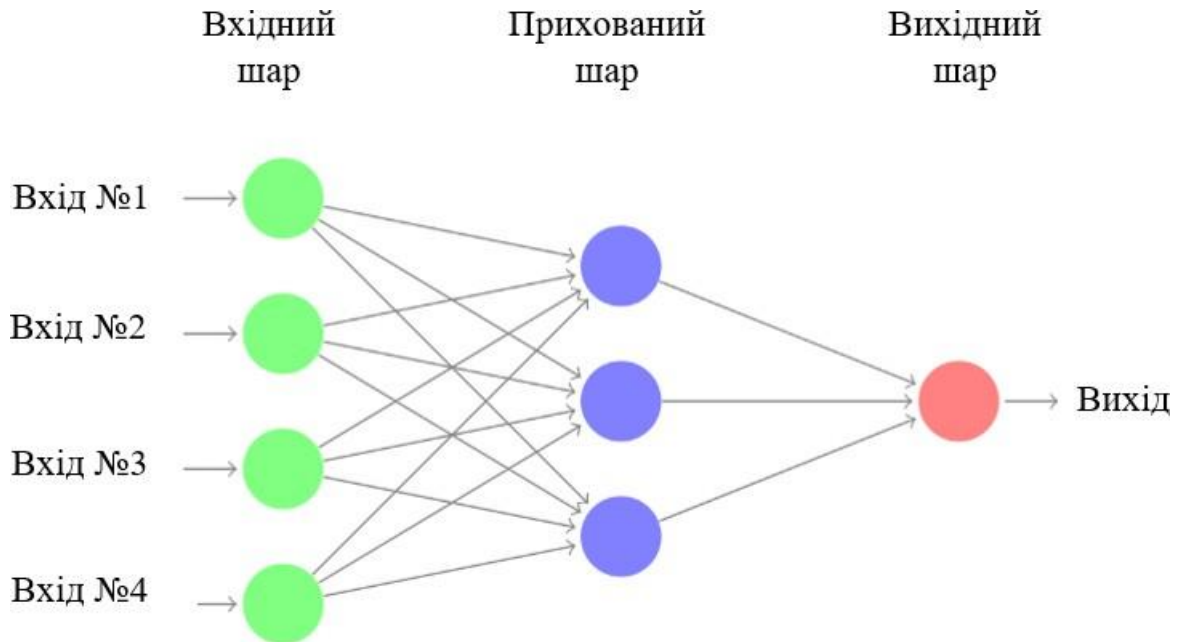


Рис.2.1. – Схематичне представлення моделі нейронних мереж

Коефіцієнти, що прикріплені до предикторів, називаються «вагами». Ваги вибираються в рамках нейронної мережі за допомогою алгоритму навчання, який мінімізує функцію витрат. Вхід у прихований нейрон j можна описати:

$$z_j = b_j + \sum_{i=1} w_{i,j} x_i \quad (2.14.)$$

У прихованому шарі це змінюється за допомогою нелінійної функції:

$$s(z) = \frac{1}{1 + e^{-z}}, \text{ щоб дати вхід для наступного шару.}$$

З даними часових рядів, лагові значення часових рядів можуть бути використані як вхідні дані для нейронної мережі. Це також називається авторегресія нейронної мережі або модель $NNAR(p,k)$, p – лагові входи і k – кількість вузлів у прихованому шарі. Для сезонної моделі застосовується модель з додатковим параметром $NNAR(p,P,k)_m$

Коли справа доходить до прогнозування, мережа застосовується ітераційно. Для прогнозування на крок вперед використовуємо наявні історичні дані. Для

прогнозування на два кроки вперед використовують одноетапний прогноз як вхідний, разом з історичними даними. Цей процес триває доти не отримують всі необхідні прогнози [28].

BATS модель – модель, яка поєднує у собі експоненціальне згладжування, перетворення Бокса-Кокса та ARMA модель для залишків. Трансформація Бокса-Кокса призначена для роботи з нелінійними даними, а модель ARMA для залишків для декорелювання часових рядів. Алгоритм BATS можна записати наступним чином:

$$\begin{aligned}
 y_t^w &= \begin{cases} \frac{y_t^w - 1}{w}, & w \neq 0 \\ \log y_t, & w = 0 \end{cases} \\
 y_t^w &= \ell_{t-1} + \phi b_{t-1} + \sum_{i=1}^m s_{t-i} + d_t \\
 \ell_t &= \ell_{t-1} + \phi b_{t-1} + \alpha d_t \\
 b_t &= \phi b_{t-1} + \beta d_t \\
 s_t &= s_{t-m} + \gamma d_t \\
 d_t &= \sum_{i=1}^p \varphi_i d_{t-i} + \sum_{i=1}^q \theta_i \varepsilon_{t-i} + \varepsilon_t
 \end{aligned} \tag{2.15.}$$

$\varepsilon_t \sim NID(0, \sigma^2)$, де $i = i$ -ий сезонності.

До переваг моделі варто віднести наступні:

1. Трансформація Бокса-Кокса може обробляти дані нелінійно, а потім дещо робить дисперсію постійною;
2. Модель ARMA для залишків може вирішити проблему автокореляції;
3. Не потрібно турбуватися про початкові значення;
4. Можна отримати не тільки точні прогнози, але й інтервальні.

До недоліків варто віднести:

1. Неможливо додати пояснювальні змінні;
2. Для сезонності на високих частотах параметр буде дуже великим [20].

Theta-метод – модель базується на концепції перетворення часових рядів. Перетворення здійснюється за допомогою θ коефіцієнта, який застосовується безпосередньо до другої похідної часових рядів [29, с. 2]. Якщо $\theta < 1$, тоді друга похідна зменшується в результаті кращого наближення довгострокової поведінки часових рядів. Якщо $\theta = 0$, то розкладена лінія перетворюється на пряму з постійним нахилом, що є звичайною лінійною регресією. З іншого боку, коли $\theta > 1$ локальна кривизна збільшуються, збільшуючи короточасні рухи часового ряду. Розкладені лінії називаються тета-лініями, позначені $Z_t(\theta)$. Ці рядки мають однакове середнє значення та нахил з вихідними даними. Однак локальні викривлення фільтруються або посилюються, залежно від значення коефіцієнта θ .

Іншими словами, процес декомпозиції має перевагу в використанні більшої кількості інформації в даних, які зазвичай не можуть бути повністю захоплені та змодельовані за рахунок екстраполяції вихідного часового ряду. Тета-лінії можна розглядати як нові часові ряди і прогнозуватись окремо, використовуючи відповідний метод прогнозування. Після того, як екстраполяція кожної тета-лінії завершена, декомпозиція проводиться за комбінованою схемою для обчислення точкових прогнозів початкового часового ряду. Формула для обчислення тета-ліній подана наступним чином:

$$Z_t = \theta y_t + (1 - \theta)(\hat{\alpha} + \hat{\beta}t), \quad (2.16.)$$

де y_t часовий ряд, $\hat{\alpha}, \hat{\beta}$ звичайні оцінки найменших квадратів. Згідно з цією точкою зору, тета-лінії можна інтерпретувати як функцію моделі лінійної регресії, застосованої безпосередньо до даних.

Для того, щоб використати тета метод потрібно виконати наступні кроки:

1. Тест на сезонність: часовий ряд перевіряється на статистично значущу сезонну поведінку, використовуючи функцію автоматичної кореляції;
2. Десезоналізація: потрібно позбутися сезонності часового ряду, застосовуючи класичний метод декомпозиції;

3. Декомпозиція: часовий ряд розкладається на дві тета-лінії, лінійну лінію регресії $Z_t(0)$ та тета-лінію для $Z_t(2)$;
4. Екстраполяція: лінійну регресію екстраполюють звичайним способом, а другу лінію екстраполюють методом простого експоненціального згладжування;
5. Поєднання: прогнози, отримані в результаті екстраполяції двох ліній, поєднуються з рівними вагами (50% -50%);
6. Ресезоналізація: результати прогнозів перетворюють з врахуванням сезонності [30].

Перетворення та коригування часових рядів є важливою частиною при побудові моделей, особливо для рядів з великими коливаннями, такі як вартість акцій та інше. Слід зауважити, перед тим як здійснювати моделювання, потрібно здійснити дослідження часових рядів, для того, щоб встановити чи взагалі потрібне коригування. Існує 4 види коригування: календарне, коригування з поправкою на населення, інфляційне коригування та математичне коригування. Мета цих коригувань це спрощення в історичних даних шляхом усунення відомих джерел варіацій.

У часових рядах, які показують значну варіацію що збільшується або зменшується з часовим лагом, може бути корисною трансформація. До найбільш популярних методів належить логарифмічна трансформація. Якщо справжні спостереження y_1, \dots, y_T , то вони трансформуються у w_1, \dots, w_T , де $w_t = \log(y_t)$. Логарифмічні перетворення є практичними, тому що вони мають свою інтерпретацію. Так, якщо ми логарифмуємо значення ціни акцій, то ми отримуємо їх дохідність.

Іноді застосовують перетворення квадратних коренів та кубів коренів. В математичній формі це виглядає як: $w_t = y_t^p$ [28].

Метод, який включає логарифмування та степеневе перетворення називається перетворення Бокса-Кокса. В математичній формі дане перетворення записується наступним чином:

$$w_t = \begin{cases} \log(y_t), \lambda = 0; \\ (y_t^\lambda - 1) / \lambda, \lambda \neq 0. \end{cases} \quad (2.17.)$$

Якщо $\lambda = 0$ перетворення є логарифмічним, а якщо $\lambda = 1$ перетворення є таким ж як і початковий ряд, за інший значень λ перетворенням буде між ними. Складність полягає у підборі параметра λ .

Після здійснення перетворення прогнозуються трансформовані дані. Для отриманих прогнозів потрібно зробити зворотну трансформацію, щоб отримати значення в початкових одиницях виміру. Зворотнє перетворення Бокса-Кокса задається методом:

$$y_t = \begin{cases} \exp(w_t), \lambda = 0; \\ (\lambda w_t + 1)^{1/\lambda}, \lambda \neq 0. \end{cases} \quad (2.18.) [21]$$

Описані методи статистичного прогнозування надають змогу з високою точністю прогнозувати курс криптовалют.

2.2. Специфіка перевірки адекватності моделей динаміки реальним процесам.

Після побудови моделі її потрібно оцінити, щоб зрозуміти чи дійсно дану модель можна використовувати, чи вона взагалі не відображає природи описуваного явища, тобто є неадекватною. Для оцінки адекватності моделі потрібно перевірити залишки моделі. Вони мають відповідати наступним вимогам:

1. Залишки є некорельовані. Якщо між залишками є кореляції, то в залишках залишається інформація, яку слід використовувати при обчисленні прогнозів;
2. Залишки мають нульове середнє. Якщо залишки мають середнє, відмінне від нуля, то прогнози є упередженими;
3. Залишки мають постійну дисперсію;
4. Залишки мають нормальний розподіл.

Будь-який метод прогнозування, який не задовольняє цим властивостям, може бути покращений. Можна мати кілька різних методів прогнозування для одного набору даних, кожен із яких задовольняє цим властивостям. Перевірка цих властивостей важлива, щоб побачити, чи використовує метод всю доступну інформацію.

Якщо будь-яка з цих властивостей не задовольняється, то метод прогнозування можна змінити, щоб дати кращі прогнози. Коригувати зміщення здійснюється наступним чином: якщо залишки мають середнє m , то просто додайте m до всіх прогнозів, і проблема зміщення буде вирішена. [п 3.3, 28]

На додачу до вищеперерахованих чотирьох властивостей, залишки потрібно перевірити на автокореляцію. Одним із таких методів є ACF та PACF графіки. Проте варто звернути увагу на більш точніші методи.

Першим з них є тест Бокса-Пірса.

$$Q = T \sum_{k=1}^h r_k^2 \quad (2.19.)$$

де h це максимальний лаг, що розглядається і T – це кількість спостережень. Якщо кожен r_k^2 близький до нуля, то й відповідно Q буде малим. Найоптимальніше використовувати $h = 10$ для несезонних даних і $h = 2m$ для сезонних даних, де m - період сезонності. Однак тест не є хорошим, коли h великий, тому якщо ці значення більше $T/5$, то краще використовувати $h = T/5$.

Пов'язаний (і більш точний) тест – це тест Льюнга-Бокса:

$$Q^* = T(T + 2) \sum_{k=1}^h (T - k)^{-1} r_k^2 \quad [28] \quad (2.20.)$$

Для перевірки точності моделі загальноприйнятою є практика відокремлення даних на навчальну та тестові вибірки. Навчальні дані використовуються для оцінки будь-яких параметрів методу прогнозування, а тестові дані використовуються для оцінки його точності.

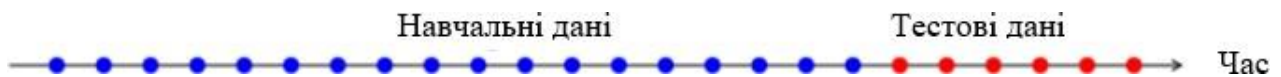


Рис. 2.2. – Схематичне зображення поділу вибірки

Розмір тестової вибірки становить до 20% від загальної вибірки та визначається як правило дослідником.

Вибір найкращої прогнозової моделі здійснюється за допомогою крос перевірки часових рядів. При горизонті прогнозування на один крок вперед вся вибірка ділиться на навчальну та тестову. Тестова вибірка складається з одного спостереження. Навчальна вибірка формується зі спостережень, які відбулися до спостережень, які формують тестовий набір даних. Оскільки неможливо отримати точний прогноз на основі малого навчального набору, ранні спостереження не розглядаються як тестові набори. Рис 2.2 ілюструє серію навчальних і тестових наборів, де синім кольором позначаються навчальні набори, а червоним тестові набори.

Процедура крос перевірки часових рядів обчислює помилки прогнозу, отримані шляхом застосування прогностичної функції до часових рядів. Для обрання кращий прогностичного методу, потрібно мінімізувати похибку прогнозу.

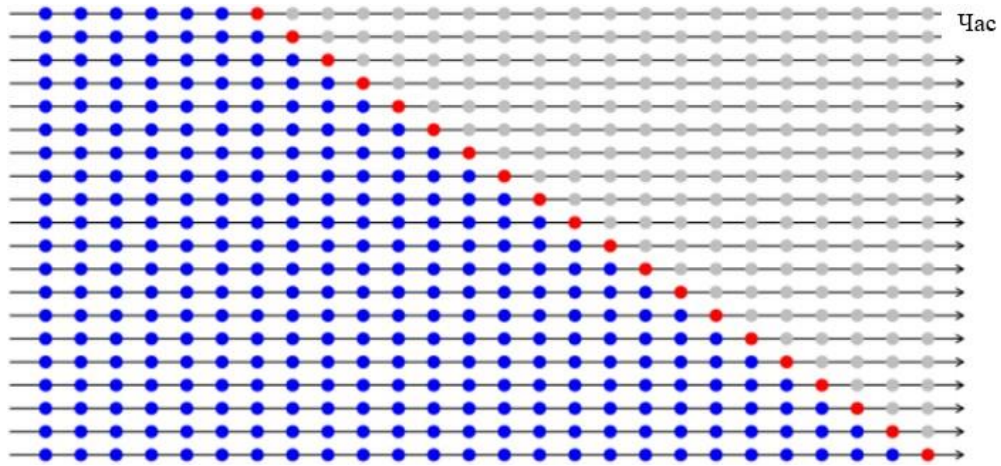


Рис 2.3. – Схематичне зображення процедури крос перевірки

Порівнюючи методи прогнозування, застосовані до одного часового ряду або до декількох часових рядів з однаковими одиницями, є два популярних показники вимірювання точності прогнозу: MAE та RMSE.

$$RMSE = \sqrt{\frac{1}{p} \sum (y_t - \hat{y}_t)^2} \quad (2.21.)$$

$$MAE = \frac{1}{p} \sum |y_t - \hat{y}_t| \quad (2.2) \quad (2.22.) \quad [34]$$

де p – кількість періодів, на які розраховують прогноз, \hat{y}_t – прогноз значення часового ряду у t -ому періоді, y_t – значення часового ряду у t -ому періоді.

При прогнозуванні часових рядів прогнози на один крок вперед можуть бути не настільки актуальними, як багатоетапні прогнози. У цьому випадку процедуру перехресної перевірки, щоб використовувати багатоетапні помилки.

Перевагою використання крос перевірки часових рядів є точність обчислення похибок прогнозу, що обраховуються на протяжності усього часового ряду. До недоліків можна віднести складність та тривалість обчислень.

2.3. Інформаційне забезпечення моделювання курсу Біткоїна, Ефіріума та Сенд.

При моделюванні ринку криптовалют важливим є зручність та швидкість отримання даних, адже ринок є дуже динамічний. Загалом, для вирішення даного завдання можна використовувати різні сайти, які пропонують історичні дані у різних форматах за певні проміжки часу. Перевагами даного рішення є простота та швидкість отримання даних, а недоліками – недостатні інтервали даних, необхідність постійного скачування файлу та його форматування. Ще одним рішенням є API (Application Programming Interface), який надають самі криптовалютні біржі та інші сервіси. Цей спосіб є набагато зручніший, адже достатньо сформулювати запит один раз та коли є необхідність отримувати дані миттєво.

Для дослідження ми будемо використовувати дані для криптовалют Біткоїн (BTC), Ефіріум (ETH) та Сенд (SAND) з періодичністю 1 день довжиною 180 днів (Рис.2.3-2.5).

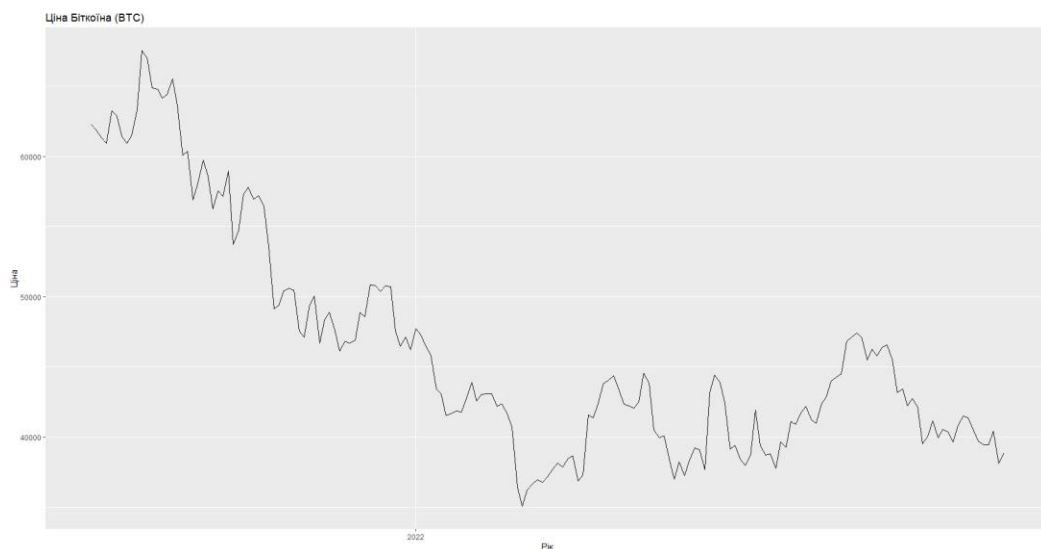


Рис. 2.4. – Графік ціни Біткоїна, періодичність 1 день

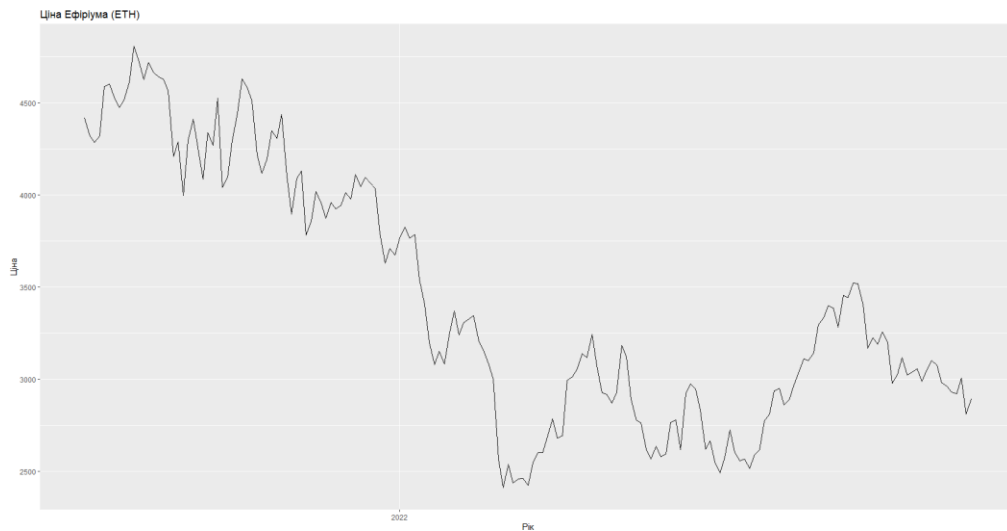


Рис. 2.5. – Графік ціни Ефіріума, періодичність 1 день

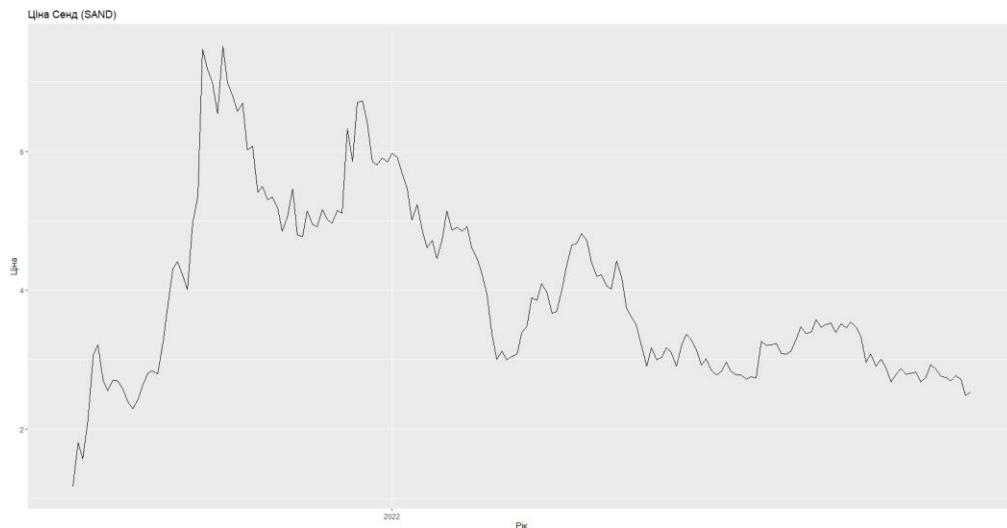


Рис. 2.6. – Графік ціни Сенд, періодичність 1 день

Як видно з графіків, ціна досить швидко змінюється від періоду до періоду, що вказує на наявність значної волатильності. Для стабілізації варіації, поширеною є практика логарифмування або перетворення Бокса-Кокса. Також можна перейти відразу до моделювання доходності криптовалют. Також важливим є дослідження ряду на стаціонарність, оскільки це буде впливати на ефективність побудованих моделей.

Augmented Dickey-Fuller Test

```
data: sand_price
Dickey-Fuller = -2.5837, Lag order = 5, p-value = 0.3323
alternative hypothesis: stationary
```

```

Augmented Dickey-Fuller Test

data: eth_price
Dickey-Fuller = -1.5933, Lag order = 5, p-value = 0.7462
alternative hypothesis: stationary

Augmented Dickey-Fuller Test

data: btc_price
Dickey-Fuller = -1.8403, Lag order = 5, p-value = 0.643
alternative hypothesis: stationary

```

Рис. 2.7. – Результати тесту Діккі-Фулера на стаціонарність рядів

Оскільки $p\text{-value} > 0.05$, для усіх трьох рядів то вони вважаються нестаціонарним. Для перетворення даних у стаціонарний ряд потрібно провести диференціацію.

```

Augmented Dickey-Fuller Test

data: dsand_price
Dickey-Fuller = -4.6351, Lag order = 5, p-value = 0.01
alternative hypothesis: stationary

Augmented Dickey-Fuller Test

data: deth_price
Dickey-Fuller = -5.4324, Lag order = 5, p-value = 0.01
alternative hypothesis: stationary

Augmented Dickey-Fuller Test

data: dbtc_price
Dickey-Fuller = -5.4937, Lag order = 5, p-value = 0.01
alternative hypothesis: stationary

```

Рис. 2.8. – Результати тесту Діккі-Фулера після диференціації першого порядку

Після диференціації першого порядку ряд став стаціонарним. Загалом, моделі, які будуть побудовані не вимагають строгої стаціонарності. Це потрібно для визначення порядку диференціації ARIMA моделі.

Також при проведенні логарифмування або перетворення Бокса-Кокса, стаціонарність зберігається, тому при побудові моделей будемо використовувати дані без застосування перетворень.

РОЗДІЛ III. СТАТИСТИЧНЕ МОДЕЛЮВАННЯ ТА ПРОГНОЗУВАННЯ КУРСУ КРИПТОВАЛЮТ

3.1. Порівняльний аналіз моделей прогнозування цін на ринку криптовалют.

При інвестуванні у криптовалюту необхідно постійно слідкувати за ситуацією, яка складається на ринку. Це включає в себе не лише відстеження динаміки ціни криптовалюту, але й побудова короткострокових та довгострокових прогнозів. Методи моніторингу можна поділити на дві великі групи: фундаментальні та технічні.

Фундаментальний аналіз заснований на економічних, політичних, соціальних та інших чинниках, але цей метод не варто застосовувати для ринку криптовалют, адже він значною мірою відокремлений від загальноприйнятої фінансової системи.

Технічний аналіз в свою чергу враховує всі події, що відбуваються на ринку. Тобто, для прогнозування ціни криптовалюту беруться до уваги лише історичні значення, що враховують стан ринку та очікування учасників. Технічний аналіз є досить точним, адже до нього відносять економіко-математичні, статистичні та інші моделі

У дослідженні використано економіко-математичні моделі та проведено порівняння моделей щодо їх продуктивності. Для моделювання та прогнозування було побудовано наступні моделі:

- Модель Холта-Вінтерса;
- Модель експоненціального згладжування;
- Theta модель;
- BATS модель;
- Моделі нейронних мереж;
- ARIMA модель.
- SARIMA модель;
- ARIMAX модель.

Результати дослідження можливо використовувати для прогнозування ціни на ринку криптовалют та відповідно будувати власну стратегію торгів, або ж використовувати для моделювання інших високоволатильних часових рядів.

Модель Холта-Вінтерса дуже добре працює з часовими рядами, що мають виражену сезонність або циклічність. Для відстеження сезонності доцільно розкласти ряд на складові:

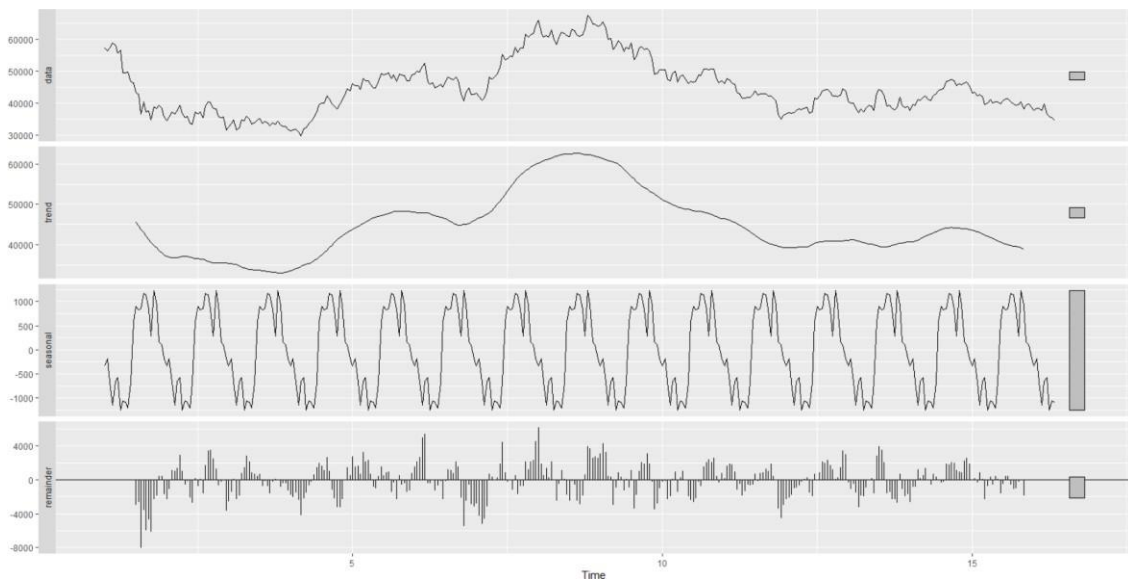


Рис. 3.1. – Розклад Ряду ціни Біткоїна на складові (тренд, сезонність, похибка)

Як видно з Рис. 3.1. сезонність присутня у ряді ціни Біткоїна, і вона складає 1-2 тижні. Також варто звернути увагу на тренд та похибку. За проміжок у 180 днів визначити глобальний тренд важко, проте дуже добре простежуються локальні тренди, що є типовим для криптовалют. Похибка або ж випадкова частина ряду є досить значною у даному ряді та складає у максимумі близько 4000\$, тобто від 5 до 10% ціни Біткоїна. Останнє зауваження ускладнює процес моделювання.

Загалом було побудовано 4 моделі Холта-Вінтерса та обрано найкращу за критерієм Акаїке.

Табл. 3.1. Значення критерія Акаїке для моделі Холта-Вінтерса

Модель	BTC	ETH	SAND
Адитивна модель	7708,54	5942,26	1272,94
Мультиплікативна модель	7800,62	6021,73	891,95
Адитивна модель із затухаючим трендом	7704,04	5939,07	1171,84
Мультиплікативна модель із затухаючим трендом	7709,93	5942,55	971,12

Виявилось, що для Біткоїна та Ефіріума кращою є адитивна модель із затухаючим трендом, а для Сенд – мультиплікативна модель із затухаючим трендом. Оскільки тренд постійно змінюється, слід підкреслити, що затухаючий тренд у цьому випадку буде показувати найкращий результат при прогнозуванні.

У модель експоненціального згладжування закладемо такі параметри (ETS): E – похибка, T – тренд та S – сезонність. Загалом можливі наступні комбінації: похибка = [A, M], тренд = [N, A, M] та сезонність = [N, A, M], де N – відсутність параметру в моделі, A – адитивний ефект, M – мультиплікативний ефект.

У випадку моделювання криптовалют варто розглянути моделі з адитивною та мультиплікативною похибкою та трендом. Сезонність не варто враховувати, як показує практика сезонні моделі погано працюють для фінансових часових рядів. Отже, варто розглянути наступні комбінації моделей: ETS(A, A, N); ETS(M, M, N); ETS(M, A, N).

Вибір моделі відбувається за інформаційним критерієм Акаїке.

Табл. 3.2. Значення критерія Акаїке для моделі ETS

Модель	BTC	ETH	SAND
ETS(A, A, N)	7697,4	2720,1	546,2
ETS(M, M, N)	7691,0	2710,9	533,6
ETS(M, A, N)	7690,7	2713,4	507,9

Оптимальні моделі для BTC – ETS(M, A, N), для ETH – ETS(M, M, N), для SAND – ETS(M, A, N) з наступними параметрами:

Табл. 3.3. Параметри моделі ETS

Параметри	BTC	ETH	SAND
α	0,91	0,99	0,99
β	0,0001	0,0001	0,0782
φ	0,82	-	0,83
l	53432,73	4448,73	0,46
b	869,28	0,99	0,83

При використанні несезонної моделі ARIMA застосуються параметри p , d , q , де p – порядок AR процесу, d – порядок диференціювання, а q – порядок MA процесу.

При відборі ARIMA моделі варто спочатку врахувати стаціонарність даних для того, щоб визначити порядок диференціювання (d), а вже потім визначати параметри p та q . Як ми уже визначили у п. 2.2, для усіх трьох криптовалют потрібно провести порядок диференціювання першого порядку. Далі потрібно дослідити ACF та PACF графіки стаціонарних даних.

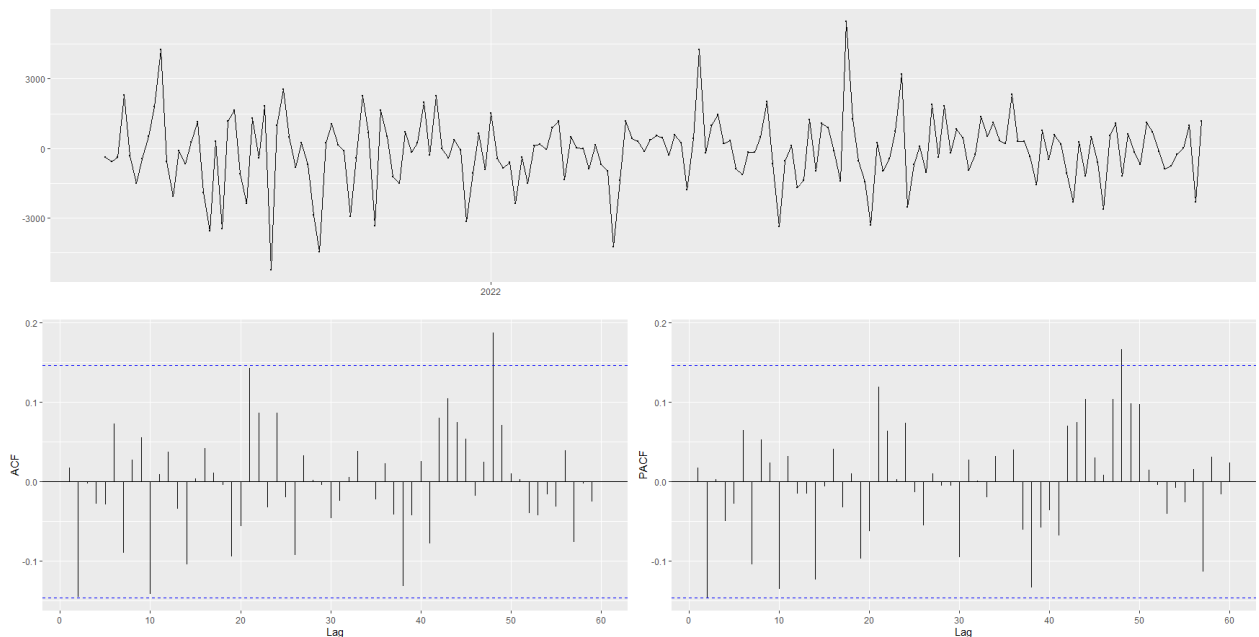


Рис. 3.2. – ACF та PACF графіки стаціонарних даних

З графіків ACF та PACF можна зробити висновок, що значних сплесків лагів за критичну зону немає, таким чином найбільш ймовірно оптимальною буде ARIMA(0, 1, 0). Щоб впевнитись у своїх припущеннях варто побудувати різні варіації ARIMA та обрати кращу за критерієм Акаїке.

ARIMA(2,1,2)	with drift	: 6518.869
ARIMA(0,1,0)	with drift	: 6520.585
ARIMA(1,1,0)	with drift	: 6521.683
ARIMA(0,1,1)	with drift	: 6520.876
ARIMA(0,1,0)		: 6518.692
ARIMA(1,1,1)	with drift	: 6522.364

Рис. 3.3. – Підбір оптимальної ARIMA моделі

Такими ж вийшли результати для рядів ETH та SAND, що дає привід стверджувати про оптимальність моделі ARIMA(0, 1, 0) для усіх трьох рядів.

Також варто враховувати, що є можливість побудувати сезонні ARIMA моделі для цих часових рядів.

Табл. 3.4. Оптимальні моделі SARIMA

BTC	ETH	SAND
ARIMA(2,1,2)(2,0,0)[24].	ARIMA(1,1,1)(2,0,0)[24]	ARIMA(3,1,0)(2,0,1)[24]

Ще одним різновидом ARIMA моделі є модель ARIMAX, яка додатково включає в себе зовнішні фактори. Дослідити вплив факторів на ряд ціни криптовалюти можна за допомогою кореляційної матриці та регресійного аналізу. Цілком ймовірно, що до таких факторів можна віднести об'єми торгів та індекс страху та жадібності.

Табл. 3.5. Кореляційна матриця для зовнішніх факторів та ціни

Кореляція	Ціна BTC	Ціна ETH	Ціна SAND
Об'єм	0,07	0,02	0,53
Індекс страху та жадібності	0,73	0,53	-0,22

Кореляційний аналіз показує значну залежність ціни BTC та ETH від індексу страху та жадібності, у SAND така залежність менш суттєва та обернена. Об'єм показує малу залежність для BTC та ETH, але для SAND така залежність є суттєвою. Це можна пояснити тим, що BTC та ETH є основними криптовалютами найбільшими за капіталізацією, тому вони створюють основну картину на ринку, індексу страху та жадібності реагує як індикатор усього ринку. SAND у свою чергу у рази менший за капіталізацією, тому значні маніпуляції з об'ємами торгів призводять до швидкого реагування ціни.

Регресійний аналіз показує таку ж ситуацію, тобто в якості зовнішніх факторів для BTC та ETH будемо використовувати індексу страху та жадібності, а для SAND об'єми торгів.

Табл. 3.6. Порівняння значення критерію Акаїке ARIMA-ARIMAX

Критерій Акаїке	BTC	ETH	SAND
ARIMA	6518,7	4714,5	114,6
ARIMAX	6518,8	4714,9	114,7

Вплив факторів не вносить значної зміни у якість моделі, а навпаки, ARIMA модель все-таки є кращою в порівнянні із ARIMAX.

Слід зауважити, Theta модель будується з дотриманням наступних етапів:

- Тест на сезонність;
- Позбавлення ряду сезонності якщо в цьому є потреба;
- Оцінка параметру α для підборі моделі простого експоненціального згладжування;
- Прогноз.

Табл. 3.7. Параметри моделі THETA

Параметри	BTC	ETH	SAND
α	0,926	0,978	0,99

Для побудови BATS моделі необхідно протестувати дуже багато різновидів моделей, адже модель враховує багато факторів, з-поміж наступних:

- Трансформація Бокса-Кокса для початкового ряду;
- Наявність або відсутність тренду;
- Наявність затухаючого тренду;
- Використання ARMA для залишків моделі;
- Сезонна або несезонна модель;
- Різні кількості гармонік, що використовуються для моделювання сезонних ефектів.

Кінцева модель обирається за критерієм Акаїке. Відповідно, кращими вийшли моделі з наявністю затухаючого тренду, використання ARMA для залишків моделі та відсутністю сезонності. Застосування трансформації Бокса-Кокса є доцільним лише для ряду Ефіріума та Сенд.

Табл. 3.8. Параметри моделі BATS

	BTC	ETH	SAND
λ	0,003	-	0,64
α	0,91	1,05	0,93
β	0,003	-0,08	0,04
<i>Damping Parameter</i>	0,8	0,8	0,80

Треба відзначити, для побудови моделі нейронних мереж скористаємось пакетом *forecast*. Відповідна функція має параметри $NNAR(p,P,k)_m$, де p – кількість лагів, P – кількість сезонних лагів, k – кількість нейронів у прихованому шарі. Для наших рядів вдалось підібрати моделі з наступними параметрами:

Табл. 3.9. Параметри моделі нейронних мереж

BTC	ETH	SAND
NNAR(1,1,2)[24]	NNAR(1,1,2)[24]	NNAR(3,1,2)[24]

Як показали результати дослідження, найкращими виявились сезонні моделі нейронних мереж. Для ряду Біткоїна та Ефіріума у моделі як входи використовуються лише останнє спостереження та два приховані шари у нейронній мережі, в той час як у Сенд три останні спостереження та два приховані шари.

3.2. Оцінювання точності моделей за допомогою крос перевірки часових рядів.

Вибір моделей одного класу проводиться переважно за критерієм Акаїке, коли ж потрібно обрати модель різних класів, то потрібно порівнювати похибки моделей. У даному випадку будемо застосовувати процедуру крос перевірки часових рядів, в якій порівнюються похибки прогностичної функції, застосовані до підбраної моделі.

Застосуємо процедуру крос перевірки часових рядів для наших побудованих моделей. Для оцінки точності прогнозів використаємо MAE та RMSE. Метод прогнозування, що мінімізує MAE, призведе до медіани прогнозів, тоді як мінімізація RMSE призведе до середнього прогнозу.

Табл. 3.10. Крос перевірка часового ряду BTC, RMSE

Метод	h=1	h=2	h=3	h=4	h=5
Модель Холта-Вінтерса	10 100,46	10 329,56	10 210,41	10 257,72	10 318,36
ARIMA	10 347,70	10 295,52	10 244,49	10 186,84	10 126,92
SARIMA	10 260,48	10 348,33	10 237,11	10 177,54	10 056,35
ARIMAX	10 348,70	10 299,52	10 247,49	10 187,84	10 130,92
ETS	10 225,45	10 174,99	10 125,55	10 068,73	10 009,36
THETA	10 262,66	10 230,67	10 199,59	10 161,02	10 119,81
NNETAR	10 233,91	10 170,90	10 112,13	10 042,65	9 972,38
BATS	10 195,74	10 136,56	10 079,63	10 016,39	9 951,48
Мінімальна похибка	10 100,46	10 136,56	10 079,63	10 016,39	9 951,48

Джерело: побудовано автором.

Для ряду Біткоїна при горизонті прогнозу на 1 крок вперед найкращим виявилась модель Холта-Вінтерса, якщо ж розглядати прогнози на довший період, то дана модель навпаки показує значне збільшення похибки, в той час як інші моделі показують навпаки зменшення похибки. Проте, якщо горизонт прогнозування більш як на 1 крок уперед, то найкращою у даному випадку виявиться BATS модель. Досить непогано результати показують моделі нейронних мереж, модель експоненційного згладжування.

Табл. 3.11. Крос перевірка часового ряду BTC, MAE

Метод	h=1	h=2	h=3	h=4	h=5
Модель Холта-Вінтерса	7 004,84	7 238,25	7 126,62	7 188,21	7 268,75
ARIMA	7 235,89	7 193,32	7 151,40	7 106,92	7 061,61
SARIMA	7 150,69	7 249,34	7 147,06	7 099,91	6 996,21
ARIMAX	7 239,89	7 196,32	7 154,40	7 107,92	7 066,61
ETS	7 117,83	7 077,01	7 036,69	6 993,08	6 948,37
THETA	7 152,73	7 129,65	7 107,45	7 082,03	7 055,54
NNETAR	7 125,78	7 073,36	7 024,46	6 969,16	6 914,34
BATS	7 090,84	7 042,04	6 994,85	6 945,34	6 895,53
Мінімальна похибка	7 004,84	7 042,04	6 994,85	6 945,34	6 895,53

Джерело: побудовано автором.

У випадку мінімізації MAE результат залишається незмінним – при горизонті прогнозу на 1 крок вперед найкращим виявилась модель Холта-Вінтерса, а при більшому горизонті найкращою є BATS модель.

Для побудови прогнозу розділимо дані на навчальну та тестову вибірки, та порівняємо прогноз із тестовою вибіркою.

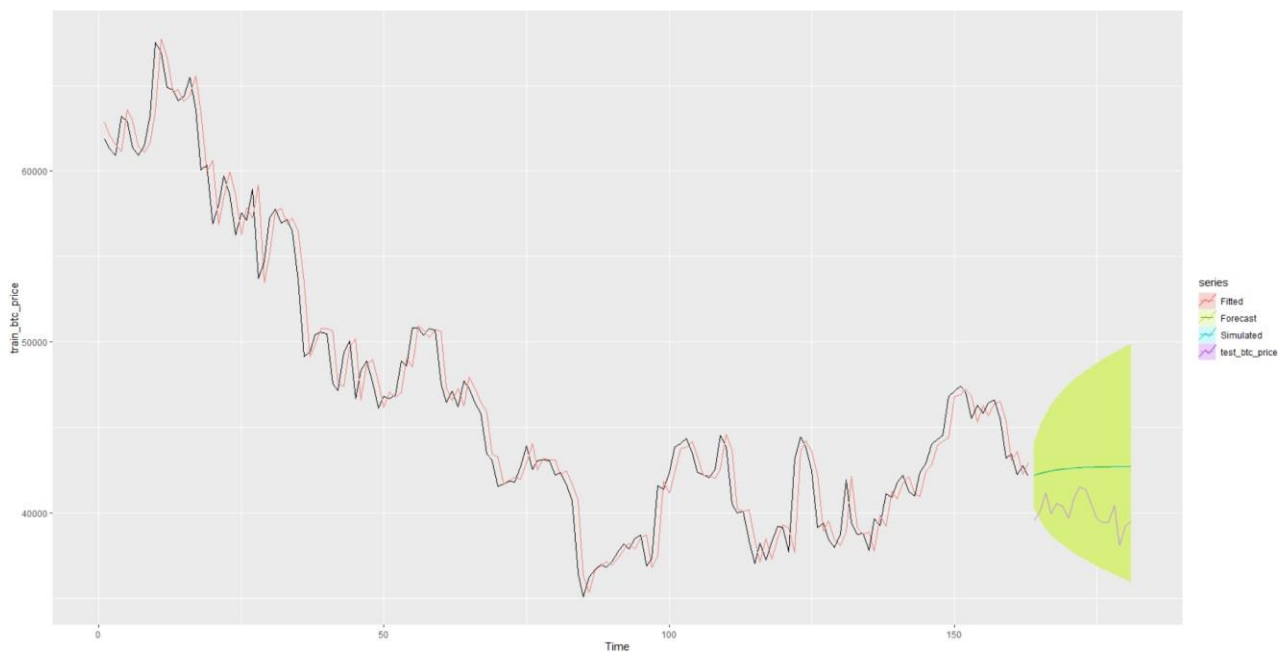


Рис. 3.4. – Реальні дані ціни BTC та прогноз

Джерело: побудовано автором.

Прогноз для навчальної вибірки виглядає більш лінійно, тому будувати прогноз на багато кроків уперед не є доцільно. Також варто звертати увагу на довірчі інтервали, що важливо при формуванні стратегії інвестора.

Табл. 3.12. Крос перевірка часового ряду ETH, RMSE

Метод	h=1	h=2	h=3	h=4	h=5
Модель Холта-Вінтерса	832,69	863,32	864,53	871,42	885,57
ARIMA	844,50	842,53	840,44	837,62	834,44
SARIMA	845,64	848,05	844,67	840,60	833,55
ARIMAX	845,50	844,53	842,44	839,62	835,44
ETS	845,77	844,20	842,50	840,06	837,27
THETA	847,62	847,16	846,57	845,24	843,55
NNETAR	858,31	858,35	858,84	858,26	857,59
BATS	841,39	838,41	835,46	831,90	828,10
Мінімальна похибка	832,69	838,41	835,46	831,90	828,10

Джерело: побудовано автором.

Для ряду Ефіріума найкращою моделлю при прогнозі на один крок уперед виявилась модель Холта-Вінтерса, а при більшому горизонті – BATS модель. Ситуація в цілому дуже схожа як і у випадку з моделюванням ряду для ціни Біткоїна.

Табл. 3.13. Крос перевірка часового ряду ETH, MAE

Метод	h=1	h=2	h=3	h=4	h=5
Модель Холта-Вінтерса	623,59	648,73	649,40	655,17	667,54
ARIMA	633,14	631,06	628,91	626,35	623,60
SARIMA	634,09	635,63	632,40	628,80	622,91
ARIMAX	634,14	635,06	630,91	629,35	624,60
ETS	634,20	632,45	630,62	628,38	625,95
THETA	635,74	634,90	633,99	632,69	631,21
NNETAR	644,72	644,37	644,43	643,81	643,25
BATS	630,59	627,67	624,81	621,64	618,38
Мінімальна похибка	623,59	627,67	624,81	621,64	618,38

Джерело: побудовано автором.

У випадку мінімізації MAE результат залишається незмінним – при горизонті прогнозу на 1 крок вперед найкращим виявилась модель Холта-Вінтерса, а при більшому горизонті найкращою є BATS модель.

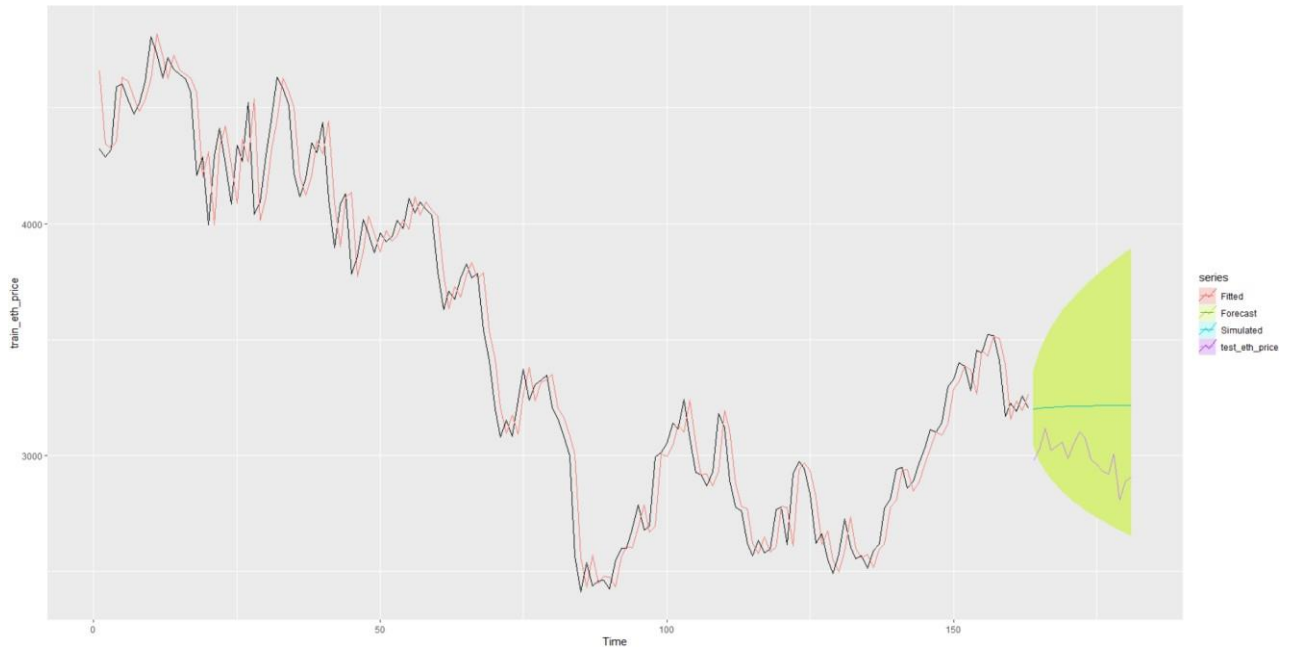


Рис. 3.5. – Реальні дані ціни ETH та прогноз

Джерело: побудовано автором.

У даному випадку прогноз ідентичний як і для ряду Біткоїна – він є лінійним.

Табл. 3.14. Крос перевірка часового ряду SAND, RMSE

Метод	h=1	h=2	h=3	h=4	h=5
Модель Холта-Вінтерса	1,94	1,99	1,99	1,99	1,98
ARIMA	1,92	1,92	1,92	1,93	1,93
SARIMA	1,89	1,90	1,90	1,91	1,91
ARIMAX	1,93	1,94	1,94	1,94	1,94
ETS	1,92	1,93	1,93	1,93	1,93
THETA	1,93	1,93	1,94	1,94	1,95
NNETAR	1,84	1,83	1,81	1,80	1,79
BATS	1,92	1,92	1,92	1,92	1,93
Мінімальна похибка	1,84	1,83	1,81	1,80	1,79

Джерело: побудовано автором.

Для ряду Сенд найкращою виявилась модель нейронних мереж при усіх горизонтах прогнозування. У випадку мінімізації MAE ситуація залишається такою ж.

Табл. 3.15. Крос перевірка часового ряду SAND, MAE

Метод	h=1	h=2	h=3	h=4	h=5
Модель Холта-Вінтерса	1,48	1,54	1,54	1,54	1,53
ARIMA	1,45	1,46	1,46	1,46	1,47
SARIMA	1,41	1,43	1,43	1,44	1,44
ARIMAX	1,49	1,49	1,47	1,48	1,49
ETS	1,46	1,46	1,46	1,47	1,47
THETA	1,47	1,47	1,48	1,48	1,49
NNETAR	1,36	1,34	1,33	1,31	1,31
BATS	1,46	1,46	1,46	1,46	1,46
Мінімальна похибка	1,36	1,34	1,33	1,31	1,31

Джерело: побудовано автором.

Якщо для ряду Біткоїна та Ефіріума вибір моделей виявився однаковим, то для Сенд ситуація зовсім інша. Це пояснюється тим, що Біткоїн та Ефіріум найбільші за капіталізацією криптовалюти, що формують ринок загалом, а динаміка Сенд може дещо відрізнятись від загальної, що формується на ринку.

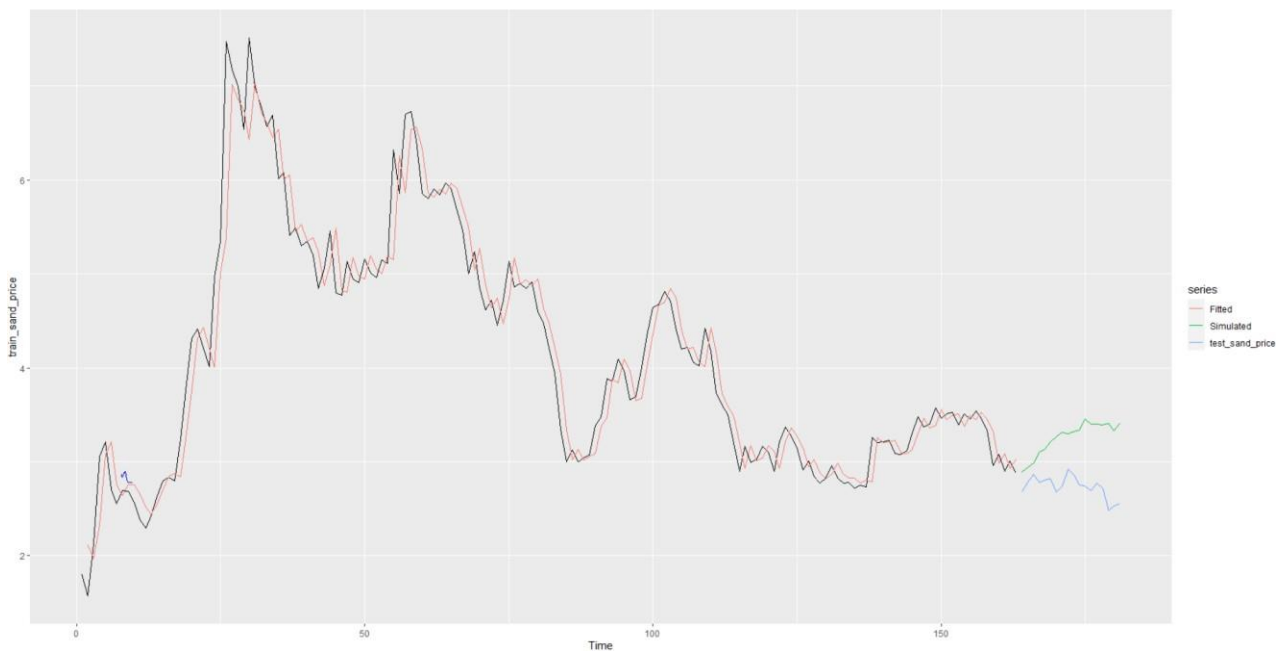


Рис. 3.6. – Реальні дані ціни SAND та прогноз

Джерело: побудовано автором.

Прогноз побудований за допомогою моделі нейронних мереж уже не є лінійним. Також можливим є симуляція прогнозів. У даному випадку було проведено 50 симуляцій та виведено середнє.

ВИСНОВКИ

Ринок криптовалют вносить значні зміни у загальний фінансовий ринок, адже це поява кардинально нової технології. Як і фондовий ринок у 30-х роках минулого століття, ринок криптовалют відкриває можливість усіх охочих інвестувати та отримувати досить значні прибутки, завдяки чому його популярність росте з кожним роком. Проте, криптовалюти – це не лише спосіб швидкого заробітку, але й технології, які уже змінюють світ у незворотному напрямку. Особливо це стосується застосування криптовалют як основного платіжного засобу у світі, адже провідні держави починають формулювати спільне бачення на визначення юридичного статусу криптовалют та розглядають можливість заміни національних валют на криптовалюти, але з державним регулюванням.

Для інвесторів ринок криптовалют – це чудова можливість заробляти великі статки за досить короткий термін часу, адже ринок є високоволатильним, проте збільшуються ризики при такому інвестуванні. Звідси впливає важливість постійного аналізу ринку та формування прогнозів.

Прогнозування є важливою частиною у будь-якій сфері для розробки ефективних цілей та планів і забезпечення ефективного функціонування будь-якого суб'єкта економічної діяльності. Побудова обґрунтованих прогнозів допомагає усвідомити майбутнє становище, при використанні теперішньої стратегії у майбутньому.

При побудові прогнозу потрібно дотримуватись п'яти основних кроків:

- 1) Визначення проблеми – розуміння того для чого будується прогноз, як він буде використовуватись, хто його буде використовувати;
- 2) Збір інформації;
- 3) Попередній аналіз – визначення певних закономірностей, виявлення наявної сезонності, циклічності у часових рядах;
- 4) Вибір та підбір моделі – на цьому етапі дослідникові потрібно зробити перетворення часового ряду. Якщо часовий ряд сезонний, то потрібно зробити його сезонно адаптованим. Наступним кроком йде підбір відповідної

моделі прогнозування. Зазвичай моделей для прогнозування вибирається декілька і них вибирається одна найкраща. Найзручніше вибрати найкращу модель за обчисленими похибками прогнозу використовуючи крос перевірку часових рядів.

5) Використання моделі прогнозування – після вибору моделі та її оцінки, модель використовують для прогнозування.

У даній роботі було дотримано усіх п'яти кроків та побудовано прогнози за допомогою моделей, що були обрані відповідно до похибки прогнозової моделі.

Для дослідження використані часові ряди ціни для криптовалют – Біткоїна, Ефіріума та Сенд, побудовано ефективні моделі та проведено тестування за допомогою крос валідації та оцінено середньоквадратичні та абсолютні похибки прогнозів. Для ряду ціни Біткоїна та Ефіріума найкращими виявились модель Холта-Вінтерса при прогнозі на один крок уперед та BATS модель при прогнозуванні на більш довший період. Для ряду ціни Сенд найкращою виявились моделі нейронних мереж, при чому на всіх горизонтах прогнозування. Така різниця у виборі моделей обґрунтовується тим, що Біткоїн та Ефіріум є найбільшими за капіталізацією криптовалютами та мають торгові пари до інших криптовалют.

При прогнозуванні високоволатильних часових рядів на значний горизонт неможливо досягнути високої точності прогнозу, тому варто завжди враховувати довірчі інтервали прогнозу при формуванні торгової стратегії.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Артем Г. Блокчейн: Как это работает и что нас ждет завтра / Г. Артем, М. Алексей., 2018. – 58 с.
2. В Україні остаточно легалізували криптовалюту: Зеленський підписав закон [Веб-сайт]. – Режим доступу <https://minfin.com.ua/ua/2022/03/16/82176211/>
3. Волатильність [Веб-сайт]. – Режим доступу <https://uk.wikipedia.org/wiki/Волатильність>
4. Волатильность криптовалют – объясняем простыми словами [Веб-сайт]. – Режим доступу: <https://coinpost.ru/p/volatilnost-kriptoalyut-obuyasnyаем-prostymi-slovami>
5. Волатильность криптовалют - особенности, причины колебаний стоимости цифровых монет[Веб-сайт]. – Режим доступу: <http://profinvestment.com/volatilnost-kriptoalyut/>
6. Дученко М. М. Вплив криптовалют на економіку країни / М. М. Дученко, Т. В. Павленко. // ЕКОНОМІКА І СУСПІЛЬСТВО. – 2018. – №19. – С. 1002–1009.
7. Е-гривня [Електронний ресурс] // Wikipedia [Інтернет-портал] . – Електронні дані. – URL <https://uk.wikipedia.org/wiki/Е-гривня>
8. Економетрика // [Черняк О. І., Комашко О.В., Ставицький А. В., Баженова О. В.] – Київ: Київський національний університет імені Тараса Шевченка, ВПЦ "Київський університет", 2009 – 395 с.
9. Загальна капіталізація ринку криптовалют [Веб-сайт]. – Режим доступу <https://ru.tradingview.com/markets/cryptocurrencies/global-charts/>
10. Захаркін О.О. Оцінка волатильності криптовалютного ринку порівняно з іншими інструментами інвестування в Україні / О.О. Захаркін, Л.С. Захаркіна, О.О. Москальова // Збірник наукових праць : «Проблеми системного підходу в економіці». – Київ : НАУ, Видавничий дім «Гельветика», 2017. – Випуск 6(62). Ч. 2. – С. 85-91.

11. Історичні дані по криптовалютам [Електронний ресурс] : [Веб-сайт]. – Режим доступу <https://ru.investing.com/crypto/>
12. Криптовалюта [Електронний ресурс] : [Веб-сайт] – Режим доступу: <https://ru.wikipedia.org/wiki/Криптовалюта>
13. Національний Банк України [Електронний ресурс]: [Веб-сайт]. – Режим доступу https://old.bank.gov.ua/control/uk/publish/article?art_id=123193&cat_id=122978
14. Національний Банк України[Електронний ресурс] : [Веб-сайт]. – Режим доступу: <https://bank.gov.ua/news/all/spilna-zayava-finansovih-regulyatoriv-schodo-statusu-kriptoalyut-v-ukrayini>
15. Перевіряємо, що вигідніше: біткоїн чи біткоїн-ETF [Веб-сайт]. – Режим доступу <https://minfin.com.ua/ua/currency/articles/scho-vigidnishe-bitkoin-chi-bitkoin-etf-pereviryaemo-na-praktici/>
16. Петрук О. М. Сутність криптовалюти як методологічна передумова її облікового відображення [Електронний ресурс] / О. М. Петрук, О. С. Новак // ВІСНИК ЖДТУ. 2017. № 4 (82). – 2017. – Режим доступу до ресурсу: <http://ven.ztu.edu.ua/article/view/119350/116061>.
17. Пряников М. М. Блокчейн как коммуникационная основа формирования цифровой экономики: преимущества и проблемы [Електронний ресурс] / М. М. Пряников, А. В. Чугунов // International Journal of Open Information Technologies. – 2017. – Режим доступу до ресурсу: <https://cyberleninka.ru/article/n/blokcheyn-kak-kommunikatsionnaya-osnova-formirovaniya-tsifrovoy-ekonomiki-preimuschestva-i-problemy/viewer>
18. 3RD GLOBAL CRYPTOASSET BENCHMARKING STUDY [Електронний ресурс] / [В. Apolline, P. Dr. Gina, W. Yue та ін.] // Cambridge Centre for Alternative Finance. – 2020. – Режим доступу до ресурсу: <https://www.jbs.cam.ac.uk/wp-content/uploads/2021/01/2021-ccaf-3rd-global-cryptoasset-benchmarking-study.pdf>.
19. ARIMA [Електронний ресурс] // Wikipedia [Інтернет-портал] . – Електронні дані. – URL <https://ru.wikipedia.org/wiki/ARIMA>

20. BATS and TBATS model [Веб-сайт]. – Режим доступу <https://yintingchou.com/posts/bats-and-tbats-model/>
21. Christoph Bergmeir. Bagging Exponential Smoothing Methods using STL Decomposition and Box-Cox Transformation [Електронний ресурс] / Christoph Bergmeir, Rob J Hyndman, Jose M Benitez // Monash University. – 2014. – Режим доступу до ресурсу: <https://robjhyndman.com/papers/BaggedETS-wpaper.pdf>.
22. Cryptocompare API [Веб-сайт]. – Режим доступу: <https://min-api.cryptocompare.com/>
23. Exponential_smoothing [Веб-сайт]. – Режим доступу: https://en.wikipedia.org/wiki/Exponential_smoothing
24. Grayscale [Веб-сайт]. – Режим доступу <https://grayscale.com/>
25. Kiril D. Forecasting cryptocurrency markets through the use of time series models [Електронний ресурс] / D. Kiril, K. Stanimir, D. Desislav // ВЕН - Business and Economic Horizons Volume 15 | Issue 2 | 2019 |pp.242-253. – 2019. – Режим доступу до ресурсу: https://ageconsearch.umn.edu/record/301145/files/201907240621_14_BEH_2019_Vol15_Issue2_Kiril-Desev_et_al_Forecasting_cryptocurrency_markets_242-253.pdf.
26. Nashirah A. B. Autoregressive Integrated Moving Average (ARIMA) Model for Forecasting Cryptocurrency Exchange Rate in High Volatility Environment: A New Insight of Bitcoin Transaction [Електронний ресурс] / A. B. Nashirah, R. Sofian // International Journal of Advanced Engineering Research and Science (IJAERS). – 2017. – Режим доступу до ресурсу: https://www.researchgate.net/publication/321375630_Autoregressive_Integrated_Moving_Average_ARIMA_Model_for_Forecasting_Cryptocurrency_Exchange_Rate_in_High_Volatility_Environment_A_New_Insight_of_Bitcoin_Transaction#read.
27. Ripple [Електронний ресурс] // Wikipedia [Інтернет-портал] . – Електронні дані. – URL <https://uk.wikipedia.org/wiki/Ripple>
28. Rob J H. Forecasting: Principles and Practice [Електронний ресурс] / Н. Rob J, A. George // Monash University, Australia. – 2018. – Режим доступу до ресурсу: <https://otexts.com/fpp2>

29. Rob J Hyndman. Unmasking the Theta method // Rob J Hyndman, Baki Billah 2001 – 7p [Электронный ресурс]– Режим доступа <https://robjhyndman.com/papers/Theta.pdf>

30. The Optimised Theta Method [Электронный ресурс] / José Augusto Fioruci, Tiago Ribeiro Pellegrini, Francisco Louzada, Fotios Petropoulos – Режим доступа до ресурсу: <https://arxiv.org/ftp/arxiv/papers/1503/1503.03529.pdf>

ДОДАТКИ

```

library(forecast)
library(ggplot2)
library(fpp)
library(seasonal)
library(httr)
library(jsonlite)
library(xml2)
library(openxlsx)
library(rugarch)
#####
## Отримання даних через API
#####
base <- "https://min-api.cryptocompare.com"
endpoint <-
"/data/v2/histoday?fsym=BTC&tsym=USDT&limit=180&e=Binance&api_key="
api_key<-
"08cfba9880df3a43df6a7cae249cb612dde3f663a6d720b2b7895245985d7748"
call <- paste0(base, endpoint, api_key)
get_btc_prices <- GET(call)
get_btc_prices_text <- content(get_btc_prices,"text", encoding = "ISO-8859-1")
# Convert to "character"
get_btc_prices_json <- fromJSON(get_btc_prices_text, flatten = TRUE) # Flatten
into list
rates <- ts(get_btc_prices_json$Data$Data$close, start = 1, frequency = 365)
btc_price<-ts(rates,start = c(2021,10), frequency = 24)
btc_plot <- autoplot(btc_price)+ylab('Ціна')+xlab('Рік')+ggtitle("Ціна Біткоїна
(BTC)")

get_fear <- GET('https://api.alternative.me/fng/?limit=180&date_format=world')

```

```

get_fear_text <- content(get_fear,"text", encoding = "ISO-8859-1") # Convert to
"character"
get_fear_json <- fromJSON(get_fear_text, flatten = TRUE) # Flatten into list
fear_greed <-
ts(get_fear_json$data$value_classification[length(get_fear_json$data$value_classification):1], start = Sys.Date() - 180)
value_fear_greed <-
ts(get_fear_json$data$value[length(get_fear_json$data$value):1], start = Sys.Date() -
180)
volume <-
ts(get_btc_prices_json$Data$Data$volumeto[2:length(get_btc_prices_json$Data$Data$a$volumeto)], start = Sys.Date() - 180)
train_index <- ts(fear_greed[1:round(0.8*length(fear_greed), digits = 0)])
test_index <- ts(fear_greed[length(train_index):length(fear_greed)])
train_volume <- ts(volume[1:round(0.8*length(volume), digits = 0)])
test_volume <- ts(volume[length(train_volume):length(volume)])
#####
## Базові графіки та перевірка на стаціонарність
#####
autoplot(btc_price)+xlab("Час")+ylab("дол. США")
ggtsdisplay(btc_price)
dbtc_price <- diff(btc_price)
ggtsdisplay(dbtc_price)
adf.test(btc_price)
adf.test(dbtc_price)
#####
## Побудова моделей
#####
#HW
hw_additive<-hw(btc_price, seasonal = "additive")

```

```

checkresiduals(hw_additive)
hw_add<- function(btc_price, h) forecast(hw_additive, h = h)
hw_additive_d <- hw(btc_price, damped = T)
checkresiduals(hw_additive_d)
hw_add_d <- function(btc_price, h) forecast(hw_additive_d, h = h)
#ARIMA
arima_nons<-arima(btc_price, order = c(2,1,2))
checkresiduals(arima_nons)
ar_non<- function(btc_price, h) forecast(arima_nons, h = h)
arima_seas<-auto.arima(btc_price, seasonal = T, ic = c("aicc"))
checkresiduals(arima_seas)
ar_seas<- function(btc_price, h) forecast(arima_seas, h = h)
#ARIMAX
arimax <- auto.arima(btc_price, xreg = volume, seasonal = F)
checkresiduals(arimax)
arimaxf <- function(btc_price, h) forecast(arimax, xreg = volume, h = h)
#ETS
ets_MMN <- ets(btc_price, model = "MMN")
checkresiduals(ets_MMN)
MMN_est<- function(btc_price, h) forecast(ets_MMN, h = h)
ets_MAN<- ets(btc_price, model = "MAN")
checkresiduals(ets_MAN)
MAN_est<- function(btc_price, h) forecast(ets_MAN, h = h)
#THETA
theta<-thetaf(btc_price)
checkresiduals(theta$residuals) #not ok
theta_f<-function(seas_btc_price,h) forecast(theta,h=h)
#NNETAR
fitnnetar<-nnetar(btc_price)
checkresiduals(fitnnetar)

```

```

netar<-function(btc_price,h) forecast(fitnnetar,h=h)
#BATS
fitbats <- bats(btc_price,use.arma.errors= TRUE, use.damped.trend=TRUE)
checkresiduals(fitbats)
bats_f<-function(btc_price,h) forecast(fitbats,h=h)
#####
## Крос перевірка
#####
rmse <- data.frame()
mae <- data.frame()
colnames(rmse) <- c("HW_a", "HW_a_d", "ARIMA",
"SARIMA","ETS_MMN", "ETS_MAN", "THETA", "NNETAR", "BATS")
colnames(mae) <- c("HW_a", "HW_a_d", "ARIMA", "SARIMA","ETS_MMN",
"ETS_MAN", "THETA", "NNETAR", "BATS")
for (i in 1:5) {
  #HW
  e1<-tsCV(btc_price, hw_add, h=i)
  r1 <- sqrt(mean(e1^2, na.rm=TRUE))
  m1 <- mean(abs(e1), na.rm=TRUE)
  e2<-tsCV(btc_price, hw_add_d, h=i)
  r2 <- sqrt(mean(e2^2, na.rm=TRUE))
  m2 <- mean(abs(e2), na.rm=TRUE)
  #ARIMA
  e3<-tsCV(btc_price, ar_non, h=i)
  r3 <- sqrt(mean(e3^2, na.rm=TRUE))
  m3 <- mean(abs(e3), na.rm=TRUE)
  e4<-tsCV(btc_price, ar_seas, h=i)
  r4 <- sqrt(mean(e4^2, na.rm=TRUE))
  m4 <- mean(abs(e4), na.rm=TRUE)
  #ETS

```

```

e5<-tsCV(btc_price, MMN_est, h=i)
r5 <- sqrt(mean(e5^2, na.rm=TRUE))
m5 <- mean(abs(e5), na.rm=TRUE)
e6<-tsCV(btc_price, MAN_est, h=i)
r6 <- sqrt(mean(e6^2, na.rm=TRUE))
m6 <- mean(abs(e6), na.rm=TRUE)
#THETA
e7<-tsCV(btc_price, theta_f, h=i)
r7 <- sqrt(mean(e7^2, na.rm=TRUE))
m7 <- mean(abs(e7), na.rm=TRUE)
#NNETAR
e8<-tsCV(btc_price, netar, h=i)
r8 <- sqrt(mean(e8^2, na.rm=TRUE))
m8 <- mean(abs(e8), na.rm=TRUE)
#BATS
e9<-tsCV(btc_price, bats_f, h=i)
r9 <- sqrt(mean(e9^2, na.rm=TRUE))
m9 <- mean(abs(e9), na.rm=TRUE)
r <- vector()
m <- vector()
r <- append(r, c(r1,r2,r3,r4,r5,r6,r7,r8,r9))
m <- append(m, c(m1,m2,m3,m4,m5,m6,m7,m8,m9))
rmse <- rbind(rmse, r)
colnames(rmse) <- c("HW_a", "HW_a_d", "ARIMA",
"SARIMA","ETS_MAM", "ETS_MNN", "THETA", "NNETAR", "BATS")
mae <- rbind(mae, m)
colnames(mae) <- c("HW_a", "HW_a_d", "ARIMA",
"SARIMA","ETS_MAM", "ETS_MNN", "THETA", "NNETAR", "BATS")
}
rmse %>% write.xlsx ('rmse.xlsx', overwrite = T)

```

```

mae %>% write.xlsx ('mae.xlsx', overwrite = T)
## Побудова прогнозів
train_btc_price <- ts(btc_price[1:round(length(btc_price)*0.9,0)], frequency =
24)
test_btc_price <-
ts(btc_price[(round(length(btc_price)*0.9,0)+1):length(btc_price)], start =
(round(length(btc_price)*0.9,0)+1))
b <- bats(train_btc_price,use.arma.errors= TRUE, use.damped.trend=TRUE)
autoplot(train_btc_price) + autolayer(b$fitted, series = "Fitted") +
autolayer(forecast(b, h=length(test_btc_price))) +autolayer(test_btc_price, series =
"Test btc_price")
autoplot(train_btc_price)+autolayer(b$fitted, series =
"Fitter")+autolayer(simulate(b, nsim = length(test_btc_price)), series = "Simulated")
+autolayer(test_btc_price, series = "Test btc_price")
df <- data.frame(one = seq(1:length(test_btc_price)))
for (i in 1:50) {
  qwerty <- simulate(b, nsim = length(test_btc_price))
  df <- cbind(df, data.frame(qwerty))
}
df <- df[,-1]
fc <- ts(rowMeans(df), start = (round(length(btc_price)*0.9,0)+1))
autoplot(train_btc_price) + autolayer(fc, series = "sim") +
autolayer(test_btc_price)
autoplot(train_btc_price) + autolayer(b$fitted, series = "Fitted")+
autolayer(forecast(bats(train_btc_price,use.arma.errors= TRUE,
use.damped.trend=TRUE), simulate = T, bootstrap = T, level = c(80), h =
length(test_btc_price)), series = "Forecast")+
autolayer(test_btc_price) + autolayer(fc, series = "Simulated")

```

**КИЇВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ
ІМЕНІ ТАРАСА ШЕВЧЕНКА**

**Економічний факультет
Кафедра статистики, інформаційно-аналітичних систем і демографії**

**ЗАВДАННЯ
на кваліфікаційну роботу магістра
студента 2 курсу**

Спеціальності 051 Економіка за освітньо-науковою програмою
«Економічна аналітика та статистика»

Юрія Стащука

1. Тема роботи: «Порівняльний аналіз методів прогнозування курсу криптовалют».
2. Термін завершення роботи: 11 травня 2022 року.
3. Об'єкт дослідження: ринок криптовалют.
4. Предмет дослідження: теоретико-методологічні та практичні засади статистичного прогнозування ціни криптовалют.
5. Мета дослідження: побудова прогнозів курсу криптовалют, порівняння різних моделей, оцінювання їх якості.
6. Завдання:
 - 6.1. визначити теоретичні особливості побудови моделей часових рядів;
 - 6.2. обґрунтувати методологію дослідження ринку криптовалют;
 - 6.3. охарактеризувати інформаційне забезпечення щодо історичних даних про курс криптовалют;
 - 6.4. дослідити особливості часових рядів та провести їх графічний аналіз;
 - 6.5. побудувати моделі часових рядів цін Біткоїна, Ефіріума та Сенд;
 - 6.6. виконати порівняльний аналіз моделей прогнозування;
 - 6.7. оцінити адекватність моделей реальним процесам та точність прогнозів ціни криптовалют;
 - 6.8. розробити рекомендації щодо використання напрацьованих (апробованих) алгоритмів для моделювання високоволатильних часових рядів.

Науковий керівник: канд.екон.наук, доцент Тетяна ТРУБНІК _____

Студент: Юрій СТАЩУК _____

(підпис)

Затверджено на засіданні кафедри статистики, інформаційно-аналітичних
систем і демографії

Протокол №12 від «12» травня 2022 року

**Календарний план виконання кваліфікаційної роботи магістра
студента 2 курсу
Спеціальності 051 Економіка за освітньо-науковою програмою
«Економічна аналітика та статистика»
Юрія СТАЩУКА**

№ етапу	Етап виконання роботи	Термін виконання	Відмітки керівника про виконання
1	Вибір теми кваліфікаційної роботи	04.10.2021-10.10.2021	
2	Формулювання мети роботи та завдань дослідження	04.10.2021-10.10.2021	
3	Пошук літературних джерел та аналіз наявної статистичної інформації	22.11.2021-05.12.2021	
4	Розробка плану та складання змісту кваліфікаційної роботи	05.12.2021-12.12.2021	
5	Добір статистичного інструментарію	13.12.2021-26.12.2021	
6	Робота над першим розділом кваліфікаційної роботи	27.12.2021-16.01.2022	
7	Перевірка першого розділу науковим керівником та доопрацювання згідно із зауваженнями	17.01.2022-23.01.2022	
8	Вибір криптовалют для здійснення моделювання. Отримання даних	24.01.2022-30.01.2022	
10	Доопрацювання другого розділу роботи, перевірка його науковим керівником	31.01.2022-13.03.2022	
11	Виконання розрахунків для третього розділу	14.03.2022-27.03.2022	
12	Написання третього розділу кваліфікаційної роботи	28.03.2022-10.04.2022	
13	Перевірка третього розділу науковим керівником та доопрацювання згідно із зауваженнями	11.04.2022-17.04.2022	
14	Написання висновків, оформлення додатків, списку використаних джерел, підготовка роботи до захисту	18.04.2022-24.04.2022	
15	Отримання відгуку від керівника кваліфікаційної роботи	09.05.2022-15.05.2022	
17	Рецензування кваліфікаційної роботи	09.05.2022-15.05.2022	
18	Перевірка дотримання академічної доброчесності на текстові запозичення	09.05.2022-15.05.2022	
19	Попередній захист виконаної кваліфікаційної роботи на кафедрі	09.05.2022-15.05.2022	
20	Допуск до захисту кваліфікаційної роботи	09.05.2022-15.05.2022	
21	Захист кваліфікаційної роботи на засіданні екзаменаційної комісії	24.05.2022	

Студент 2 курсу магістратури Юрій СТАЩУК _____

Науковий керівник канд.екон.наук, доц. Тетяна ТРУБНІК _____