

Міністерство освіти і науки України
Харківський національний університет радіоелектроніки

Факультет інформаційно-аналітичних технологій та менеджменту
(повна назва)

Кафедра прикладної математики
(повна назва)

КВАЛІФІКАЦІЙНА РОБОТА Пояснювальна записка

рівень вищої освіти другий (магістерський)

Прогнозування цін на ринку криптовалют
методами машинного навчання

(тема)

Виконав:

студент 2 курсу, групи САУм-22-1

Максименко А.С.

(прізвище, ініціали)

Спеціальність 124 Системний аналіз

(код і повна назва спеціальності)

Тип програми освітньо-професійна

(освітньо-професійна або освітньо-наукова)

Освітня програма Системний аналіз і управління

(повна назва освітньої програми)

Керівник доц. Гибкіна Н.В.

(посада, прізвище, ініціали)

Допускається до захисту

Зав. кафедри ПМ

(підпис)

Сидоров М.В.

(прізвище, ініціали)

2024 р.

Харківський національний університет радіоелектроніки

Факультет інформаційно-аналітичних технологій та менеджменту

Кафедра прикладної математики

Рівень вищої освіти другий (магістерський)

Спеціальність 124 Системний аналіз

(код і повна назва)

Тип програми освітньо-професійна

(освітньо-професійна або освітньо-наукова)

Освітня програма Системний аналіз і управління

(повна назва)

ЗАТВЕРДЖУЮ:

Зав. кафедри ПМ _____

(підпис)

“ 06 ” листопада 2023 р.

ЗАВДАННЯ
НА КВАЛІФІКАЦІЙНУ РОБОТУ

студентові Максименку Андрію Сергійовичу
(прізвище, ім'я, по батькові)

1. Тема роботи Прогнозування цін на ринку криптовалют методами машинного навчання

затверджена наказом по університету від 2 листопада 2023 р. № 1277 Ст

2. Термін подання студентом роботи до екзаменаційної комісії 10 січня 2024 р.

3. Вихідні дані до роботи статистичні дані про коливання цін на ринку криптовалют, архітектури рекурентних нейронних мереж

4. Перелік питань, що потрібно опрацювати в роботі _____

1. Системний аналіз предметної області

2. Вибір і обґрунтування методу розв'язання

3. Програмна реалізація

4. Результати обчислювального експерименту

5. Аналіз можливих застосувань

5. Перелік графічного матеріалу із зазначенням креслеників, схем, плакатів, комп'ютерних ілюстрацій _____

1. Актуальність теми роботи _____

2. Постановка задачі _____

3. Системний аналіз предметної області _____

4. Метод чисельного аналізу _____

5. Результати обчислювального експерименту _____

КАЛЕНДАРНИЙ ПЛАН

№	Назва етапів роботи	Терміни виконання етапів роботи	Примітка
1	Підбір та вивчення технічної літератури за темою роботи	6 – 12 листопада 2023 р.	виконано
2	Вибір та обґрунтування методу	13 – 26 листопада 2023 р.	виконано
3	Розробка алгоритму і програми	27 листопада – 10 грудня 2023 р.	виконано
4	Проведення аналітичних досліджень та розрахунків	11 грудня – 24 грудня 2023 р.	виконано
5	Робота над текстом пояснювальної записки	25 грудня 2023 р. – 9 січня 2024 р.	виконано
6	Представлення роботи на рецензію в ЕК	10 січня 2024 р.	виконано

Дата видачі завдання 6 листопада 2023 р.

Студент _____
(підпис)

Керівник роботи _____ доц. Гибкіна Н.В.
(підпис) (посада, прізвище, ініціали)

РЕФЕРАТ

Пояснювальна записка: 61 с., 6 табл., 27 рис., 1 дод., 21 джерело.

КРИПТОВАЛЮТА, ЗАДАЧА ПРОГНОЗУВАННЯ, МАШИННЕ НАВЧАННЯ, РЕКУРЕНТНА НЕЙРОННА МЕРЕЖА, LSTM, GRU.

Об'єкт дослідження – процес змінення цін криптовалют.

Мета роботи – прогнозування цін на ринку криптовалют за допомогою методів машинного навчання.

Методи дослідження – рекурентні нейронні мережі.

У кваліфікаційній роботі розглянуто задачу прогнозування цін на ринку криптовалют за допомогою методів машинного навчання. Виконано системний аналіз предметної області, в результаті обрано у якості методу розв'язання рекурентні нейронні мережі. На основі обраного методу розроблено програмний продукт, який може використовуватись для прогнозування, проведено обчислювальні експерименти та проаналізовано отримані результати. Зроблено висновки щодо можливості його застосування у подальших дослідженнях.

ABSTRACT

Introductory note: 61 pages, 6 tables, 27 figures, 1 appendix, 21 sources.

CRYPTOCURRENCY, FORECASTING TASK, MACHINE LEARNING, RECURRENT NEURAL NETWORK, LSTM, GRU.

Object of research – the process of changing cryptocurrency prices.

Purpose of work – forecasting prices on the cryptocurrency market using machine learning methods.

Methods of research – recurrent neural networks.

The qualification work the problem of price forecasting in the cryptocurrency market using machine learning methods was considered. A system analysis of the subject area was performed, and as a result, recurrent neural networks were chosen as a solution method. Based on the chosen method, a software product that can be used for forecasting is developed, computational experiments are carried out and the results are analyzed. Conclusions are drawn regarding the possibility of its application in further research.

ЗМІСТ

	С.
Вступ	8
1 Системний аналіз предметної області та постановка задач дослідження	10
1.1 Системний аналіз задачі прогнозування цін на ринку	
криптовалют методами машинного навчання	10
1.1.1 Вербальна модель системи	10
1.1.2 Морфологічний опис системи	10
1.1.3 Функціональна модель системи	12
1.2 Аналіз сценаріїв вирішення задачі прогнозування цін на ринку	
криптовалют методами машинного навчання	15
1.2.1 Модель аналізу проблеми	15
1.2.2 Оцінювання вектора пріоритетів незадоволеностей методом	
аналізу ієрархій	17
1.2.3 Модель вирішення проблеми	21
1.3 Змістовна та формальна постановка задачі	21
1.3.1 Змістовна постановка задачі	21
1.3.2 Формальна постановка задачі	22
1.4 Постановка задач дослідження	22
2 Вибір та обґрунтування методу розв’язання	24
2.1 Загальні характеристики криптовалют	24
2.1.1 Основні відомості	24
2.1.2 Принцип роботи криптовалют	26
2.1.3 Історія виникнення криптовалют	28
2.2 Вплив різних факторів на ринок криптовалют	29
2.2.1 Політичний вплив	29
2.2.2 Економічний вплив	30
2.2.3 Інформаційний вплив	31
2.3 Рекурентні нейронні мережі	32

	7
2.3.1 Нейронна мережа типу LSTM	34
2.3.2 Нейронна мережа типу GRU	38
Висновки за розділом 2	40
3 Програмна реалізація	41
3.1 Вибір мови програмування та середовища для реалізації програми	41
3.2 Алгоритм розв’язання задачі прогнозування цін на ринку криптовалют методами машинного навчання	42
3.3 Опис програми	43
Висновки за розділом 3	45
4 Результати обчислювального експерименту та їх аналіз	46
Висновки за розділом 4	54
Висновки	55
Перелік джерел посилання	56
Додаток А Лістинг програми	58

ВСТУП

Актуальність теми. Останніми роками криптовалюти стали однією із найважливіших та швидкозростаючих галузей фінансової індустрії, яка привертає до себе увагу як з боку інвесторів, так і з боку дослідників. Зростання ринку криптовалют та його вплив на глобальну економіку роблять цю тему дуже актуальною для дослідження. Одним з ключових викликів, що стоять перед учасниками ринку, є нестабільність цін на криптовалютному ринку. Вирішення цього питання може мати стратегічне значення для інвесторів та учасників ринку. Макроекономічні та мікроекономічні фактори і висока волатильність, які можуть впливати на ринок, роблять його складним для аналізу та прогнозування, а з іншого боку, висока волатильність дає змогу для інвесторів використовувати це для збільшення прибутку на деякому проміжку часу. Для того, щоб робити правильні висновки відносно майбутньої ціни на актив, є гостра необхідність у використанні аналітичних середовищ прийняття рішень. Наявність неоднозначності та швидкі зміни умов ринку криптовалют вимагають надійних методів прогнозування, які можуть адаптуватися до змінних умов.

Мета і завдання кваліфікаційної роботи. Метою кваліфікаційної роботи є прогнозування цін на ринку криптовалют за допомогою методів машинного навчання. Для досягнення поставленої мети необхідно виконати наступні завдання:

- провести огляд і аналіз сучасного стану задачі «Прогнозування цін на ринку криптовалют методами машинного навчання»;
- розв’язати задачу прогнозування цін на ринку криптовалют методами глибинного машинного навчання за допомогою рекурентних нейронних мереж;
- провести обчислювальні експерименти за історичними даними щодо цін, отриманими з відкритих джерел;
- провести аналіз отриманих результатів.

Об’єктом дослідження є процес змінення цін криптовалют.

Предметом дослідження є дослідження можливості прогнозування цін на ринку криптовалют.

Методи дослідження. У кваліфікаційній роботі використовуються методи машинного навчання, зокрема, методи рекурентних нейронних мереж.

Публікації. Результати, отримані у кваліфікаційній роботі, було представлено на 27-му Міжнародному молодіжному форумі «Радіоелектроніка та молодь у XXI столітті» (м. Харків, 10-12 травня 2023 р.) [19].

1 СИСТЕМНИЙ АНАЛІЗ ПРЕДМЕТНОЇ ОБЛАСТІ ТА ПОСТАНОВКА ЗАДАЧ ДОСЛІДЖЕННЯ

1.1 Системний аналіз задачі прогнозування цін на ринку криптовалют методами машинного навчання

1.1.1 Вербальна модель системи

Під час проведення системного аналізу предметної області розглянуто складну систему «Прогнозування цін на ринку криптовалют методами машинного навчання».

Предметом аналізу цієї системи є «Проблема прогнозування цін на ринку криптовалют методами машинного навчання».

Точка зору: дослідник.

Ціль: на основі статистичних даних про коливання цін на ринку криптовалют, розробити нейронну мережу, що буде давати прогноз на ціни криптовалют у майбутньому з прийнятним рівнем точності.

Головним входом для досліджуваної в роботі системи є статистичні дані про коливання цін на ринку криптовалют та фактори, що впливають на зміну цін.

Із використанням досліджуваної системи дослідник на виході отримує результати щодо прогнозних значень цін на ринку криптовалют у майбутньому.

1.1.2 Морфологічний опис системи

Морфологічний опис системи включає опис зовнішнього середовища та модель типу «чорний ящик».

Всі ресурси, необхідні для функціонування системи, організація отримує із зовнішнього середовища, з якою знаходиться в стані постійного обміну.

Морфологічний опис зовнішнього середовища зображено на рисунку 1.1.

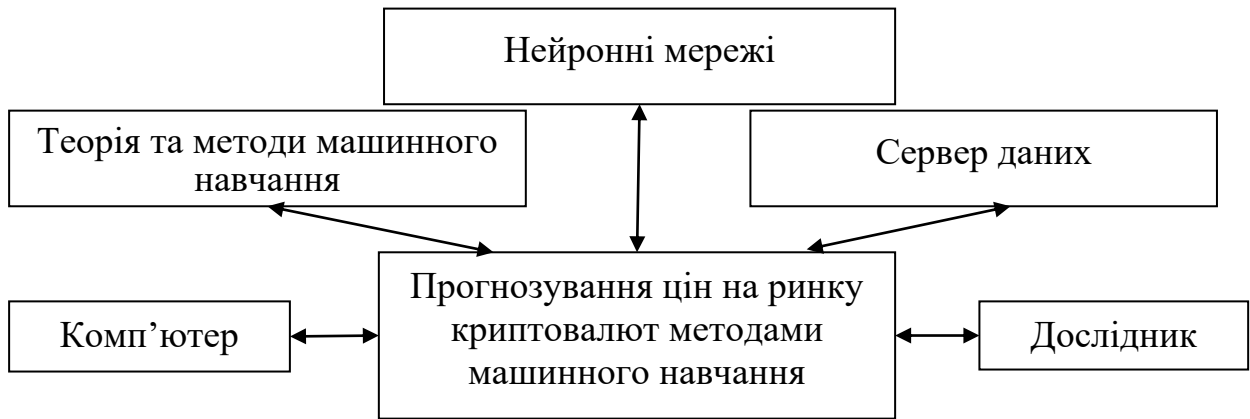


Рисунок 1.1 – Зовнішнє середовище системи

Зовнішнє середовище – це умови і фактори, які існують в навколишньому середовищі за межами системи. Вони можуть як впливати на функціонування системи, так і відчувати на собі її вплив.

Об'єкти зовнішнього середовища:

- а) за допомогою комп'ютера здійснюється процес обчислення;
- б) теорія та методи машинного навчання безпосередньо використовуються при організації комплексу, що дозволяє проводити обчислення;
- в) нейронні мережі дають змогу провести процес прогнозування на основі відомих даних;
- г) дослідник обирає методи для аналізу та за допомогою обчислювальних засобів робить обчислення для досліджуваного процесу;
- д) від серверу даних надходять статистичні дані, на основі яких проводять аналіз.

Для опису взаємодії системи з навколишнім середовищем використано модель типу «чорний ящик». На головний вхід системи подаються статистичні дані про коливання цін на ринку криптовалют та методи машинного навчання. На виході отримуємо прогноз на ціни криптовалют у майбутньому. Модель системи типу «чорна скринька» наведена на рисунку 1.2.

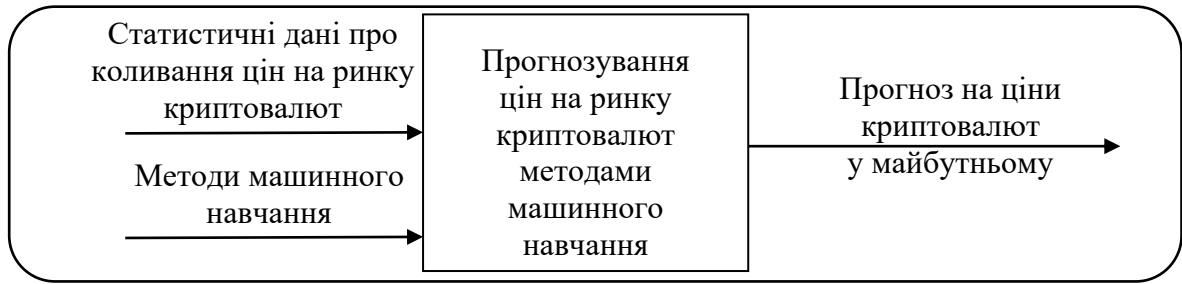


Рисунок 1.2 – Модель системи типу «чорна скринька»

1.1.3 Функціональна модель системи

В якості стандарту для моделювання різноманітних систем введена методологія функціонального моделювання IDEF0, яка виконує функції формалізації та опису бізнес-процесів [1].

Якщо розглядати процес в рамках методології IDEF0, він подається у вигляді набору елементів, які взаємодіють між собою, а також показують ресурси, що надходять до кожної роботи [1, 2]. В ієрархії діаграм IDEF0 перша діаграма відображає функціонування системи в цілому (рисунок 1.3). Такі діаграми називають контекстними. На вхід контекстних діаграм подаються вхідні дані, необхідні для виконання задачі системи; зверху – механізми керування (інструкції, теорії); знизу відображується те, що керує процесом, наприклад, дослідник; виходом контекстної діаграми є її цільовий продукт.

Після опису системи в цілому проводиться розбиття її на великі фрагменти. Цей процес має назву функціональної декомпозиції, а діаграми, які описують кожен фрагмент і взаємодію цих фрагментів, називаються діаграмами декомпозиції.

Після опису контексту будуються наступні декомпозиції в ієрархії для опису системи. Кожна діаграма більш детально описує предмет аналізу до моменту досягнення необхідного рівня деталізації опису [1, 2].

Застосування методології моделювання IDEF0 для нових систем ставить на меті визначення функцій та вимог для подальшої розробки системи,

що відповідає вимогам та реалізує виділені функції. Якщо IDEF0 використовується на існуючій системі, то є можливість аналізувати та описувати функції, які виконує система. Крім того, можна візуалізувати механізми, за допомогою яких ці функції виконуються.

В кваліфікаційній роботі розглядається функціонування системи «Прогнозування цін на ринку криптовалют методами машинного навчання». На вхід системи подаються статистичні дані про коливання цін на ринку криптовалют; засобами керування є теорія методів машинного навчання та алгоритми машинного навчання для прогнозування; до механізмів відносяться дослідник та прикладне програмне забезпечення; а на виході ми отримуємо прогноз на ціни криптовалют у майбутньому.

На рисунку 1.3. наведена контекстна діаграма розглянутої задачі. На її основі можна робити висновки щодо внутрішньої будови розглянутої системи. Основну проблему можна розбити на чотири великі задачі. Результат декомпозиції контекстної діаграми наведена на рисунку 1.4. Подальша декомпозиція наведена на рисунку 1.5.

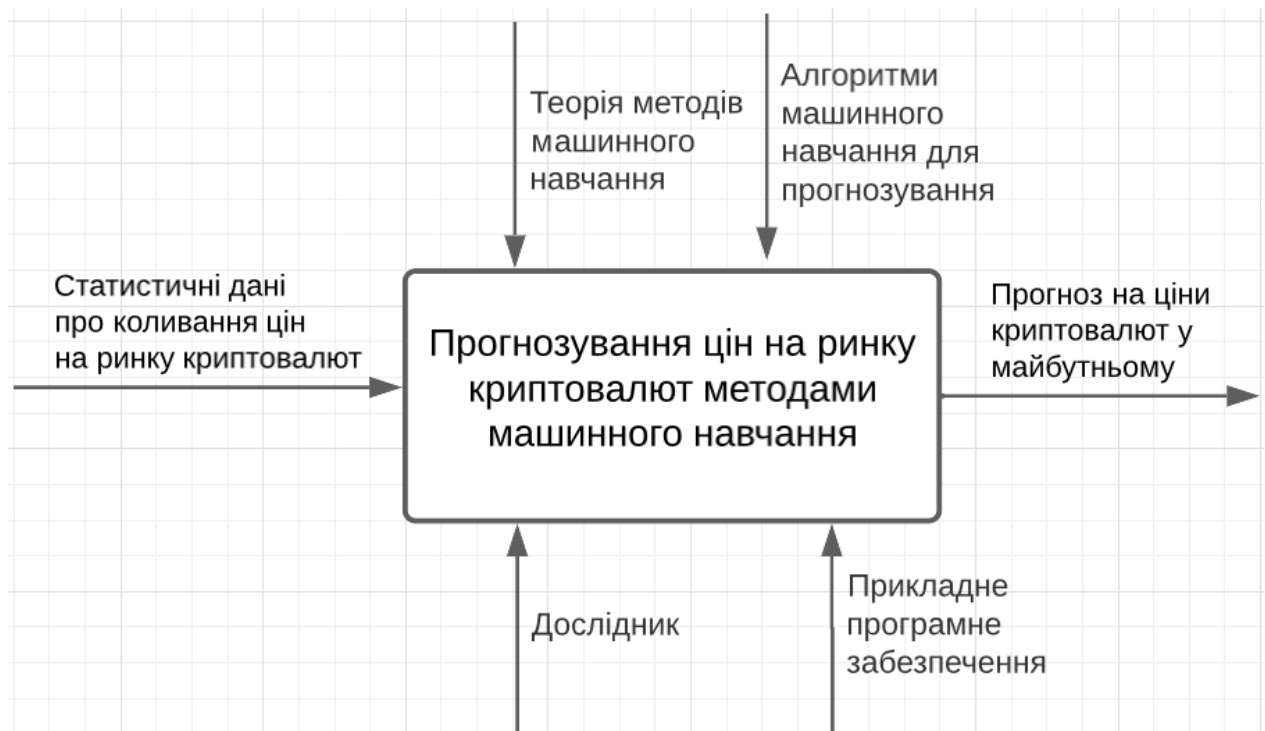


Рисунок 1.3 – Контекстна діаграма системи

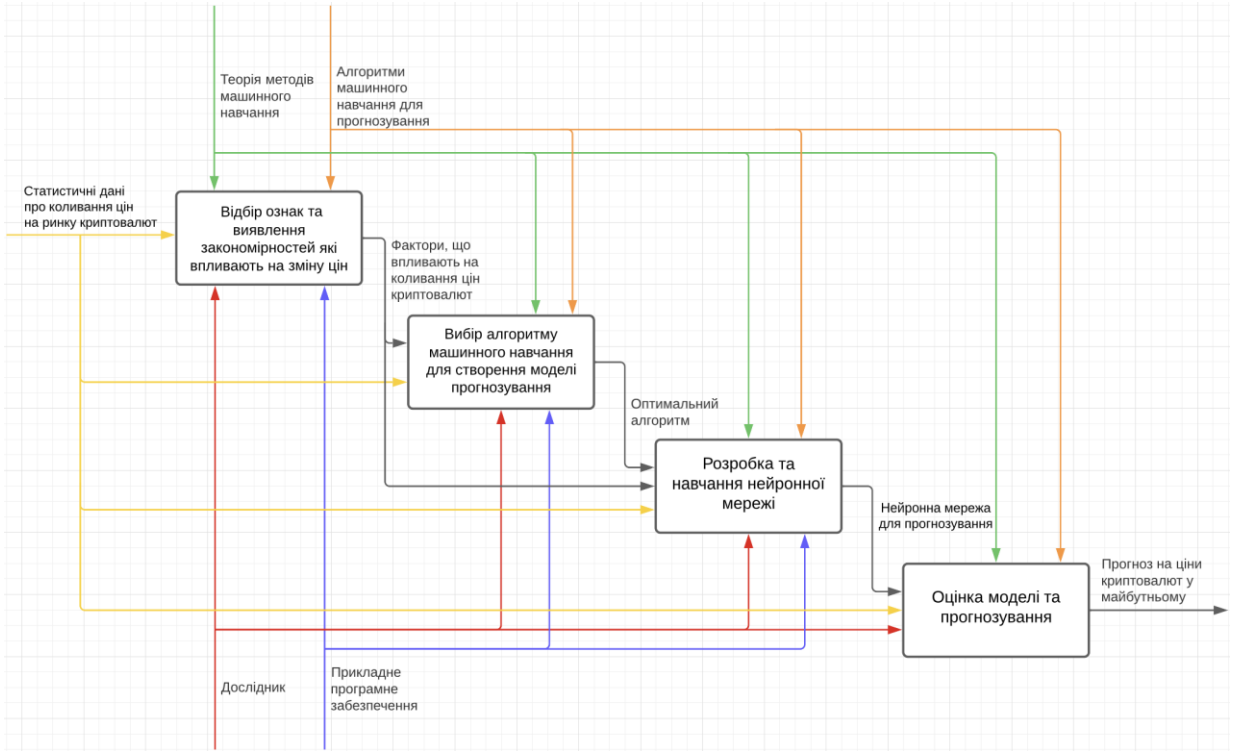


Рисунок 1.4 – Діаграма декомпозиції. Рівень А0

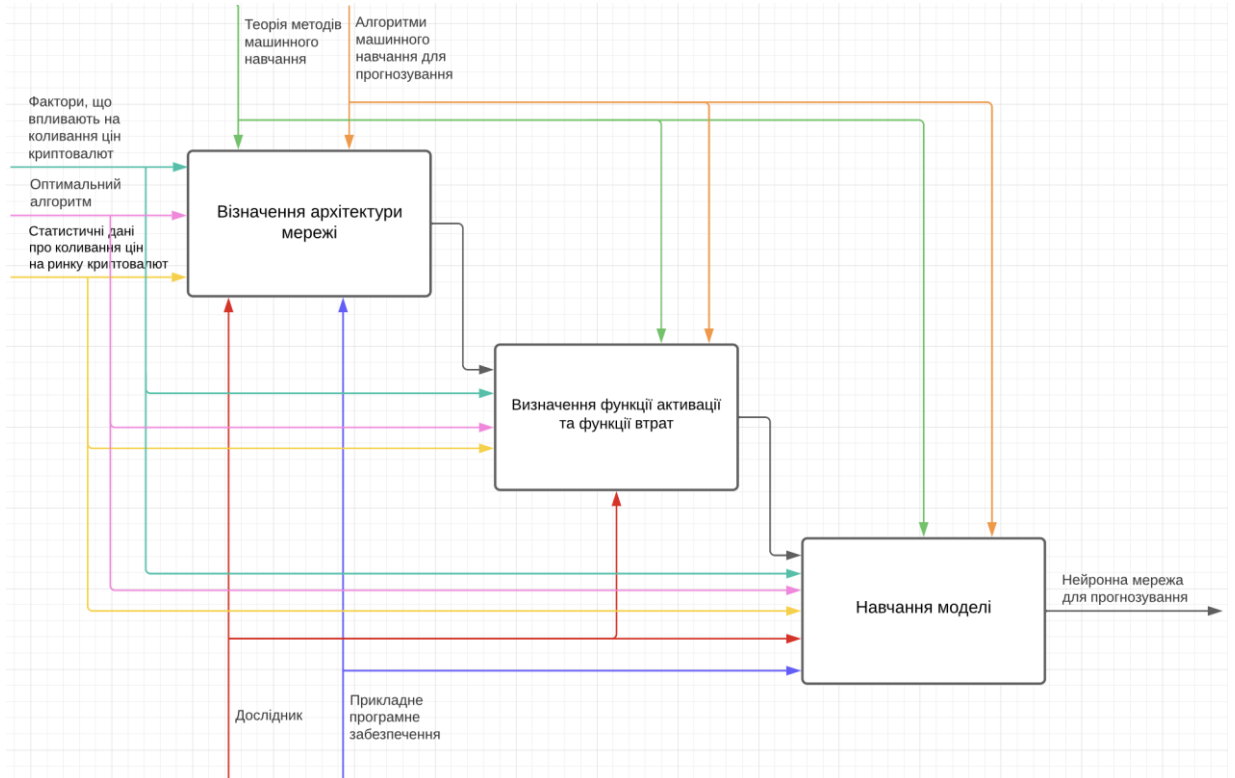


Рисунок 1.5 – Діаграма декомпозиції. Рівень А3

1.2 Аналіз сценаріїв вирішення задачі прогнозування цін на ринку криптовалют методами машинного навчання

1.2.1 Модель аналізу проблеми

Для структуризації проблеми в роботі будемо використовувати метод аналізу ієрархії (MAI), який ґрунтується на ієрархічному підході. Метод аналітичної ієрархії розкладає завдання експертної оцінки на більш прості складові частини. Ця декомпозиція визначає відносну значущість альтернатив в рамках ієрархічної системи критеріїв, і ця значущість відображується чисельно у вигляді векторів пріоритетів [2].

Значення отриманих векторів пріоритетів представляють оцінки на шкалі відношень та відповідають так названим жорстким оцінкам. Метод аналітичної ієрархії має різні модифікації, які враховують взаємозв'язки між альтернативами та критеріями, а також спосіб порівняння альтернатив.

За допомогою методу аналізу ієрархій необхідно дослідити задачу та вибрати оптимальний метод, яким буде розв'язана задача вибору алгоритму машинного навчання для створення моделі прогнозування.

Проблема буде вирішуватися за наступними критеріями:

- критерій 1: точність прогнозу;
- критерій 2: складність реалізації;
- критерій 3: середня швидкість обчислень;
- критерій 4: горизонт прогнозування.

Обирати будемо з множини альтернатив:

- альтернатива 1: лінійна регресія;
- альтернатива 2: рекурентні нейронні мережі (RNN);
- альтернатива 3: нейронні мережі прямого поширення сигналу.

Метод аналізу ієрархії полягає в структуризації задач прийняття рішень шляхом побудови багаторівневої ієрархії, що об'єднує усі компоненти задачі. Це, в свою чергу, дає змогу отримати чисельні оцінки інтенсивності

взаємовпливу елементів ієрархії, на основі яких оцінюються ступені переваги альтернатив відносно головної мети.

Використовуючи метод попарних порівнянь, побудуємо ієрархічну структуру (рисунок 1.6).

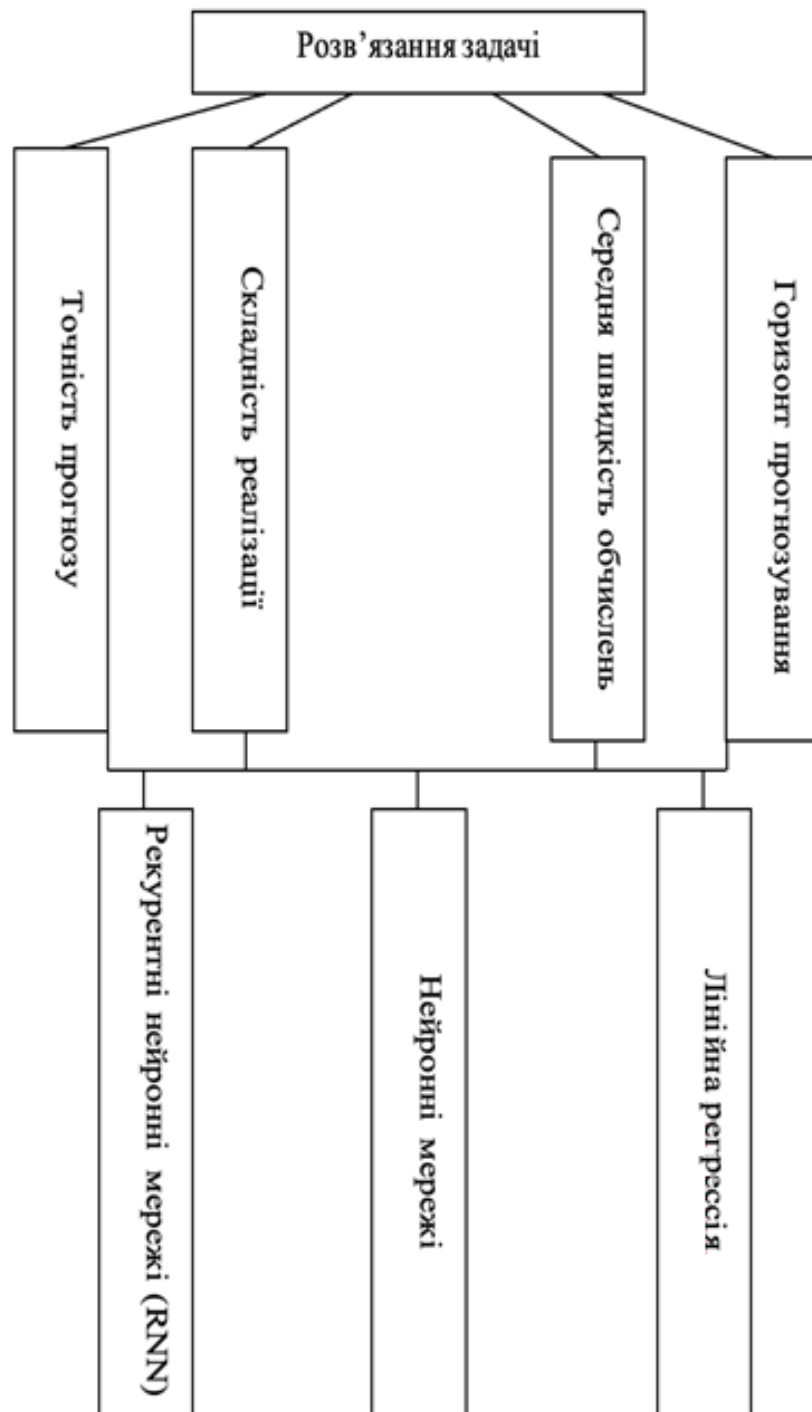


Рисунок 1.6 – Ієрархічна модель процесу аналізу розв'язання задачі

1.2.2 Оцінювання вектора пріоритетів незадоволеностей методом аналізу ієрархій

На наступному кроці методу аналізу ієрархій необхідно побудувати матриці попарних порівнянь моделі та критеріїв системи. За методом, запропонованим Т. Сааті, була сформована матриця порівнянь критеріїв, і її результати подані у таблиці 1.1 [3].

Таблиця 1.1 – Матриця попарних порівнянь

	K1	K2	K3	K4	Середнє геометричне за рядками	Вектор пріоритетів
K1	1	4	7	8	$x_1 = \sqrt[4]{1 \cdot 4 \cdot 7 \cdot 8} = 3,86867$	$p_1^K = \frac{x_1}{\Sigma} = 0,582$
K2	$\frac{1}{4}$	1	8	9	$x_2 = \sqrt[4]{\frac{1}{4} \cdot 1 \cdot 8 \cdot 9} = 2,05977$	$p_2^K = \frac{x_2}{\Sigma} = 0,310$
K3	$\frac{1}{7}$	$\frac{1}{8}$	1	$\frac{1}{2}$	$x_3 = \sqrt[4]{\frac{1}{7} \cdot \frac{1}{8} \cdot 1 \cdot \frac{1}{2}} = 0,30739$	$p_3^K = \frac{x_3}{\Sigma} = 0,046$
K4	$\frac{1}{8}$	$\frac{1}{9}$	2	1	$x_4 = \sqrt[4]{\frac{1}{8} \cdot \frac{1}{9} \cdot 2 \cdot 1} = 0,40824$	$p_4^K = \frac{x_4}{\Sigma} = 0,061$
					$\Sigma = x_1 + x_2 + x_3 + x_4 = 6,64407$	

$$\text{Індекс узгодженості (ІУ)} = \frac{4,463 - 4}{4 - 1} = 0,154.$$

$$\text{Випадкова узгодженість} = 0,13.$$

$$\text{Відносна узгодженість (ВУ)} = \frac{0,154}{0,19} = 0,15 = 15\%.$$

Для прийняття рішення щодо використання методу необхідно провести порівняльний аналіз альтернатив. У цій модифікації методу аналізу ієрархій (МАІ) альтернативи розглядаються як ієрархія з однаковою кількістю та функціональним складом. Щоб встановити відносну важливість елементів у цій

ієрархії, використовується шкала відношень Сааті. Результати порівняння альтернатив за окремими критеріями, коли експерт оцінює кожну альтернативу відповідно до їх переваг, представлені в таблицях 1.2 – 1.5.

Значення випадкової узгодженості для цих матриць дорівнює 0,58.

Таблиця 1.2 – Порівняння альтернатив за першим критерієм

K1	A1	A2	A3
A1	1	$\frac{1}{6}$	$\frac{1}{7}$
A2	6	1	3
A3	7	$\frac{1}{3}$	1

Таблиця 1.3 – Порівняння альтернатив за другим критерієм

K2	A1	A2	A3
A1	1	$\frac{1}{8}$	$\frac{1}{6}$
A2	8	1	5
A3	6	$\frac{1}{5}$	1

Таблиця 1.4 – Порівняння альтернатив за третім критерієм

K3	A1	A2	A3
A1	1	4	$\frac{1}{6}$
A2	$\frac{1}{4}$	1	$\frac{1}{7}$
A3	6	7	1

Таблиця 1.5 – Порівняння альтернатив за четвертим критерієм

K4	A1	A2	A3
A1	1	$\frac{1}{3}$	4
A2	3	1	5
A3	$\frac{1}{4}$	$\frac{1}{5}$	1

Для кожної наведеної вище матриці розраховуємо вектор пріоритетів:

$$\vec{p}_1^A = \begin{pmatrix} 0,066 \\ 0,619 \\ 0,315 \end{pmatrix},$$

$$\vec{p}_2^A = \begin{pmatrix} 0,057 \\ 0,719 \\ 0,224 \end{pmatrix},$$

$$\vec{p}_3^A = \begin{pmatrix} 0,187 \\ 0,070 \\ 0,743 \end{pmatrix},$$

$$\vec{p}_4^A = \begin{pmatrix} 0,280 \\ 0,627 \\ 0,094 \end{pmatrix}.$$

Оскільки подані матриці парних порівнянь альтернатив мають третій порядок, то $RI^A = 0,58$. Індеси узгодженості і відношення узгодженості для цих матриць за кожним критерієм співпадають:

$$CI_{K1}^A = 0,079, \quad CI_{K2}^A = 0,0945, \quad CI_{K3}^A = 0,086, \quad CI_{K4}^A = 0,043,$$

$$CR_{K1}^A = 0,136, \quad CR_{K2}^A = 0,162, \quad CR_{K3}^A = 0,147, \quad CR_{K4}^A = 0,074.$$

Усі значення відношень узгодженості близькі до 0,1, що свідчить про високу узгодженість оцінок, наданих експертом.

Далі, для обчислення вектора глобальних пріоритетів альтернатив формуємо матрицю, складену з векторів локальних пріоритетів альтернатив за кожним критерієм

$$P^A = \begin{pmatrix} 0,066 & 0,057 & 0,187 & 0,280 \\ 0,619 & 0,719 & 0,070 & 0,627 \\ 0,315 & 0,224 & 0,743 & 0,094 \end{pmatrix},$$

та обчислюємо вектор глобальних пріоритетів за формулою $\vec{p} = P^A \vec{p}^K$.

Отже, вектор глобальних пріоритетів дорівнює:

$$\vec{p} = \begin{pmatrix} 0,066 & 0,057 & 0,187 & 0,280 \\ 0,619 & 0,719 & 0,070 & 0,627 \\ 0,315 & 0,224 & 0,743 & 0,094 \end{pmatrix} \cdot \begin{pmatrix} 0,582 \\ 0,310 \\ 0,046 \\ 0,061 \end{pmatrix} = \begin{pmatrix} 0,082 \\ 0,625 \\ 0,292 \end{pmatrix}.$$

Розрахуємо індекс узгодженості і відношення узгодженості для всієї наведеної ієрархії:

$$CI = CI^K + \left(\vec{p}^K, \vec{CI}^A \right) = 0,154 + 0,582 \cdot 0,079 + 0,310 \cdot 0,0945 +$$

$$+ 0,046 \cdot 0,086 + 0,061 \cdot 0,043 = 0,23585,$$

$$RI = RI^K + RI^A = 0,90 + 0,58 = 1,48,$$

$$CR = \frac{CI}{RI} = 0,159,$$

що також можна вважати гарним показником узгодженості.

1.2.3 Модель вирішення проблеми

У таблиці 1.6 наведені результати, з яких можна зробити висновок, що максимальна компонента вектору глобальних пріоритетів відповідає другій альтернативі, тобто методу рекурентних нейронних мереж.

Таблиця 1.6 – Кінцеві дані

Критерій /Альтернатива	K1	K2	K3	K4	Узагальнені пріоритети
A1	0,066	0,057	0,187	0,280	0,082
A2	0,619	0,719	0,070	0,627	0,625
A3	0,315	0,224	0,743	0,094	0,292

1.3 Змістовна та формальна постановка задачі

1.3.1 Змістовна постановка задачі

У даній роботі розглядається задача прогнозування цін на криптовалюти за допомогою методів машинного навчання. Задача полягає в прогнозуванні ціни криптовалют на ринку на основі історичних даних, що відображують коливання цін на ринку у минулому, а також можливо інших індикаторів, таких як обсяг торгів, RSI, MACD тощо. Прогнозування цін криптовалют є важливим завданням для трейдерів та інвесторів, які прагнуть приймати успішні рішення на ринку.

Метод рекурентних нейронних мереж (RNN) використовується для аналізу часових рядів і динамічних залежностей між даними в різні моменти часу. Тому модель RNN може навчатися на історичних даних та використовуватись для прогнозування цін криптовалют у майбутнє.

1.3.2 Формальна постановка задачі

Нехай:

X_t – вектор вхідних даних для моменту часу t (цінові дані, обсяги торгів, технічні індикатори);

Y_t – ціна криптовалюти для часу t ;

H_t – стан прихованого шару RNN в момент часу t ;

W – матриця ваг, яка використовується для прогнозування ціни;

b – вектор зміщення;

σ – функція активації.

Наша мета – створити RNN-модель, яка здатна передбачити ціну криптовалюти на момент деякого проміжку часу, на основі інформації з моменту часу t та попередніх даних.

Модель описана такою рекурентною формулою:

$$H_t = \sigma(WX_t + UH_{t-1} + b),$$

$$Y_{t+1} = VH_t + c,$$

де W, U, V – матриці ваг;

b, c – вектори зміщення.

1.4 Постановка задач дослідження

Метою цієї роботи є розробка моделі RNN, яка буде навчатися на основі історичних даних щодо динаміки цін на криптовалюту та чинників, які мали на це вплив, і використовувати їх для передбачення ціни криптовалюти у майбутні моменти часу. Модель може бути налаштована з урахуванням конкретних характеристик ринку криптовалют та використовувати різноманітні технічні та фундаментальні показники як вхідні дані.

Виходячи із постановки задачі сформуємо перелік задач, які необхідно виконати під час дослідження:

- провести огляд і аналіз сучасного стану задачі «Прогнозування цін на ринку криптовалют методами машинного навчання»;
- розв’язати задачу прогнозування цін на ринку криптовалют методами глибинного машинного навчання за допомогою рекурентних нейронних мереж;
- провести обчислювальні експерименти за історичними даними щодо цін, отриманими з відкритих джерел;
- провести аналіз отриманих результатів.

2 ВИБІР ТА ОБҐРУНТУВАННЯ МЕТОДУ РОЗВ'ЯЗАННЯ

2.1 Загальні характеристики криптовалют

2.1.1 Основні відомості

Криптовалюта – це децентралізована цифрова або віртуальна валюта, що базується на технології блокчейн та призначена для роботи, як засіб обміну валюти через комп'ютерну мережу. На відміну від офіційних валют, криптовалюти є децентралізованими, що передбачає за собою відсутність централізованого контролю та впливу на роботу системи. Принцип роботи криптовалют базується на технології блокчейн, яка є розподіленою базою даних, застосовуваною для реєстрації та підтвердження транзакцій.

У загальному вигляді криптовалютами можна вважати цифрові дані, належність яких може бути підтверджена цифровим підписом, облік яких можливий за допомогою криптографічних методів, а існування доведено за допомогою незалежної системи. Криптоактиви можуть мати ознаки та слугувати цифровим вираженням дуже широкого спектру різноманітних матеріальних та нематеріальних активів та прав на них.

Технічні, технологічні та організаційні аспекти використання криптовалют вказують на те, що вони майже тотожні з електронним видом грошей. Є значний потенціал у схожості грошей електронної форми із криптовалютами, як міри вартості, засобу обігу платежу та засобу накопичення, але висока волатильність стає цьому на заваді. Зростання поширеності та інтеграції криптовалют у традиційні платіжні системи та зниження транзакційних витрат дає можливість мати криптовалюті ці функції. Однак, повне ототожнення криптовалют з електронними грошима є хибним. Схожість проявляється лише в тому, що криптовалюти як і електронні гроші не потребують ідентифікації власника та обертаються поза межами банківської системи в електронному вигляді [5].

Основні відмінності криптовалюти від електронних грошей:

– доступність до криптовалюти обмежується лише доступом до мережі Інтернет, в той час, як доступність до електронних грошей пов'язана із доступом до мобільних девайсів і агентської мережі;

– більшість країн поки що не мають правового регулювання обігу криптовалюти, електронні гроші регулюються вповноваженим органом (найчастіше – центральним банком);

– емітентом криптовалютит є блокчейн та майнери, а емітентом електронних грошей виступає юридично уповноважена особа-емітент;

– виробляється криптовалюта в процесі майнінгу, електронні ж гроші випускаються за наказом уповноваженого органу в кількості, яка дорівнює масі фіатних грошей;

– при використанні електронних грошей для запобігання фінансових ризиків вимагається процедура ідентифікації клієнта, а при використанні криптовалюти зберігається анонімність;

– цінність криптовалюти визначається попитом та пропозицією, довірою користувачів до системи, а цінність електронних грошей дорівнює цінності та кількості фіатної валюти [5].

Відповідно до опису та основних характеристик, криптовалюта повинна відповідати наступним умовам:

– система являється децентралізованою, тобто не потребує єдиного органу керування;

– система зберігає огляд одиниць криптовалюти та їх власності;

– здатність створення нових одиниць криптовалюти визначена системою до виходу проекту на ринок; лімітований випуск цифрових валют забезпечує контроль інфляції, адже загальна кількість криптомонет відома заздалегідь;

– підтвердити право власності на одиницю криптовалюти можна лише криптографічним шляхом;

– записи про всі криптотранзакції зберігаються у прозорому та захищеному від злому реєстрі: як тільки транзакція потрапила до блокчейну, її

не можна видалити або відкоригувати, доступним залишається лише перегляд, що може зробити будь-хто;

– якщо одночасно введено дві різні інструкції щодо зміни власника одних і тих же криптографічних одиниць, система виконує щонайбільше одну з них.

Біткойн, який вважається найпершим на ринку і найдорожчою віртуальною валютою світу, був створений з метою зниження споживчих витрат шляхом витіснення банківських установ з ланцюжка однорангових фінансових угод. Поява біткоіна розв'язала проблему подвійних витрат, покладених на споживачів фінансових послуг, оскільки індивідууми мають відкритий доступ до всіх документальних записів стосовно їхніх віртуальних транзакцій [6]. Також сильним поштовхом стала глобальна фінансова криза 2008 року, що була викликана іпотечною мильною бульбашкою у США. Після цієї кризи люди усвідомили, що центральні банки навіть найбільших економік світу є надто вразливими, і що потрібно переосмислити усю банківську структуру загалом.

2.1.2 Принцип роботи криптовалют

Більшість криптовалют базуються на технології блокчейн, де покупці та продавці можуть взаємодіяти безпосередньо. Технологія блокчейн є неперервно обновлюваною базою даних, що реєструє всі транзакції. Кожен учасник отримує унікальний криптографічний ключ для підпису та контролю доступу до своїх цифрових активів. Підтвердження транзакцій відбувається під час одночасного доступу покупця та продавця, забезпечуючи стійкість до змін у ланцюгу блоків. Перш ніж транзакції додаються до блоку, вони проходять підтвердження в мережі вузлів, забезпечуючи високий рівень безпеки та запобігаючи зовнішнім втручанням. Кожен новий блок утворює послідовне з'єднання з попереднім блоком у ланцюжку. Важливо зауважити, що блокчейн не дозволяє змінювати вже існуючі блоки; замість цього, можна лише додавати нові. Ця обмеженість виникає з того, що будь-яка зміна в одному блоку

призводить до зміни хеш-коду, який використовується в наступних блоках. Такий підхід забезпечує стійкість до видалення або модифікацій, оскільки будь-яка зміна в ланцюжку блоків негайно стає помітною.

Розглянемо основні операції, які можна здійснювати з криптовалютами:

Зберігання: для безпечного зберігання криптовалют і управління криптографічними ключами, криптовалюти зберігають в програмних або апаратних електронних гаманцях, або на криптобіржах.

Обмін: криптовалюти можна обмінювати на інші криптовалюти або традиційні фіатні валюти на криптобіржах. Це дозволяє користувачам купувати та продавати криптовалюти за поточними курсами.

Транзакції: криптовалюти можуть використовуватися для здійснення платежів та переказів коштів. Користувачі можуть ініціювати транзакції, вказуючи отримувача та кількість криптовалюти, яку вони бажають передати. Транзакції підтверджуються та обробляються мережею вузлів.

Майнінг: під майнінгом розуміють процес створення нових блоків та підтвердження транзакцій у блокчейні криптовалюти. В процесі майнінгу виконується процес вирішення складних обчислювальних задач. Наприклад, у Bitcoin застосовується механізм доказу роботи (Proof of Work). Майнери, учасники мережі, конкурують між собою, намагаючись розв'язати математичну задачу. Ця задача вимагає великих обчислювальних потужностей та багатьох спроб, але легко перевіряється іншими учасниками мережі. Майнер, який першим успішно розв'язує задачу, отримує можливість створити новий блок у блокчейні. Після створення блока в нього включаються підтвержені транзакції з мережі. За свою роботу майнер отримує винагороду в криптовалюті, такій як Bitcoin. Крім того, майнери можуть отримувати комісійні за транзакції, які вони включають у блок. Майнінг допомагає забезпечити безпеку мережі, оскільки вирішення завдання потребує великих обчислювальних потужностей, ускладнюючи спроби зломисників змінювати попередні блоки. Також майнінг регулює випуск нових криптовалют у обіг, оскільки винагороди за майнінг поступово зменшуються з часом.

2.1.3 Історія виникнення криптовалют

Початкові концепції створення криптовалюти виникли завдяки внеску американського криптографа Девіда Чаума, який у 1983 році висунув ідею криптографічних електронних грошей під назвою «ecash». Не менш важливою роллю в розвитку цього напрямку відзначилися імена Вея Дая, Ніка Сабо та Адама Бека. Вони були серед піонерів, що зробили значний внесок у створення концепцій блокчейну, які в подальшому стали фундаментом для функціонування сучасних криптовалют.

У сучасності розвиток криптовалют почався із запуску першої та найвідомішої криптовалюти – Bitcoin. Засновником Bitcoin вважається особа або група осіб під псевдонімом Сатосі Накамото (Satoshi Nakamoto), які в 2008 році оприлюднили білий папір під назвою «Bitcoin: A Peer-to-Peer Electronic Cash System». У цьому документі розглядалася методика вирішення проблеми подвійного витрачання цифрових активів без посередництва чи участі третіх сторін, пропонуючи ідею децентралізованої електронної валюти.

В січні 2009 року було запущено першу мережу Bitcoin, а Сатосі Накамото видобув перший блок Bitcoin, відомий як «Genesis Block». Також в цьому році була вирішена перша транзакція між Сатосі Накамото та розробником Bitcoin Халом Фінні.

Bitcoin став першою і, на той час, єдиною криптовалютою. Однак у 2011 році з'явилися перші альтернативні криптовалюти, такі як Namecoin та Litecoin, які вдосконалювали різні аспекти протоколу Bitcoin.

В 2013 році вартість Bitcoin стрімко зросла, досягнувши значення понад 1000 доларів за один Bitcoin. Це викликало значний інтерес громадськості та залучення уваги від ЗМІ та фінансових інституцій.

У 2017 році ринок криптовалют пережив велике зростання, з'явилося багато нових криптовалют та проєктів на базі блокчейн. Зростання популярності призвело до запуску багатьох Initial Coin Offerings (ICO), що надали новим проєктам фінансову підтримку.

Після рекордного 2017 року ринок зазнав корекції і багато криптовалют втратили значну частину своєї вартості. А вже у 2018 році почалась активна дискусія про регулювання криптовалют та ICO.

У 2020-2021 роках відбулось збільшення інтересу від інституційних інвесторів. Також було розгорнуто нові технології та покращено версії попередніх криптовалют, наприклад, Ethereum 2.0.

Наступні роки видалися важкими для ринку криптовалют. Багато криптоактивів зазнали судових звинувачень збоку комісії цінних паперів США.

2.2 Вплив різних факторів на ринок криптовалют

2.2.1 Політичний вплив

Політичний вплив становить найбільшу силу на криптовалютний ринок, оскільки рішення країн щодо обмежень чи заборони використання криптовалют впливають на їхню прийнятність і розповсюдження. Коли певна країна вводить заборону на використання криптовалюти, це призводить до припинення прийому її компаніями та обмеження доступу для населення через законодавчі обмеження, чи навіть до кримінальної відповідальності за її використання.

На сьогодні в світі існує вісім країн, де криптовалюта повністю заборонена. Першою країною, яка введе такий заборонний захід, була Болівія у 2014 році. Китай є найбільшою країною, де заборонено використання криптовалюти. Раніше Китай відзначався своєю лідерською позицією у галузі майнінгу криптовалют, і його рішення призвело до спаду цін на біткоїн на 5%. Серед інших країн, де криптовалюта заборонена, також зазначаються авторитарні або недемократичні режими, такі як Непал, Афганістан, Бангладеш, Марокко, Алжир і Єгипет. Крім того, багато країн вводять часткові обмеження на використання криптовалют [7].

Наявність чи відсутність заборони – це лише один з аспектів політичного впливу на криптовалютний ринок. Політичні конфлікти між країнами також

можуть визначати використання та сприйняття криптовалют. Наприклад, введення економічних санкцій однією країною проти іншої може збільшити інтерес до криптовалют для уникнення санкцій та здійснення міжнародних транзакцій.

Деякі уряди розглядають можливість створення своїх централізованих цифрових валют (ЦЦВ), які будуть випускатися та регулюватися центральними банками. Впровадження ЦЦВ може вплинути на криптовалютний сектор, пропонуючи альтернативу децентралізованим криптовалютам і потенційно впливаючи на їх прийняття. Це може бути однією з причин заборони криптовалюти в Китаї в 2021 році, оскільки саме тоді було випробувано цифровий юань [8, 9].

Крім того, висловлювання та дії політиків можуть формувати громадську думку та ставлення до криптовалют. Позитивні або негативні коментарі можуть впливати на впевненість на ринку та настрої інвесторів, спричиняючи коливання цін та зміни на ринку.

2.2.2 Економічний вплив

Ціни на криптовалюту можуть миттєво змінюватися під впливом різноманітних факторів, таких як інвестиційний настрій, технологічний прогрес, зміни в регулюванні та макроекономічні умови. Вартість криптовалют також може визначатися пропозицією, яка включає винагороди за майнінг та механізми випуску токенів.

Криптовалютні ринки виявляють велику чутливість до настроїв, які часто формуються внаслідок спекулятивної торгівлі. Позитивні новини, укладення партнерських угод або досягнення технологічного прогресу можуть призводити до зростання цін, оскільки інвестори очікують подальшого зростання вартості. Зворотно, негативні новини, проблеми з безпекою або ринкове маніпулювання можуть спричинити продажі та падіння цін.

Одним з важливих факторів, що суттєво впливає на ринок криптовалют, була економічна криза 2022 року. Загальна капіталізація всіх криптовалют відчутно зменшилася з 2,5 трлн доларів США наприкінці 2021 року до менше 1 трлн у червні 2022 року. У період нестабільності люди починають переводити свої активи з менш стабільних криптовалют в більш стабільні фіатні валюти, такі як долар США, євро, швейцарський франк, золото та срібло [10].

Також зростає усвідомленість екологічних проблем, пов'язаних з енергоспоживанням криптовалют, особливо в процесі майнінгу біткойнів. Екологічний вплив криптовалют може впливати на інвестиційний настрій, регулювання та загальну динаміку ринку.

2.2.3 Інформаційний вплив

Інформаційний вплив відіграє ключову роль на криптовалютному ринку. Актуальні новини, події та зміни в регулюванні мають велике значення для учасників цього динамічного ринку. Інформація впливає на настрої, ціни та поведінку інвесторів, надаючи їм можливість відслідковувати та реагувати на зміни.

Величезний вплив мають повідомлення відомих директорів, експертів та авторитетних видань. Це особливо відзначається у випадку з повідомленнями Ілона Маска у соціальних мережах, які часто викликають значні коливання вартості криптовалют. Наприклад, твіт Маска про Dogecoin у початкових місяцях 2021 року призвів до зростання його вартості більше ніж на 800%. Аналогічно, висловлювання Маска щодо занепокоєння стосовно високого енергоспоживання біткойна та припинення компанії Tesla приймати його як оплату, призвело до зниження ціни біткойна понад 10% [11, 12].

2.3 Рекурентні нейронні мережі

Ціни криптовалют дуже чутливі різноманітних факторів, що ускладнює роботу з ними. Тому є гостра необхідність у використанні сучасних інформаційних та обчислювальних технологій для обробки історичних даних, які відображують змінення цін з часом, та виконання на їх основі прогнозів щодо майбутніх значень цих цін. Одним з сучасних підходів у цьому напрямку є використання технологій глибокого навчання, зокрема, рекурентних нейронних мереж. Розглянемо їх більш детально.

Рекурентні нейронні мережі (Recurrent Neural Network, RNN) – клас моделей глибокого машинного навчання, заснований на використанні попередніх станів мережі для обчислення поточного. З'єднання між вузлами утворюють граф, орієнтований у часі. Це створює внутрішній стан мережі, що дозволяє їй проявляти динамічну поведінку в часі. На відміну від нейронних мереж прямого поширення, рекурентні нейронні мережі можуть використовувати свою внутрішню пам'ять для обробки довільних послідовностей входів. Такі мережі зручно застосовувати в тих випадках, коли вхідні дані завдання є нефіксованою послідовністю значень, як, наприклад, текстові дані, де текстовий фрагмент представлений нефіксованою кількістю речень, фраз і слів. Кожен символ у тексті, окремі слова, розділові знаки і навіть цілі фрази – все це може бути атомарним елементом вхідної послідовності [13].

За останні роки було запропоновано багато різних архітектурних рішень для рекурентних мереж від найпростіших до складних. Останнім часом найбільшого поширення набули мережа з довготривалою та короткочасною пам'яттю (LSTM) та керований рекурентний блок (GRU). На рисунку 2.1 зображена частина нейронної мережі: A отримує деякі дані X на вхід та подає на вихід деяке значення H . Циклічний зв'язок дозволяє передавати інформацію від поточного кроку мережі до наступного [14].

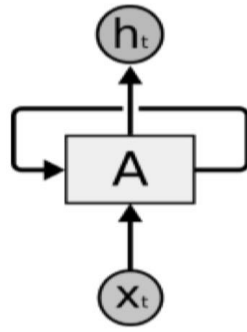


Рисунок 2.1 – Частина рекурентної нейронної мережі

Існує багато різновидів, рішень та конструктивних елементів рекурентних нейронних мереж. Проблема рекурентної мережі полягає в тому, що якщо враховувати кожен крок часу, то стає необхідним для кожного кроку часу створювати свій шар нейронів, що викликає серйозні обчислювальні складнощі. Крім того, багат шарові реалізації виявляються обчислювально нестійкими, тому що в них, як правило, зникають або стрімко зростають ваги.

Якщо обмежити розрахунок фіксованим часовим вікном, то отримані моделі не відображатимуть довгострокові тренди. Різні підходи намагаються вдосконалити модель історичної пам'яті та механізм запам'ятовування та забування.

Рекурентні нейронні мережі не так сильно відрізняються від звичайних нейронних мереж. Їх можна уявити собі, як безліч копій однієї й тієї ж мережі, причому кожна копія передає повідомлення наступній копії. На рисунку 2.2 показано розгорнутий цикл рекурентної мережі [14].

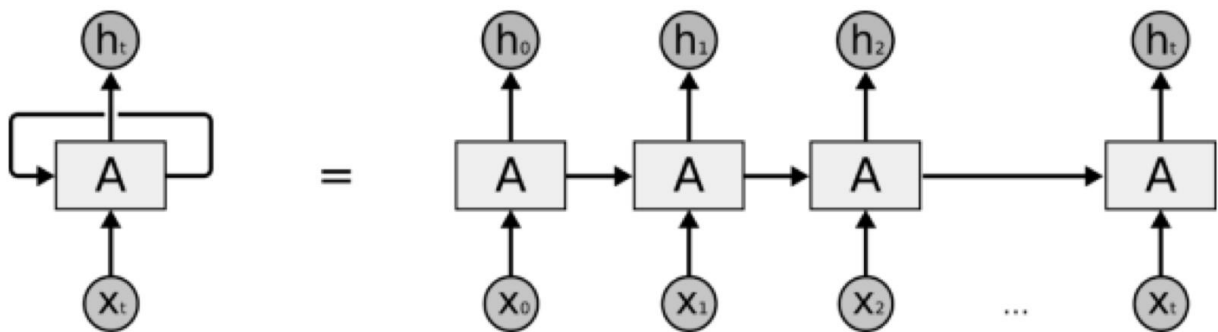


Рисунок 2.2 – Цикл рекурентної нейронної мережі

Така “ланцюгова” сутність показує, що рекурентні нейронні мережі за своєю природою тісно пов’язані з послідовностями та списками.

Базова архітектура, що була розроблена у 1980-х роках, будується з вузлів, кожен із яких з’єднаний з усіма іншими вузлами. У кожного нейрона поріг активації змінюється згодом і є дійсним числом. Всі нейрони мають змінні дійснозначні ваги. Вузли поділяються на вхідні, вихідні та приховані. В задачах навчання з учителем на кожному дискретному кроці часу на вхідні вузли подаються дані, а інші вузли завершують свою активацію, та вихідні сигнали готуються для передачі нейронам наступного рівня. Якщо, наприклад, мережа відповідає за розпізнавання мови, в результаті до вихідних вузлів надходять вже мітки (розпізнані слова).

У задачах навчання з підкріпленням (reinforcement learning), де немає вчителя, який забезпечує цільові сигнали для мережі, використовується функція придатності або функція оцінки (reward function). За цією функцією проводиться оцінка якості роботи мережі, при цьому значення на виході впливає на поведінку мережі на вході.

Зокрема, якщо мережа реалізує гру, на виході вимірюється кількість пунктів виграшу або оцінка позиції. Кожен ланцюжок обчислює помилку як сумарну девіацію за вихідними сигналами мережі. Якщо є набір навчальних прикладів, помилка обчислюється з урахуванням помилок кожного окремого прикладу.

2.3.1 Нейронна мережа типу LSTM

Мережі довготривалої пам’яті (Long Short Term Memory) або LSTM – особливий вид RNN, здатних до навчання довгостроковим залежностям. Вони показують дуже добрі результати при розв’язанні різноманітних проблем і в даний час широко застосовуються. LSTM спеціально спроектовані таким чином, щоб уникнути проблеми довгострокових залежностей. Запам’ятовувати інформацію на

тривалий період часу – це практично їхня поведінка за замовчуванням, а не щось таке, що вони лише намагаються зробити.

Така мережа, крім вхідних значень та попереднього стану, використовує також фільтри (gates), що визначають, яким чином інформація буде використовуватися для обчислення вихідних значень на поточному шарі y_t , так і значень прихованого шару на наступному кроці h_{t+1} .

У мережах LSTM вдалося уникнути проблеми зникнення або вибухового зростання градієнтів у процесі навчання шляхом зворотного поширення помилки. Мережа LSTM зазвичай керується за допомогою рекурентних вентилів, які називаються вентилями (gates) «забування». Помилки поширюються назад у часі через потенційно необмежену кількість віртуальних шарів.

Таким чином навчання в LSTM відбувається із збереженням пам'яті тисяч і навіть мільйонів часових інтервалів у минулому. Топології мереж типу LSTM можуть розроблятися відповідно до специфіки завдання. У мережі LSTM можуть враховуватися навіть великі затримки між значними подіями, і тим самим високочастотні та низькочастотні компоненти можуть змішуватись.

Усі рекурентні нейронні мережі мають форму ланцюга модулів, що повторюються (repeating module). У стандартній RNN (рисунок 2.3) ці модулі, що повторюються, матимуть дуже просту структуру, наприклад, всього один шар гіперболічного тангенсу (tanhtanh) [14].

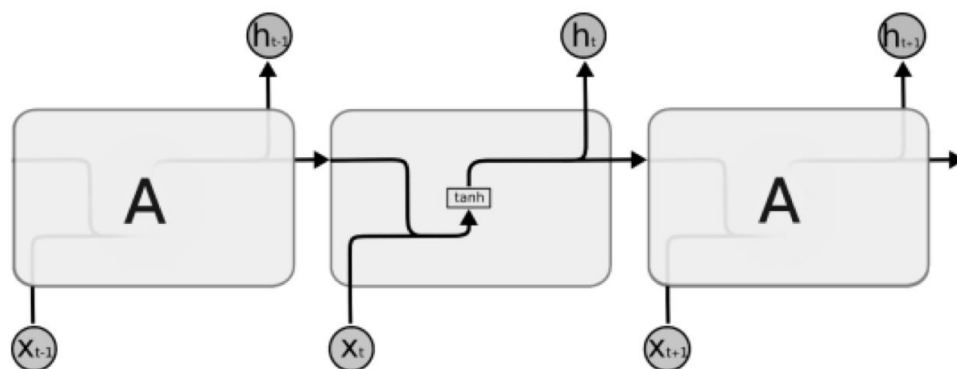


Рисунок 2.3 – Архітектура рекурентної нейронної мережі

LSTM теж мають таку ланцюгову структуру, але модуль, що повторюється, побудований інакше. Замість одного нейронного шару їх чотири, причому вони взаємодіють особливим чином (рисунок 2.4) [14].

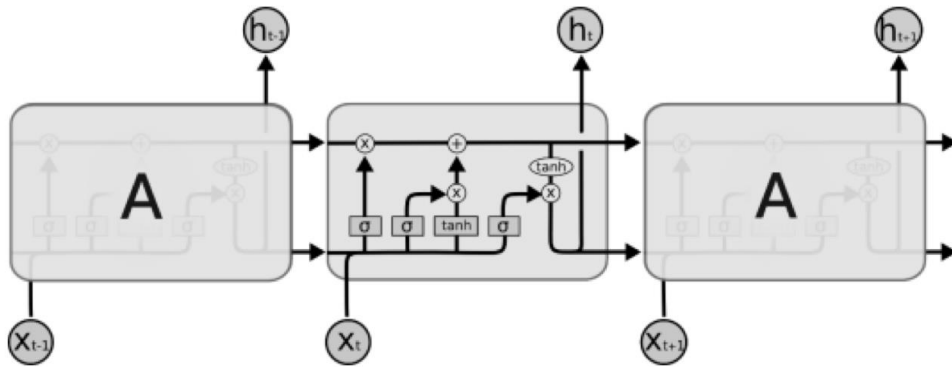


Рисунок 2.4 – Архітектура рекурентної нейронної мережі LSTM

Розглянемо докладніше структуру LSTM-шару. Центральним поняттям тут є запам'ятовуючий блок (memory cell), який, так само, як і стан мережі h , обчислюється на кожному кроці з використанням поточного вхідного значення x_t та значення блоку на попередньому кроці c_{t-1} .

Ключ до LSTM – клітинний стан (cell state) – горизонтальна лінія, що проходить крізь верхню частину діаграми (рисунок 2.5). Клітинний стан являє собою щось на кшталт стрічки конвеєра. Вона проходить вздовж усього ланцюга тільки з невеликими лінійними взаємодіями. Інформація може просто передаватися по ній без змін [14].

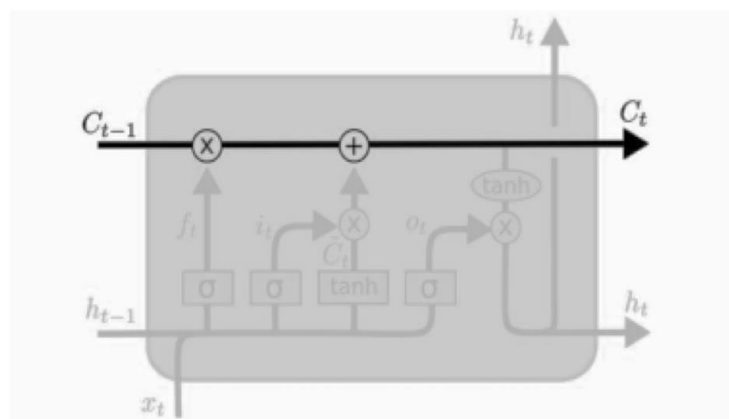


Рисунок 2.5 – Обчислення запам'ятовуючого блоку

LSTM може видаляти інформацію зі стану блоку. Цей процес регулюється структурами, які називаються фільтрами (gates). Фільтри дозволяють пропускати інформацію на підставі деяких умов. Вони складаються з шару сигмоїдальної нейронної мережі та операції поточкового множення. Сигмоїдальний шар повертає числа від нуля до одиниці, які позначають, яку частку кожного блоку інформації слід пропустити далі через мережу [13]. Нуль у даному випадку означає, що не треба пропускати нічого, а одиниця – пропустити все.

Спочатку LSTM визначає, які дані можна виключити зі стану блоку. Це рішення приймає сигмоїдальний шар, що називається шаром фільтра забування (forget gate layer). Він аналізує h_{t-1} і x_t та повертає число від 0 до 1 кожного числа зі стану блоку c_{t-1} .

Сигмоїдальний шар, що називається шаром фільтра забування, можна описати функцією, що має наступний вигляд:

$$f_t = \sigma(W_f x_t + U_f h_{t-1}). \quad (2.1)$$

Після цього LSTM вирішує, яка нова інформація буде зберігатися у стані блоку. Цей етап складається з двох частин – вхідного фільтра (input gate) та обчислення стану блоку пам'яті на поточному кроці.

Вхідний фільтр (input gate) i_t визначає, наскільки значення блоку пам'яті на поточному етапі має впливати на результат. Значення фільтра варіюються від 0 (повністю ігнорувати вхідні значення) до 1, що забезпечується областю значень сигмоїдальної функції:

$$i_t = \sigma(W_i x_t + U_i h_{t-1}). \quad (2.2)$$

Стан блоку пам'яті c_t на поточному кроці можна описати наступною функцією:

$$\tilde{c}_t = \tanh(W_c x_t + U_c h_{t-1}), \quad (2.3)$$

$$c_t = f_t c_{t-1} + i_t \tilde{c}_t. \quad (2.4)$$

Він (2.4) обчислюється на основі всіх даних, що надходять у момент часу t , з використанням фільтрів (2.1) та (2.2) [15].

Далі LSTM вирішує, яку інформацію отримаємо на виході. Вихідні дані будуть засновані на стані блоку та до них буде застосований вихідний фільтр (output gate) [15].

$$o_t = \sigma(W_o x_t + U_o h_{t-1}). \quad (2.5)$$

Підсумкове значення LSTM-шару можна описати функцією, що має наступний вигляд:

$$h_t = o_t \tanh(c_t). \quad (2.6)$$

Підсумкове значення (2.6) визначається вихідним фільтром (2.5) та нелінійною трансформацією над станом блоку пам'яті (2.4).

2.3.2 Нейронна мережа типу GRU

У 2014 році в роботі була представлена модель GRU (Gated Recurrent Unit), яка заснована на тих же принципах, що і LSTM, але використовує менше фільтрів і операцій для обчислення h_t . Фільтр оновлення (update gate) z_t і фільтр скидання стану (reset gate) r_t обчислюються за формулами:

$$z_t = \sigma(W_z x_t + U_z h_{t-1}), \quad (2.7)$$

$$r_t = \sigma(W_r x_t + U_r h_{t-1}). \quad (2.8)$$

Архітектура шару GRU має наступний вигляд (див. рис. 2.6) [13].

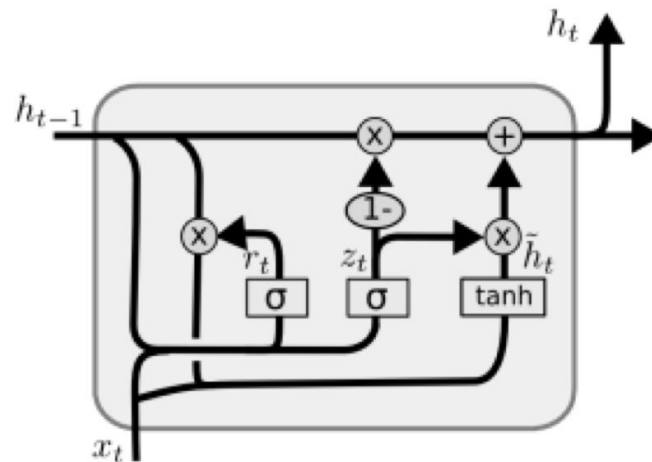


Рисунок 2.6 – Архітектура шару GRU

Вихідне значення вираховується на основі проміжного значення h , яке за допомогою фільтра скидання стану (2.8), визначає, які значення попереднього кроку \tilde{h}_{t-1} слід виключити (пряма аналогія з фільтром забуття з LSTM) [17]:

$$\tilde{h}_t = \tanh(Wx_t + r_t + Uh_{t-1}). \quad (2.9)$$

Підсумкове значення можна описати функцією, що має такий вигляд:

$$h_t = z_t h_{t-1} + (1 - z_t) \tilde{h}_t. \quad (2.10)$$

Підсумкове значення використовує фільтр оновлення (2.7) і проміжне значення (2.9).

Мережа GRU завдяки простішій структурі порівняно з LSTM, може бути ефективнішою з точки зору обчислювальних ресурсів. З іншого боку LSTM більш гнучкий в управлінні інформацією за рахунок використання більшої кількості фільтрів та змінних стану. Вибір між LSTM і GRU залежить від конкретного завдання та доступних ресурсів.

Висновки за розділом 2

У другому розділі було розглянуто означення криптовалюти, її загальні характеристики та принципи використання. Було досліджено вплив різних факторів на ринок криптовалют, зокрема, політичний, економічний, інформаційний впливи тощо.

Окремої уваги було приділено питанням використання сучасних технологій, які можуть бути використані для розв'язання задачі прогнозування цін на криптовалюти, – рекурентним нейронним мережам. Розглянуто загальну структуру рекурентної нейронної мережі та її окремі особливі архітектури – нейронні мережі LSTM та GRU. Обидва типи нейронних мереж можуть бути ефективними під час розв'язання різних задач, пов'язаних з опрацюванням послідовних даних і відображенням часових залежностей, і вибір між ними може бути здійснений експериментально під час розв'язання конкретної прикладної задачі.

3 ПРОГРАМНА РЕАЛІЗАЦІЯ

3.1 Вибір мови програмування та середовища для реалізації програми

Машинне навчання є одним із найактуальніших і важливих напрямів у галузі штучного інтелекту. Воно застосовується у багатьох областях, включаючи розпізнавання зображень, аналіз текстів, прогнозування та інші, де наявні великі обсяги різноманітних даних. Для реалізації проектів у галузі машинного навчання використовують багато різноманітних мов програмування. Кожна з них має свої особливості, які можуть впливати на вибір мови для розв'язання конкретної задачі.

Python є популярною мовою програмування для машинного навчання через його зручність, високу продуктивність і величезне різноманіття інструментів та бібліотек. Автор мови (Гвідо ван Россум) зробив акцент на читабельність коду та простоту синтаксису, тому Python дозволяє розробнику робити акцент на аналізі та оптимізації отриманих результатів, а не витратити зайвий час на пошук помилок у складних синтаксичних конструкціях коду під час реалізації задач.

Важливим є те, що середовище мови Python та безліч її фреймворків є безкоштовними та доволі простими у використанні.

До переваг використання Python для розв'язання складних аналітичних задач можна віднести:

- підтримку високорівневих математичних операцій;
- розширену підтримку векторизованих операцій із використанням NumPy;
- велику бібліотеку інструментів для візуалізації даних з використанням Matplotlib та Seaborn;
- підтримку безлічі алгоритмів машинного навчання з використанням бібліотек, таких як scikit-learn та TensorFlow;
- підтримку багатопоточності та масштабування;

В якості середовища для розробки було обрано Visual Studio code. Із основних переваг VS code треба виділити те, що він є швидким та потужним з точки зору реалізації коду.

3.2 Алгоритм розв'язання задачі прогнозування цін на ринку криптовалют методами машинного навчання

Виділимо основні етапи розв'язання задачі прогнозування цін на ринку криптовалют.

Етап 1. Підготовка даних. На першому етапі програма зчитує дані з таблиці та виводить їх у вигляді часового ряду. Для уникнення помилок треба очистити дані від пропущених значень і аномалій та нормалізувати дані для полегшення навчання нейронної мережі.

Ділимо історичні данні на часові вікна, які будуть використовуватись для прогнозування. Відповідно до розміру даних задається розмір вибірки, відносно якої дані діляться на навчальний та тестові набори.

Етап 2. Наступним етапом є вибір типу та архітектури мережі, в роботі було розглянуто два типи рекурентних мереж: LSTM і GRU. Для покращення результатів прогнозу треба налаштувати гіперпараметри:

- визначити кількість шарів;
- визначити кількість нейронів у кожному шарі;
- обрати функцію активації для рекурентних шарів;
- обрати функцію втрат для вимірювання помилок між прогнозами та фактичними значеннями;
- обрати метод оптимізації, такий як Adam, SGD або RMSprop, для навчання моделі;
- визначити кількість зразків у пакеті (Batch_Size) для кожної епохи навчання;
- епохи відповідають за кількість разів, коли весь навчальний набір пройде через мережу під час навчання;

– налаштувати Dropout, який відключає частину нейронів у випадку перенавантаження.

Етап 3. Після налаштування всіх необхідних параметрів та налаштування мережі можна переходити до навчання моделі на навчальному наборі, який було визначено на першому етапі, та оцінки якості роботи навченої моделі на тестових даних.

Навчена модель далі буде використовуватись для прогнозування майбутніх значень цін. Для візуалізації та подальшого порівняння результатів будемо графіки.

3.3 Опис програми

В процесі виконання роботи було розроблено та реалізовано систему, яка робить прогноз на криптовалютну пару за допомогою рекурентних нейронних мереж.

Для навчання і подальшого прогнозування було обрано криптовалюту Litecoin (LTC). Litecoin (LTC) був одним із перших створених альткоїнів. Його блокчейн, створений колишнім інженером Google, Charlie Lee у 2011 році, був розроблений на основі відкритого вихідного коду Bitcoin. Але Litecoin представив певні модифікації, такі як більш висока швидкість генерації блоків та інший алгоритм майнінгу Proof of Work (PoW) під назвою Scrypt [20].

Litecoin має обмежену загальну пропозицію у 84 мільйони. Подібно до Bitcoin, Litecoin може бути отриманий шляхом майнінгу і має механізм халвінгу, який відбувається кожні 840 000 блоків (приблизно 4 роки). Останній халвінг LTC відбувся у серпні 2019 року, коли винагорода за блок була зменшена вдвічі з 25 LTC до 12,5 LTC [20].

Вхідними даними для обчислювальних експериментів є історичні дані про ціни на LTC, які було взято з інтернет ресурсу – Yahoo Finance [21]. Дані подані у вигляді таблиці, що містить інформацію про щоденні ціни на криптовалюту з 1 січня 2018 року до 1 січня 2024 року (2214 записів).

Дана база даних містить наступні параметри:

Date – дата, станом на 00:00 кожного дня, для якої наведені наступні показники.

Open – початкова ціна криптовалюти з початку нового дня.

High – найвища ціна на криптовалюту впродовж певного дня.

Low – найнижча ціна на криптовалюту впродовж певного дня.

Close – ціна криптовалюти, станом на закриття торгів певного дня.

Задачу прогнозування курсів криптовалют з використанням рекурентних нейронних мереж LSTM і GRU було реалізовано мовою програмування Python. В роботі ми використовували наступні бібліотеки: numpy, pandas, matplotlib.pyplot, keras, sklearn.

За допомогою бібліотеки Pandas ми зчитуємо данні з .csv файлу та сортуємо їх за датою. Конвертуємо стовпець «Close» у числовий тип даних float та будуємо графік реальних цін на LTC.

За для стабільності та ефективності навчання нейронних мереж використали MinMaxScaler, який нормалізує дані у діапазон значень від 0 до 1. Нормалізація зменшує вплив розмірів вхідних даних на модель та полегшує збіжність алгоритму оптимізації. Для часового вікна обрали 60 днів. Розмір вибірки було задано у відсотковому співвідношенні – 60% були використані для навчання, 40% використали для тренування.

Побудова моделі LSTM була реалізована засобами бібліотеки Keras. В якості оптимізатора використовували Adam. Для того, щоб результати прогнозування були точними і відповідали дійсності, експериментальним шляхом необхідно правильно налаштувати параметри та гіперпараметри моделі.

Для оцінки точності прогнозу використали три метрики: MSE (Mean Squared Error, середньоквадратична помилка), MAE (Mean Absolute Error, середня абсолютна помилка), та RMSE (Root Mean Squared Error, квадратична коренева середня помилка).

На основі всіх отриманих даних програма робить прогноз на основі моделі LSTM та GRU. Як результат, ми отримуємо графіки з прогнозами цін на криптовалюту LTC та оцінку їх точності.

Програма має зручний інтерфейс та легку читабельність коду, що дасть змогу інвесторам легко використовувати графіки з прогнозами для аналізу та планування своїх дій та змінювати налаштування моделі для отримання більш точного результату.

Висновки за розділом 3

В даному розділі розглянуто особливості мови програмування Python як однієї з найпопулярніших сучасних мов програмування, які використовуються для реалізації проєктів глибокого машинного навчання. Виділено особливості цієї мови, які дозволили нам обрати її для розв'язання задачі, якій присвячена кваліфікаційна робота.

Наведено основні етапи розв'язання задачі прогнозування цін на ринку криптовалют. Першим етапом є підготовка даних і подання їх у зручному для подальшого аналізу вигляді. Другий етап передбачає вибір типу та архітектури нейронної мережі та налаштування її гіперпараметрів. Наступним етапом є навчання моделі та оцінка її якості. Навчена модель може використовуватись для прогнозування.

4 РЕЗУЛЬТАТИ ОБЧИСЛЮВАЛЬНОГО ЕКСПЕРИМЕНТУ ТА ЇХ АНАЛІЗ

На рисунку 4.1 подано історичні дані цін на криптовалюту Litecoin (LTC), для якої будемо виконувати аналіз. Дані подано у вигляді часового ряду.

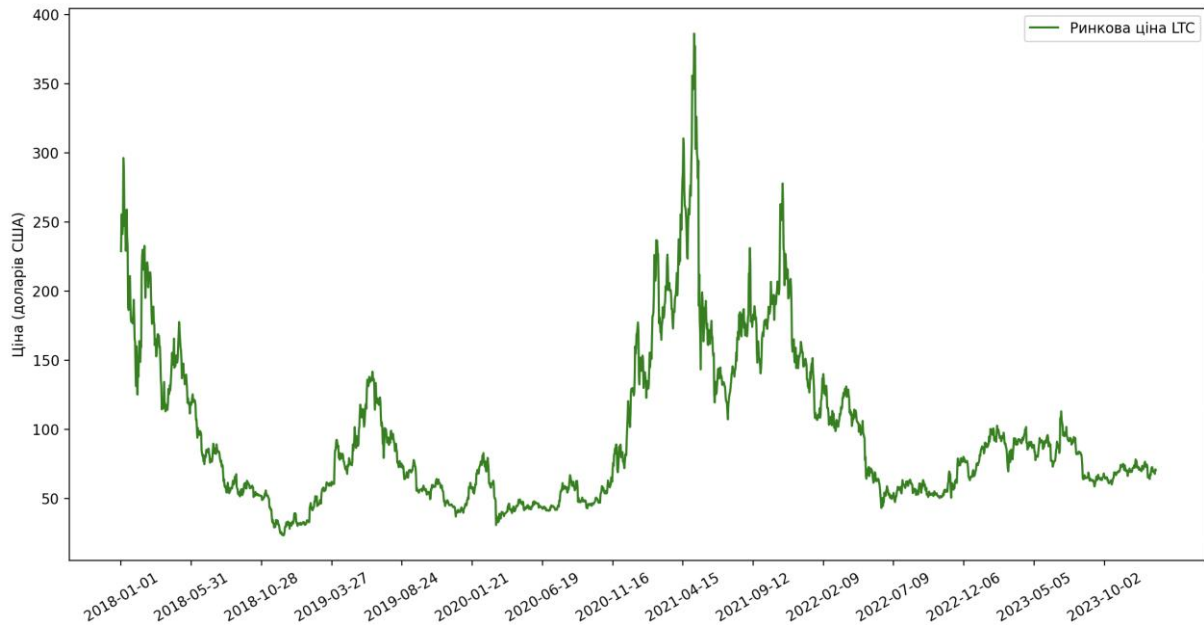


Рисунок 4.1 – Динаміка цін LTC

Перед використанням датасету для навчання нейронної мережі було виконано нормалізацію даних. Після підготовки даних програма розбиває на дасет на два піднабори – навчальних даних та тестових даних і переходимо до навчання нейронної мережі.

На рисунках 4.2 і 4.3 наведено результати прогнозування цін для тестового набору даних, коли навчання відбувалось протягом однієї епохи.

Модель LSTM, прогноз якої на тестових даних (червона крива) виведений на рисунку 4.2, має 3 шари по 50 нейронів на кожному, Dropout – 0,2 на кожному непарному шарі та розглянуто 30 часових кроків.

Модель GRU, прогноз якої на тестових даних (синя крива) виведений на рисунку 4.3, має 4 шари, кожен з яких має 50 нейронів, Dropout – 0,2 на кожному непарному шарі, та розглянуто 30 часових кроків.

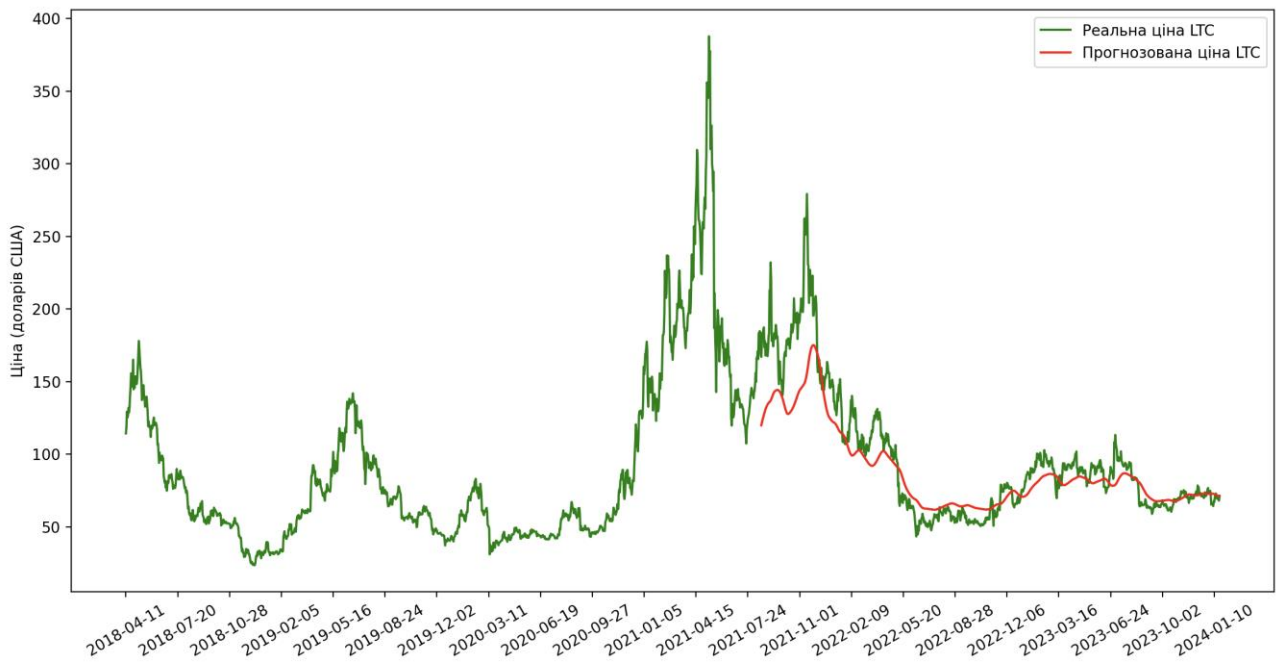


Рисунок 4.2 – Прогнозування цін нейронною мережею LSTM, тестовий набір

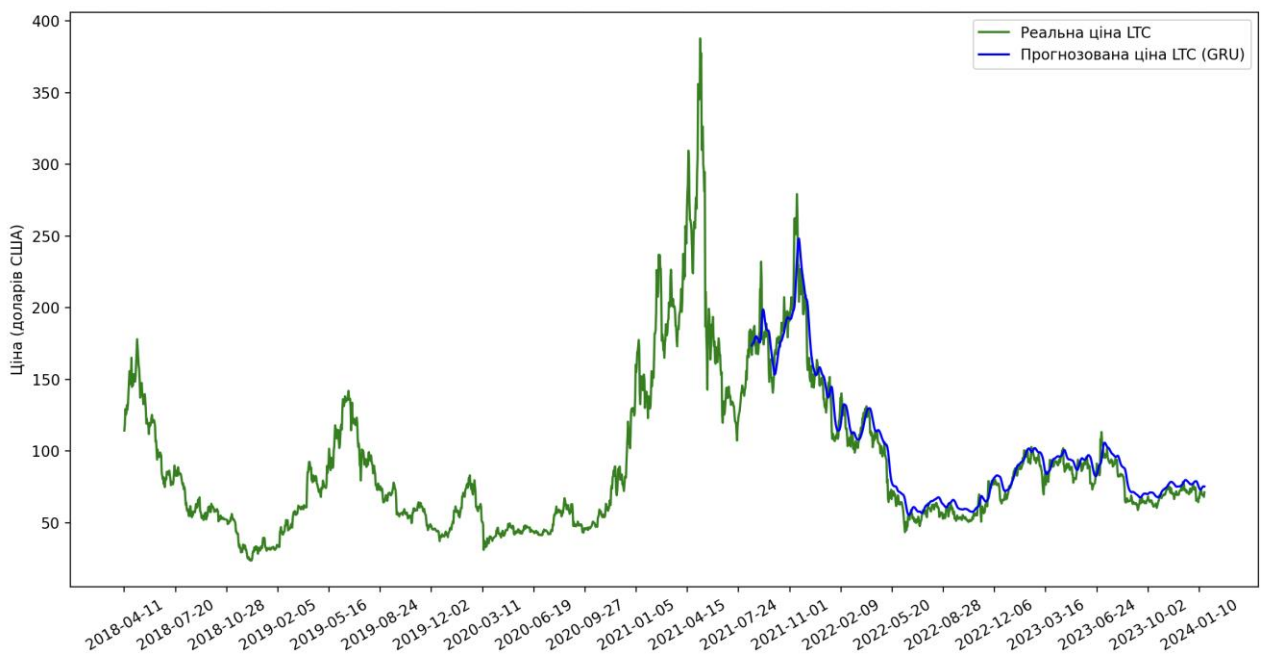


Рисунок 4.3 – Прогнозування цін нейронною мережею GRU, тестовий набір

На рисунках 4.2 – 4.3 наведений приклад не досить вдало підібраних параметрів. Ми бачимо, що графіки (червона та синя криві) частково повторюють реальні тенденції (зелена крива), але різниця між прогнозами та реальними значеннями є суттєвою, про що свідчить також значення середньо-абсолютної похибки (рисунок 4.4 – 4.5).

MSE: 283.78850496405965
MAE: 11.709289826374635
RMSE: 16.846023416939076

Рисунок 4.4 – Метрики для LSTM моделі

MSE: 114.37576378582744
MAE: 7.838850267625574
RMSE: 10.694660526909091

Рисунок 4.5 – Метрики для GRU моделі

Для покращення результатів збільшимо кількість епох. Результати експерименту наведені на рисунках 4.6 – 4.9.

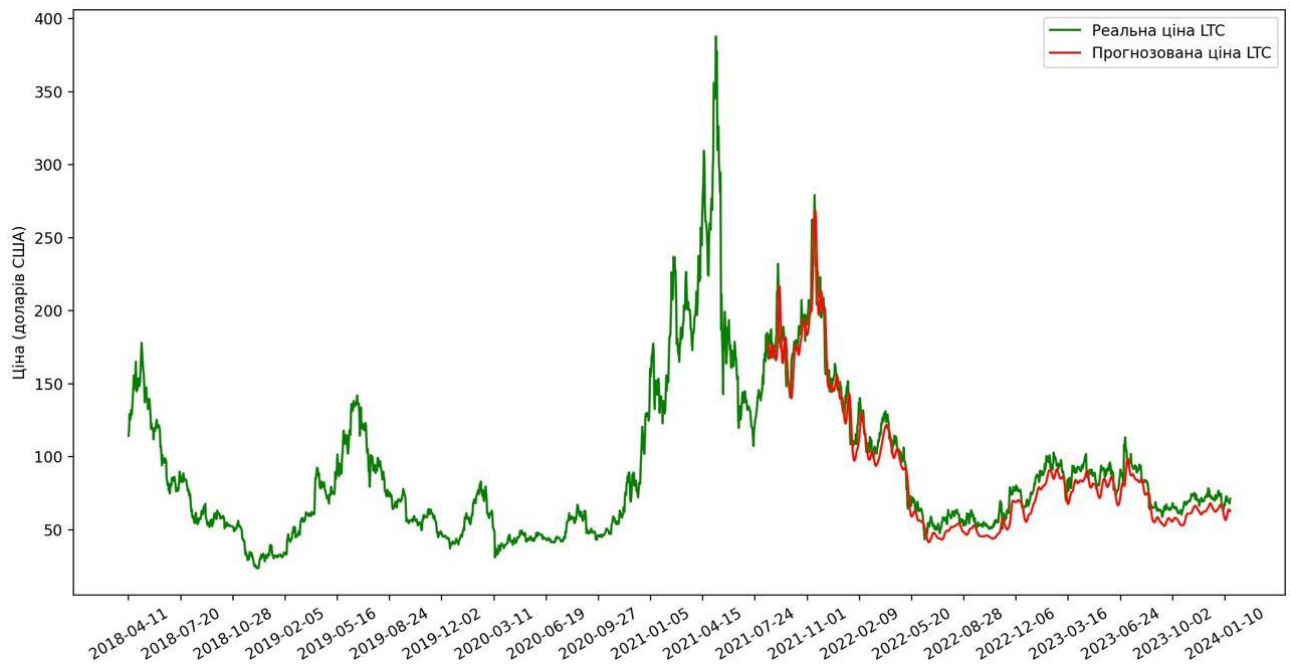


Рисунок 4.6 – Прогнозування цін нейронною мережею LSTM, тестовий набір

MSE: 70.139853722735
MAE: 5.432435583900236
RMSE: 8.374953953469536

Рисунок 4.7 – Метрики для LSTM моделі

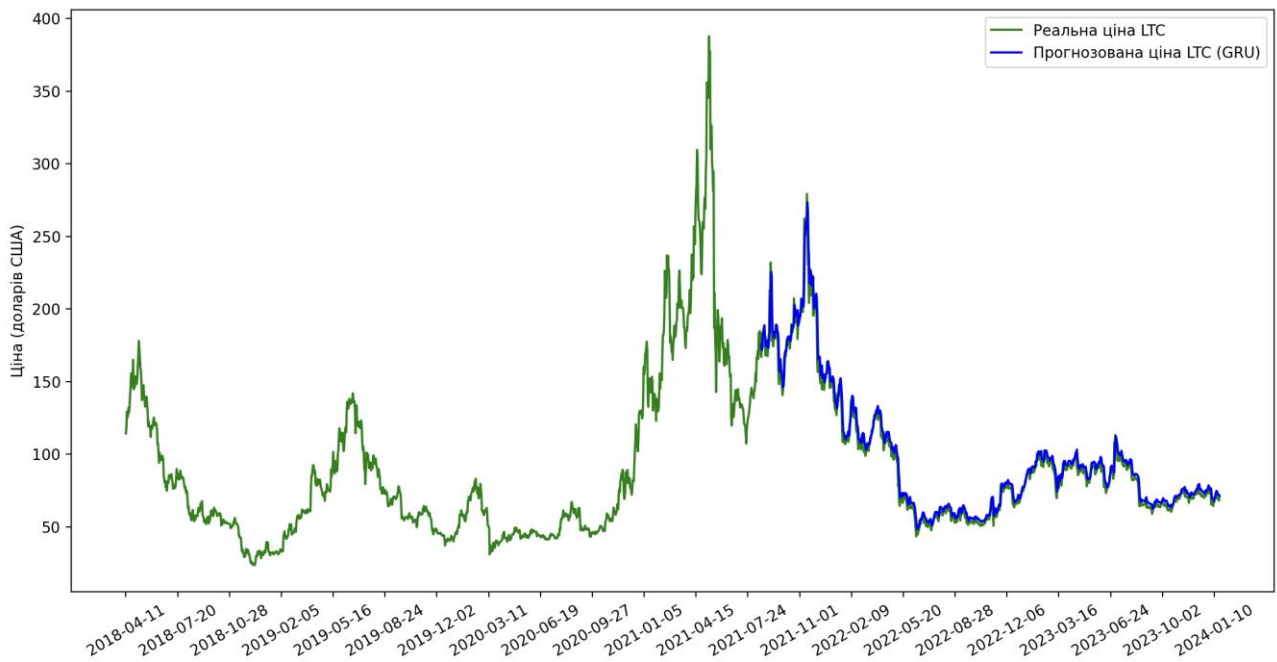


Рисунок 4.8 – Прогнозування цін нейронною мережею GRU, тестовий набір

MSE: 37.75993532776699
 MAE: 4.183086357019664
 RMSE: 6.144911336037892

Рисунок 4.9 – Метрики для GRU моделі

Моделі, прогнози яких зображені на рисунках 4.6 та 4.8, мають по 3 шари по 50 нейронів на кожному, Dropout – 0,2 на кожному непарному шарі. Кількість епох в обох моделях рівна 30, що є оптимальним значенням. Це значення було отримане завдяки моніторингу втрат навчання за допомогою середньо-квадратичної похибки (MSE). З кожною епохою втрати зменшуються, а коли зменшення стає мінімальним або ж втрати починають зростати, то це і є оптимальною кількістю епох, яку потрібно використовувати. Графіки загалом майже повністю відбивають реальні тенденції, а різниця між прогнозованими і реальними значеннями є відносно малою. Значення метрик (рисунки 4.7, 4.9) суттєво покращились порівняно з результатами тренування на попередніх мережах.

При даному наборі параметрів обидві моделі краще навчаються і здійснюють точніший прогноз. При цьому модель GRU дає точніший прогноз

порівняно з моделлю LSTM. Про це свідчать значення метрик MSE, MAE, RMSE (рисунок 4.9).

У наступному експерименті додамо ще прихованих шарів та збільшимо кількість нейронів для кожного шару. Результати експериментів з прогнозування цін на оцінки якості навчання відповідних моделей наведено на рисунках 4.10, 4.11.

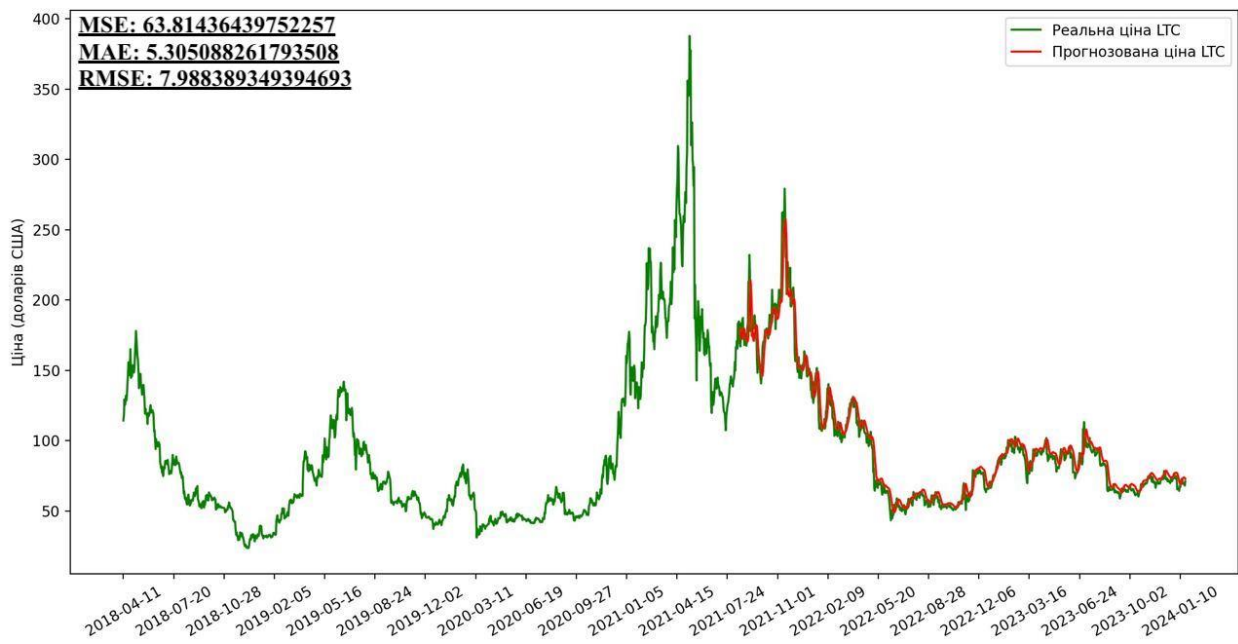


Рисунок 4.10 – Прогнозування цін нейронною мережею LSTM, тестовий набір

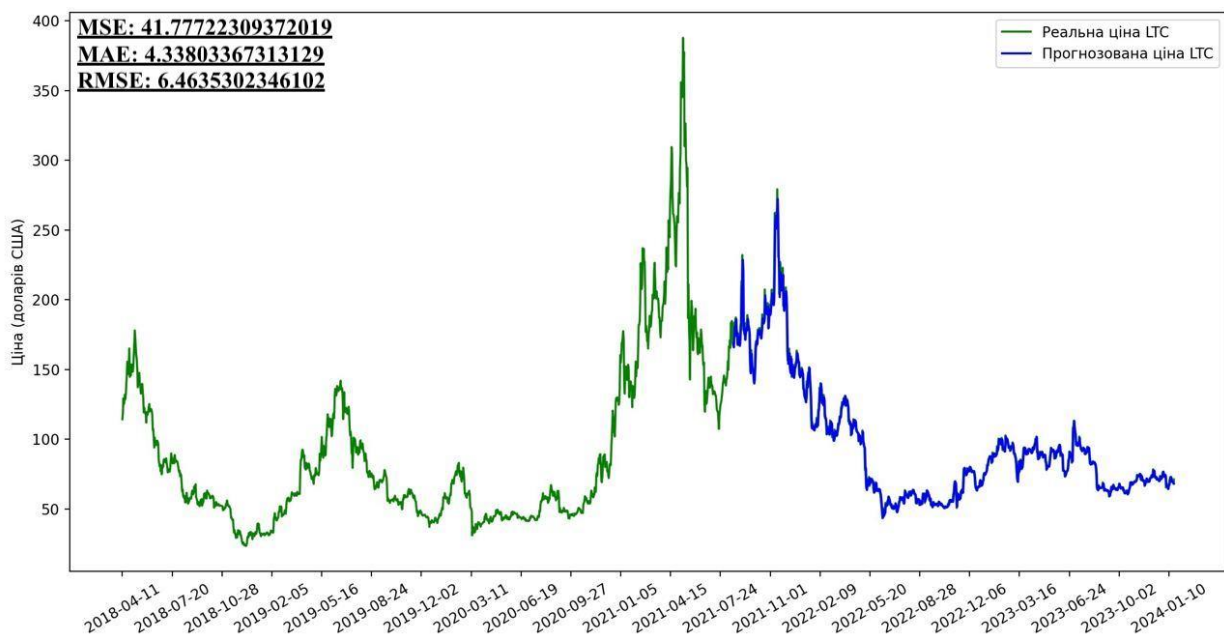


Рисунок 4.11 – Прогнозування цін нейронною мережею GRU, тестовий набір

На рисунках 4.10 – 4.11 зображені моделі, які мають по 4 шари по 50 нейронів на кожному, Dropout – 0,2 на кожному непарному шарі. Аналізуючи значення метрик, обчислених на тестовому наборі даних, можна робити висновок, що збільшення кількості шарів в нашому випадку робить результат навчання більш точним, але занадто велика кількість шарів і нейронів можуть призвести до перенавчання, що в свою чергу, може погіршити результат.

Збільшимо кількість нейронів для кожного шару до 100 у обох моделях та збільшимо значення шару випадкового вимкнення (Dropout) до 0,3, щоб уникнути перенавчання. Розмір зразків для пакету зменшимо до 10-ти. експериментів з прогнозування цін на оцінки якості навчання відповідних моделей наведено на рисунках 4.12, 4.13.

Підбором параметрів нам вдалось покращити результат прогнозування навчальної вибірки на скільки це можливо. Виходячи з результатів можна впевнено казати, що модель GRU набагато краще справляється з прогнозом в контексті даної роботи при коректно підібраних параметрах, про що свідчать значно менші значення метрик, що оцінюють якість навчання, порівняно з мережею LSTM.

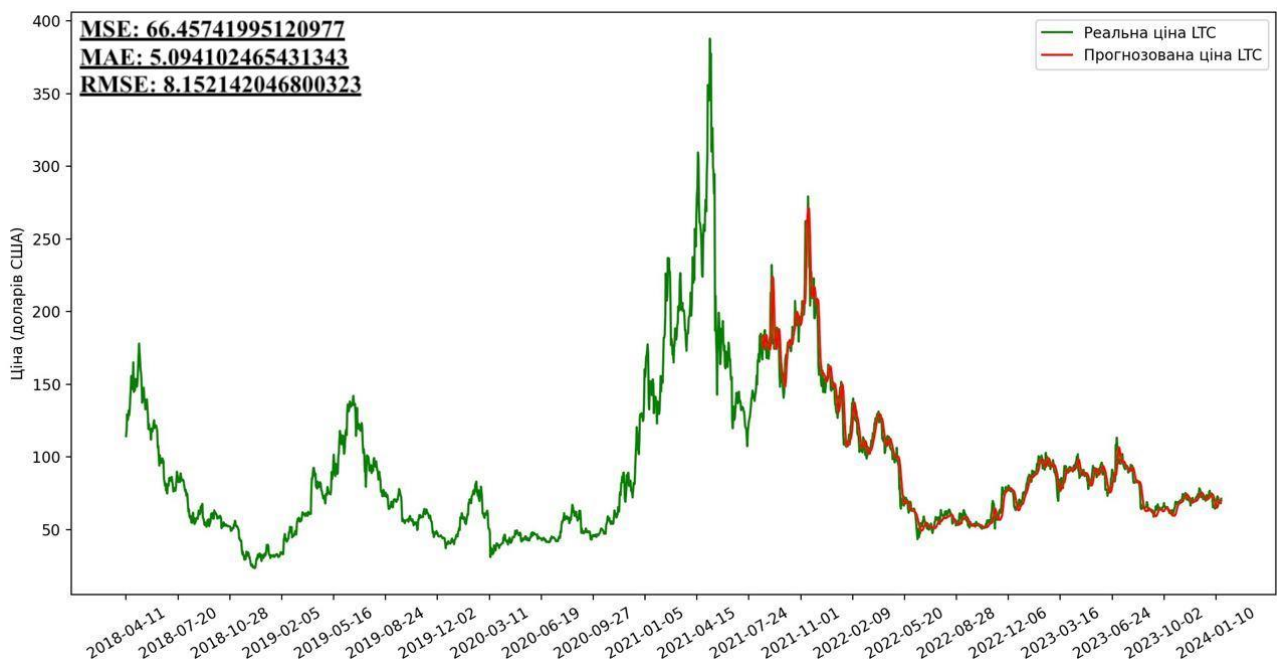


Рисунок 4.12 – Прогнозування цін нейронною мережею LSTM, тестовий набір

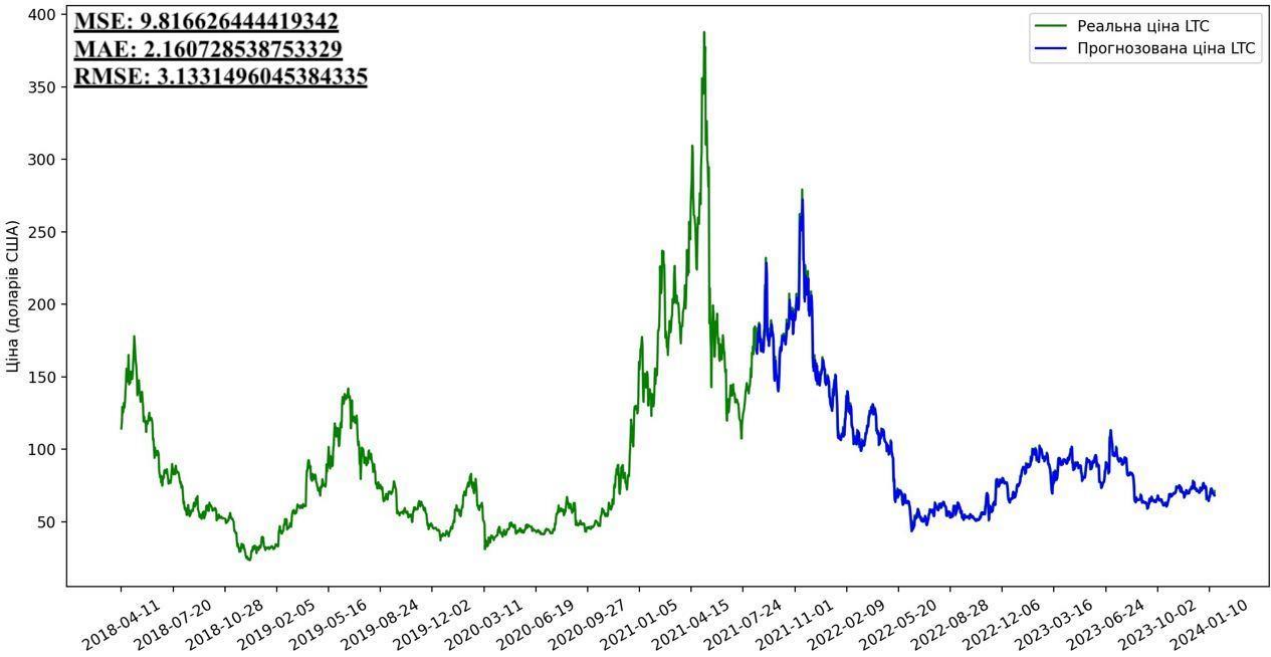


Рисунок 4.13 – Прогнозування цін нейронною мережею GRU, тестовий набір

Далі розглянемо можливості прогнозування цін побудованими нейронними мережами LSTM та GRU на наступні 30 днів. Результати прогнозування наведено на рисунках 4.14, 4.15.

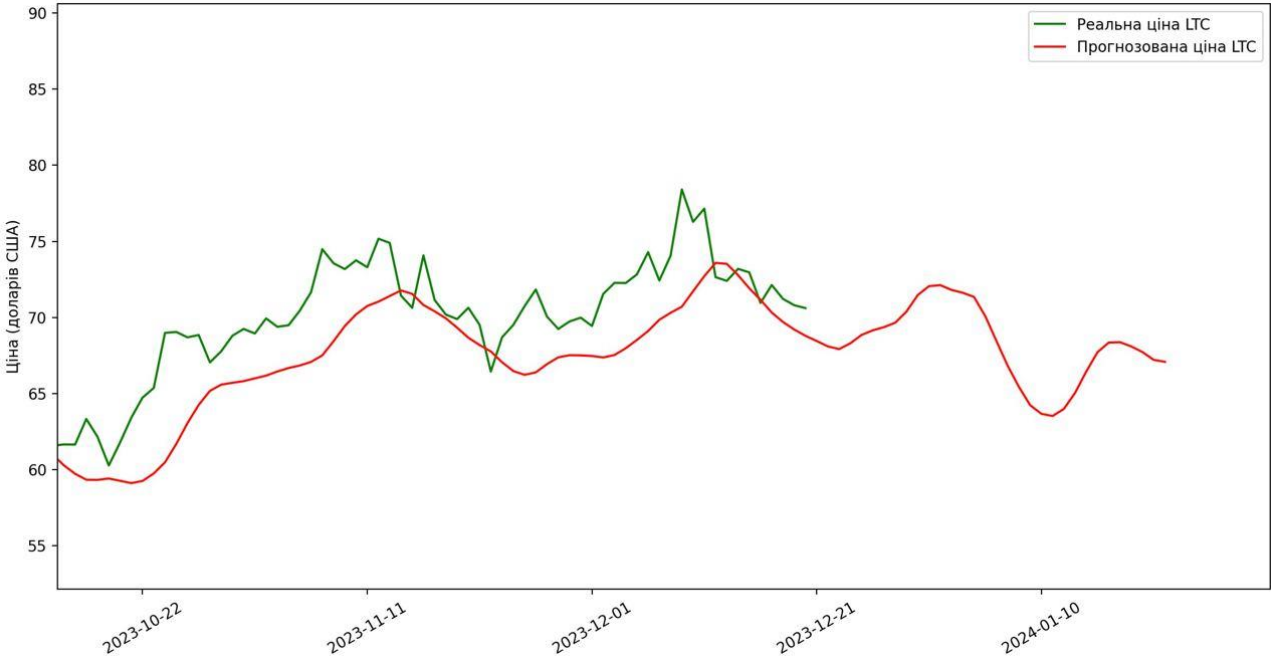


Рисунок 4.13 – Прогнозування цін нейронною мережею LSTM на наступні 30 днів

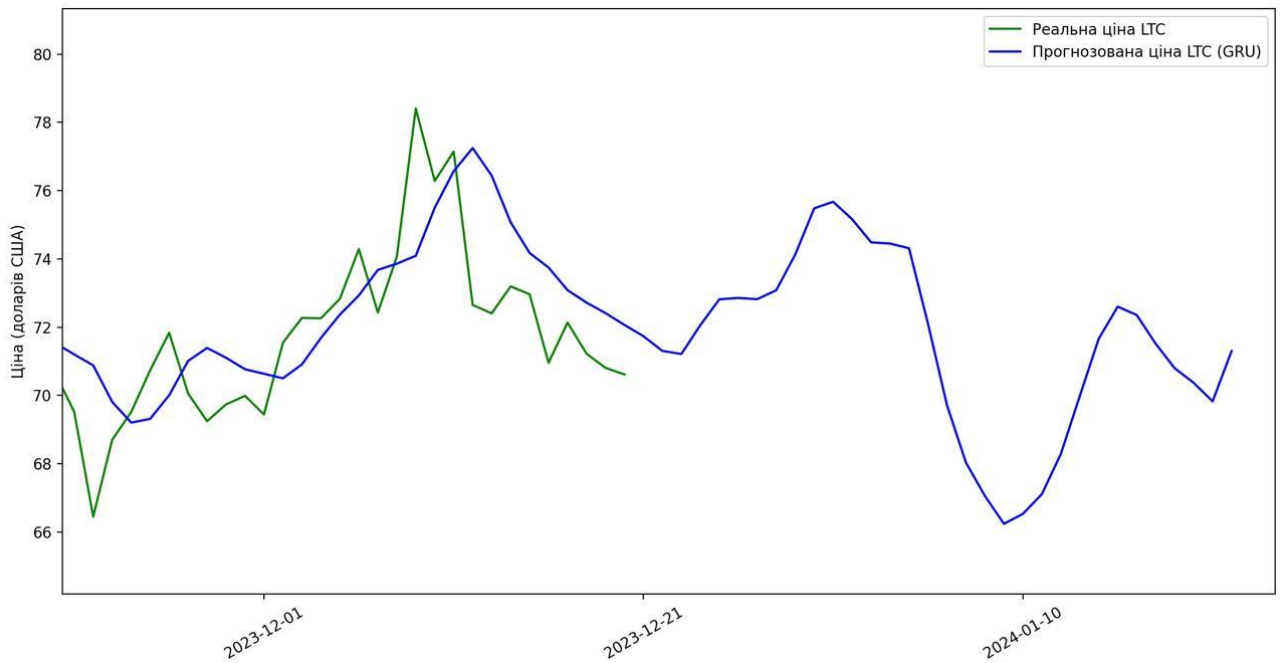


Рисунок 4.14 Прогнозування цін нейронною мережею GRU на наступні 30 днів

Аналіз графіків дозволяє зробити такі висновки. В першому випадку, де аналіз проводиться за допомогою LSTM мережі, отримані результати досить неточні., це видно з тієї частини графіка, де наявні дані тестового набору. Тому, якщо використовувати цю модель для ф'ючерсної торгівлі, треба особливо уважно обирати розмір кредитних плечей. У другому випадку, де аналіз проводиться за допомогою GRU, результат аналізу більш точний, з мінімальним значенням відхилення, тому GRU краще підходить для виконання задачі прогнозування криптовалют. Зауважимо, ще результати можуть залежати від використаного набору даних.

Графіки в обох моделях повторюють тенденції ринку, але з невеликою затримкою, що вказує на необхідність доопрацювання архітектури нейронної мережі. Затримка може бути викликана великою кількістю факторів: низька якість даних, інформаційний або політичний вплив, наявність латентних факторів. Все це може впливати на якість прогнозування та на перенавчання моделі.

Для усунення цієї проблеми необхідно провести більше експериментальних досліджень, змінюючи параметри моделі, такі як кількість прихованих шарів, кількість епох навчання, обсяг тренувальних даних. Крім

цього перспективним є використання інших архітектур нейронних мереж, які також здатні працювати з даними, що являють собою часові послідовності.

Висновки за розділом 4

У даному розділі проведений аналіз результатів експериментів, виконаних за допомогою програмного продукту, що реалізує прогнозування цін на криптовалюту. Реалізація задачі прогнозування була виконана за допомогою рекурентних нейронних мереж LSTM і GRU. Програма була написана мовою програмування Python з залученням необхідних бібліотек.

В ході виконання задачі було опрацьовано датасет, в якому зберігаються історичні відомості щодо цін на криптовалюту LTC з 2018 по 2024 рік. На основі даного датасету було проведено підбір параметрів та типу рекурентної нейронної мережі, виконано навчання її та отриманий прогноз на наступні 30 днів з подальшим аналізом. В ході дослідження було виявлено, що модель дає гарний прогноз на найближчій період часу.

ВИСНОВКИ

У кваліфікаційній роботі було розв'язано задачу прогнозування цін на ринку криптовалют методами машинного навчання. Була проаналізована проблема прогнозування цін на криптовалюту та виконаний системний аналіз даної проблеми. За допомогою методу аналізу ієрархій було обрано оптимальний за певними критеріями метод розв'язання поставленої задачі – метод нейронних мереж. Вибір методу підтверджується його популярністю при обробці великих обсягів даних, що є часовими послідовностями.

У роботі розглянуто особливості криптовалют та вплив різних факторів на формування їх ціни, а також нейромережеві технології, такі як рекурентні нейронні мережі LSTM і GRU, що можуть використовуватись для прогнозування ціни.

Для реалізації прогнозу курсу криптовалюти LTC було використано обидві розглянуті нейронні мережі. Шляхом експериментальних досліджень було виявлено оптимальні параметри та гіперпараметри для навчання моделі у контексті нашої задачі і виявлено, що GRU має більшу швидкість навчання та точність прогнозування, що робить її більш ефективною для роботи з великими обсягами даних. Також проведено порівняльний аналіз результатів, який підтвердив перевагу GRU.

За допомогою побудованих нейронних мереж був зроблений прогноз курсу криптовалюти LTC на найближчі 30 днів, і визначено точність моделі на тренувальній вибірці. Зазначено, що добре побудований прогноз на тренувальній вибірці свідчить про його потенційну можливість вказувати на майбутні тренди.

У подальших дослідженнях можна розглянути удосконалення програмного продукту шляхом використання декількох нейронних мереж, а також проведення експериментів з різними криптовалютами для визначення найкращої моделі прогнозування. Також можна розробити більш складну модель, додавши ще додаткові методи для поліпшення точності прогнозів.

ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ

1. Integration definition for function modeling (IDEF0). Draft Federal Information Processing Standards Publication 183, 1993. 325 p.
2. Azzolini J. Introduction to Systems Engineering Practices, 2001. 250 p.
3. Gordiievych A., Shubin I. Forecasting of airfare prices using time series // 2015 Information Technologies in Innovation Business Conference, ITIB 2015 – Proceedings, 2015. P. 68-71.
4. Теслюк Н. П., Жулин О. В., Назаренко Я. Я., Куїмова А. В. Аналіз розвитку криптовалют на віртуальному ринку фінансових послуг. *Ефективна економіка*, 2021. № 7. URL: <http://www.economy.nayka.com.ua/?op=1&z=9052> (дата звернення: 24.06.2023).
5. Narayanan A. Bitcoin and Cryptocurrency Technologies: A Comprehensive Introduction / Arvind Narayanan, Joseph Bonneau, Edward Felten, Andrew Miller & Steven Goldfeder. *Princeton University Press*. 2016. 304 p.
6. Bariviera A. F., Basgall M. J., Hasperué W., Naiouf M. Some stylized facts of the Bitcoin market. *Physica A: Statistical Mechanics and its Applications*. 2017. P. 82–90.
7. Bonneau J., Miller A., Clark J., Narayanan A., Kroll J., Felten E. SoK: Research Perspectives and Challenges for Bitcoin and Cryptocurrencies. IEEE Symposium on Security and Privacy, 2015. P. 104-121.
8. Pillai S., Biyani D., Motghare R., Karia D. Price Prediction and Notification System for cryptocurrency Share Market Trading // 2021 International Conference on Communication information and Computing Technology (ICCICT), 2021, P. 1-7.
9. Crypto Predictions. URL: <https://cryptopredictions.com/> (дата звернення: 25.10.2023).
10. Coincodex. URL: <https://coincodex.com/> (дата звернення: 25.03.2023).
11. Trading Beasts. URL: <https://tradingbeasts.com> (дата звернення: 25.03.2023).
12. Coin Market cap. URL: <https://coinmarketcap.com/> (дата звернення: 25.03.2023).

13. Avdieiev O. D., Kauk V. I. Research of methods of forecasting the dynamics of the course of cryptocurrencies, 2023.
14. Transformer neural network. URL: <https://builtin.com/artificial-intelligence/transformer-neural-network> (дата звернення: 03.04.2023).
15. Hochreiter S., Schmidhuber J. Long short-term memory // *Neural computation*. 1997. Vol. 9, no. 8. P. 1735 –1780.
16. Multi-resolution linear prediction based features for audio onset detection with bidirectional LSTM neural networks / E. Marchi [et al.] // *Acoustics, Speech and Signal Processing (ICASSP), 2014 IEEE International Conference IEEE*. 2014. – P. 2164- 2168.
17. Dutta A., Kumar S., Basu M. A gated recurrent unit approach to bitcoin price prediction. *Journal of Risk and Financial Management*, vol. 13, 2020. 23 p.
18. Smelyakov K., Bizkrovnyi O., Sharonov, N., Smelyakov S., Chupryna A. Building of Regression Models for Cryptocurrency Price Prediction CEUR Workshop Proceedings, 2022, 3171, P. 1216–1232.
19. Максименко А. С. Аналіз підходів до розв'язання задачі прогнозування ринку криптовалют // 27-й Міжнародний молодіжний форум «Радіоелектроніка і молодь у XXI столітті». Зб. матеріалів форуму. Т. 7. Харків : ХНУРЕ. 2023. С.167–168.
20. Що таке Litecoin (LTC)? URL: <https://academy.binance.com/uk/articles/what-is-litecoin-ltc> (дата звернення: 25.10.2023).
21. Litecoin USD (LTC-USD) Price History & Historical Data - Yahoo Finance URL: <https://finance.yahoo.com/quote/LTC-USD/history> (дата звернення: 20.12.2023).