

prise (with Translation of Topics into English)]. Lviv: Vydavnytstvo Lvivskoi politekhniki, 2017.

Honcharenko, D. "Shyfruvannia: typy i alhorytmy. Shcho tse, chym vidrizniaiutsia i de vykorystovuiutsia?" [Encryption: Types and Algorithms. What Is It, How are They Different and Where are They Used?]. *Host-Pro*. July 16, 2020. <https://hostpro.ua/wiki/ua/security/encryption-types-algorithms/>

Kurylo, O. B. et al. *Finansy pidpriemstv* [Enterprise Finance]. Kyiv: VD «Kondor», 2020.

Ostapov, S. E., Yevseiev, S. P., and Korol, O. H. *Kiberbezpeka suchasni tekhnologii zakhystu* [Cyber Security Modern Protection Technologies]. Kyiv: Novyi svit-2000, 2023.

Partyn, H. O., and Seliuchenko, N. Ye. *Finansovy menedzhment* [Financial Management]. Lviv: Vydavnytstvo Lvivskoi politekhniki, 2018.

Pritchard, C. L. *Risk Management: Concepts and Guidance*. London; New York: CRC Press, 2015.

Susidenko, O. V. *Finansova bezpeka pidpriemstva: teoriia, metody, praktyka* [Financial Security of the Enterprise: Theory, Methods, Practice]. Kyiv: TsUL, 2019.

Uninets-Khodakivska, V. P., Kostiukevych, O. I., and Liatambor, O. A. *Rynok finansovykh posluh: teoriia i praktyka* [Market of Financial Services: Theory and Practice]. Kyiv: TsUL, 2021.

Wells, J. T. *Corporate Fraud Handbook: Prevention and Detection*. Hoboken, NJ: Wiley, 2017.

УДК 330.4:519.8

JEL: C89; D21; F31; L86; O24

DOI: <https://doi.org/10.32983/2222-4459-2024-9-129-138>

МОДЕЛІ АНАЛІЗУ ДИНАМІКИ РИНКУ КРИПТОВАЛЮТ З УРАХУВАННЯМ ПОВЕДІНКОВИХ МЕТРИК СТЕЙКХОЛДЕРІВ ЗА ДАНИМИ СОЦІАЛЬНИХ МЕРЕЖ

©2024 ГУР'ЯНОВА Л. С., ЛУЦЕНКО Р. Р.

УДК 330.4:519.8

JEL: C89; D21; F31; L86; O24

Гур'янова Л. С., Луценко Р. Р. Моделі аналізу динаміки ринку криптовалют з урахуванням поведінкових метрик стейкхолдерів за даними соціальних мереж

Використання методів інтелектуального аналізу даних у контексті поведінкової економіки віртуальних активів на основі даних соціальних мереж дозволяють більш точно оцінювати цінові рухи криптовалют. У дослідженні побудовано моделі прогнозування цін на ринку криптовалют з урахуванням поведінкових факторів стейкхолдерів на основі соціальних даних із платформи TikTok. Дані для цього дослідження отримані за допомогою прикладних програмних інтерфейсів соціальних мереж. Основні етапи дослідження включали збір даних, їх обробку та агрегацію, масштабування та кореляційний аналіз, побудову та оцінку моделей. У результаті дослідження визначено ключові поведінкові метрики соціальних мереж. Кореляційний аналіз продемонстрував наявність сильних лінійних зв'язків між соціальними показниками TikTok та слабкі зв'язки із ціною біткоїна. У дослідженні побудовані лінійні моделі, поліноміальна регресія, дерево рішень та «випадковий ліс». Використані такі поведінкові метрики, як кількість публікацій, лайків, коментарів, поширень і переглядів. Проведено оцінку моделей шляхом тестування за допомогою метрик MSE і MAE. Результати свідчать про обмежену ефективність лінійної регресії для прогнозування цін на криптовалюту через нелінійну природу ринку. Модель дерева рішень продемонструвала певний успіх у прогнозуванні цін на біткоїні, проте зросли відхилення у прогнозах з часом, особливо в умовах ринкових коливань. Поліноміальна регресія і модель «випадкового лісу» демонструють вищу точність у прогнозах. На основі порівняння показників MSE і MAE «випадковий ліс» виявився найефективнішою моделлю для прогнозування цін біткоїна серед розглянутих.

Ключові слова: криптовалюти, патерни поведінки, API (application programming interface), соціальні мережі, машинне навчання, поведінкова економіка, моделі прогнозування.

Рис.: 6. **Бібл.:** 15.

Гур'янова Лідія Семенівна – доктор економічних наук, професор, професор кафедри економічної кібернетики та прикладної економіки, Харківський національний університет імені В. Н. Каразіна (майдан Свободи, 4, Харків, 61022, Україна)

E-mail: guryanovalidiya@gmail.com

ORCID: <https://orcid.org/0000-0002-2009-1451>

Researcher ID: <https://www.webofscience.com/wos/author/record/L-3402-2017>

Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=36068855600>

Луценко Ростислав Русланович – аспірант кафедри економічної кібернетики та прикладної економіки, Харківський національний університет імені В. Н. Каразіна (майдан Свободи, 4, Харків, 61022, Україна)

E-mail: roxanisen@gmail.com

ORCID: <https://orcid.org/0000-0003-0737-3902>

Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=57238374000>

UDC 330.4:519.8
JEL: C89; D21; F31; L86; O24

**Guryanova L. S., Lutsenko R. R. Models for Analyzing Cryptocurrency Market Dynamics with Behavioral Metrics of Stakeholders
Based on Social Media Data**

The use of data mining techniques in the context of the behavioral economy of virtual assets based on social media data allows for a more accurate assessment of the price movements of cryptocurrencies. The study builds models for forecasting prices in the cryptocurrency market, taking into account the behavioral factors of stakeholders based on social data from the TikTok platform. The data for this study were obtained using social media application programming interfaces. The main stages of the study included data collection, processing and aggregation, scaling and correlation analysis, building and evaluating the relevant models. As a result of the study, key behavioral metrics of social networks are identified. Correlation analysis showed strong linear links between TikTok's social performance and weak links to the price of bitcoin. The study builds linear models, polynomial regression, Decision Tree and Random Forest. Behavioral metrics such as the number of shares, likes, comments, shares, and views are used. Models are evaluated by testing with use of MSE and MAE metrics. The results suggest the limited effectiveness of linear regression for predicting the prices of cryptocurrencies due to the non-linear nature of the market. The Decision Tree model has shown some success in predicting bitcoin prices, however, deviations in forecasts have increased over time, especially in the face of market fluctuations. Polynomial regression and the Random Forest model demonstrate higher accuracy in forecasts. Based on the comparison of MSE and MAE indicators, the Random Forest turned out to be the most effective model for predicting bitcoin prices among those considered.

Keywords: cryptocurrencies, behavior patterns, API (application programming interface), social networks, machine learning, behavioral economics, forecasting models.

Fig.: 6. **Bibl.:** 15.

Guryanova Lidiya S. – D. Sc. (Economics), Professor, Professor of the Department of Economic Cybernetics and Applied Economics, V. N. Karazin Kharkiv National University (4 Svobody Square, Kharkiv, 61022, Ukraine)

E-mail: guryanovalidiya@gmail.com

ORCID: <https://orcid.org/0000-0002-2009-1451>

Researcher ID: <https://www.webofscience.com/wos/author/record/L-3402-2017>

Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=36068855600>

Lutsenko Rostyslav R. – Postgraduate Student of the Department of Economic Cybernetics and Applied Economics, V. N. Karazin Kharkiv National University (4 Svobody Square, Kharkiv, 61022, Ukraine)

E-mail: roxanisen@gmail.com

ORCID: <https://orcid.org/0000-0003-0737-3902>

Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=57238374000>

Сучасний ринок криптовалют характеризується високою волатильністю і швидкими змінами [1; 2], що призводить до неможливості застосування традиційних методів прогнозування вартості фінансових активів для визначення інвестиційної привабливості віртуальних активів. Зараз віртуальні активи все частіше розглядають як альтернативний вид інвестиційних активів, особливо в умовах зниження рівня ділової активності фондових ринків, несприятливої кон'юнктури на ринку нерухомості та ін. Зростаючий обсяг гібридних інвестиційних портфельів, які враховують як традиційні, так і альтернативні активи, зокрема криптовалюти, високий рівень волатильності ринку криптовалют потребують розробки нових теоретико-методичних підходів до побудови моделей прогнозування динаміки ринку криптовалют.

На динаміку ринку впливають соціальні та поведінкові фактори стейкхолдерів [3; 4], інформаційним джерелом дослідження яких є соціальні мережі [5; 6]. Такі поведінкові метрики, як кількість публікацій, активність користувачів соціальних мереж, частота ключових слів або фраз, індекс настроїв, часова динаміка настроїв, взаємозв'язки між користувачами соціальних мереж є важливими для прогнозування ринку віртуальних активів. Поведінкові фактори віртуальних активів є інди-

торами, які допомагають досліджувати, як психологічні та емоційні аспекти впливають на фінансові рішення учасників ринку. Такі детермінанти, як «ірраціональна впевненість», «аверсія до втрат», «ефект натовпу», «чутливість до новин», «страх упустити можливість» відображають ті поведінкові упередження, які впливають на фінансові рішення стейкхолдерів ринку віртуальних активів.

Ефективне управління інвестиціями у віртуальні активи потребує комплексної оцінки за допомогою методів інтелектуального аналізу даних, що дозволяють глибше зрозуміти динаміку ринку та поведінку інвесторів. Дослідження особливостей застосування таких методів, як лінійна регресія, множинна та поліноміальна регресія, ARIMA/SARIMA, а також алгоритми машинного навчання (дерева рішень, Random Forest, XGBoost) для аналізу динаміки ринку криптовалют, дозволить оцінити ефективність цих підходів, виявити зв'язки між активністю стейкхолдерів ринку в соціальних мережах і ціновими коливаннями криптовалют.

Отже, для побудови моделей аналізу динаміки ринку криптовалют слід враховувати не лише фінансові показники, а й поведінкові фактори та метрики за даними соціальних мереж, а також застосовувати методи інтелектуального аналізу даних.

Аналіз сучасних досягнень у сфері інтелектуального аналізу даних [5–11] підкреслює важливість його інтеграції в поведінкову економіку віртуальних активів. Публікація Zhang C., Li W., Zhang H., Zhan T. [8] містить огляд теоретичних основ методів інтелектуального аналізу даних і пропонує практичні рекомендації для підвищення точності та швидкості аналізу. У статті Могильної М. В. та Дубровіна В. І. [10] приділено увагу проблемам обробки великих даних, розглянуто технологію інтелектуального аналізу тексту, програмне забезпечення для текстового аналізу, яке використовується в машинному навчанні та науці про дані. У статті "Crypto-Currency Sentiment Analysis on Social Media" автори Erdoğan M. C., Canayaz M. аналізують зв'язки між настроями користувачів і змінами цін на криптовалюту. Вони доводять, що емоційні реакції та думки споживачів впливають на ринкові тенденції [5]. Виявлено, що використання різних моделей, таких як багатошарові перцептрони (MLP), опорні вектори (SVM) та випадкові ліси (RF), дозволяє ефективно передбачити зміни цін на криптовалюту. Дослідження підтвердило, що моделі машинного навчання є ефективними для короткострокового прогнозування цін на криптовалюту, з помірно середньою абсолютною відсотковою похибкою апроксимації (MAPE), близькою до 3,5% [7]. Додавання технічних індикаторів, таких як Simple Moving Average (SMA), Exponential Moving Average (EMA) та Relative Strength Index (RSI), до набору вхідних даних поліпшує точність моделей, допомагаючи відстежувати коротко- та довгострокові тенденції цін. У дослідженні [9] "Time-series forecasting of Bitcoin prices using high-dimensional features: a machine learning approach" автори прогнозують ціну на біткоїн за допомогою моделей машинного навчання, визначають точність прогнозування для короткострокових та середньострокових цін на біткоїн. Стаття Chaudhary D., Saroj S. K. [11] надає практичні рекомендації щодо застосування машинного навчання в прогнозуванні цін на криптовалюту. У публікації [12] приділено увагу прогнозуванню ринку криптовалют за допомогою методів машинного навчання та аналізу настроїв. Оцінювалися та порівнювалися три моделі прогнозування: багатошарові перцептрони (MLPs), опорні вектори (SVMs) і випадкові ліси (RFs) для Bitcoin, Ethereum, Ripple та Litecoin з використанням даних із соціальної мережі Twitter.

Дослідження у сфері поведінкової економіки віртуальних активів із застосуванням інтелектуальних систем аналізу даних активно висвітлюються науковцями. Так, у статті "Deep Learning-Based Cryptocurrency Price Prediction Scheme with Inter-Dependent Relations" [13] автори пропонують

схему прогнозування цін на Litecoin та Zcash із застосуванням алгоритмів GRU та LSTM. Зазначено, що при прогнозуванні цін на криптовалюту слід враховувати вплив соціальних і психологічних факторів. Традиційні моделі часових рядів, такі як VARIMA, ARIMA та GARCH, мають обмеження щодо точності прогнозування при нерівномірних даних. Використання алгоритмів машинного навчання, таких як SVM, випадкові ліси та KNN, показало позитивні результати у прогнозуванні цін на криптовалюту. У статті [14] автори порівнюють ефективність статистичних, машинних і глибоких навчальних моделей. Результати даного дослідження показують, що моделі глибокого навчання, зокрема LSTM, демонструють найкращі результати. У статті проводиться порівняння моделей глибокого навчання (DL), машинного навчання (ML) і статистичних моделей для прогнозування щоденних цін на криптовалюту. Результати показують, що рекурентні моделі глибокого навчання, LSTM виявилися найбільш ефективними. Автори зазначають, що створення більш точних прогнозів для торгових стратегій допоможе трейдерам приймати ефективні інвестиційні рішення.

Ряд досліджень акцентують увагу на важливості соціальних мереж у формуванні ринкових настроїв [1; 2; 5; 6]. Аналіз цих факторів дозволяє підвищити точність прогнозів на ринку віртуальних активів. У статті "Social Media Sentiment Analysis for Cryptocurrency Market Prediction" автори досліджують кореляцію між показниками настроїв і коливаннями цін на Bitcoin, аналізують методи розрахунку показників настроїв з тексту та оцінюють їх точність у прогнозуванні, аналізують зв'язки між показниками настроїв і змінами цін [1]. Автори акцентують увагу на причинно-наслідкових зв'язках між показниками настроїв та змінами цін. Результати показують, що аналіз настроїв є ефективним інструментом для прогнозування цін на криптовалюту.

Привертають увагу дослідників і моделі обробки природної мови для аналізу настроїв у соціальних медіа в контексті прогнозування фінансових ринків. У публікації Wołk K. [2] приділено увагу оцінці настроїв у соціальних медіа для прогнозування короткострокових цін на криптовалюту. Для аналізу настроїв застосовано багатомодельний підхід. У статті [6] проаналізовано кореляцію між ціною біткоїна та активністю користувачів у соціальних мережах, вивчено взаємозв'язок між ринковими настроями та ціновими коливаннями.

Методи експериментальної економіки також допомагають науковцям і практикам краще зрозуміти поведінку інвесторів. Так, у статті «До-

слідження поведінкових особливостей трейдерів: поєднання методів експериментальної економіки та технологій машинного навчання» Конової К. та Дека А. [3] з використанням таких методів, як «Гра на відгадування», виділяють кілька типів трейдерів: «жартівники», «прибуткові трейдери» та «професіонали». Дослідження підтверджує, що більш раціональні трейдери отримують вищі прибутки, тоді як ірраціональна поведінка може призводити до значних фінансових втрат. Дослідження "Long and Short-Term Impacts of Regulation in the Cryptocurrency Market" [4] оцінює вплив актуальних новин і подій на доходи учасників ринку криптовалют. Показано, що інвестори менш негативно реагували на найбільш неліквідні криптовалюти та на ті, які зазнали більшого ризику асиметрії інформації.

Аналіз сучасних публікацій свідчить про те, що, незважаючи на вагомий внесок науковців у вивчення динаміки ринку криптовалют, вплив поведінкових факторів стейкхолдерів ринку віртуальних активів, питання інтеграції метрик соціальних мереж у моделі прогнозування та аналізу залишається недостатньо дослідженими. Наявні моделі не завжди враховують специфіку поведінкової економіки, яка є важливим елементом для розуміння змін на ринку криптовалют. Це стало основою для формулювання мети даного дослідження.

Метою дослідження є проведення комплексного аналізу моделей динаміки ринку криптовалют з урахуванням поведінкових метрик стейкхолдерів за даними соціальних мереж, що впливають на ринкову активність. Відповідно до мети дослідження основними завданнями є:

- ✦ здійснити збір даних про активність користувачів у соціальних мережах та цінні дані біткоіна;
- ✦ виконати обробку й агрегацію даних із соціальних мереж;
- ✦ провести аналіз кореляції між активністю в соціальних мережах і ціною біткоіна;
- ✦ визначити, які метрики активності в соціальних мережах можуть служити показниками або детермінантами змін ціни біткоіна;
- ✦ побудувати моделі прогнозування на основі лінійної, поліноміальної регресії, а також методів машинного навчання (Random Forest, Decision Tree) для визначення динаміки ціни біткоіна;
- ✦ оцінити точність та ефективність моделей за допомогою таких метрик, як MSE (Mean Squared Error) та MAE (Mean Absolute Error).

Реалізація мети та завдань дозволить комплексно оцінити потенціал використання поведінкових моделей для аналізу динаміки ринку крипто-

валют, а також сформулювати рекомендації щодо вибору методів прогнозування на основі даних соціальних мереж.

Дані для даного дослідження отримані за допомогою АПІ соціальних мереж за розробленим алгоритмом моніторингу [15]. *Основні етапи методології:* збір даних, їх обробка й агрегація, масштабування та кореляційний аналіз, побудова та оцінка моделей. Отримано дані із соціальних мереж, такі як кількість лайків, коментарів, переглядів, поширень контенту. Для збору даних щодо ціни біткоіна використані бібліотеки `urfinance`, проведено обробку та агрегацію даних за днями для зручного поєднання їх з цінними даними. Далі здійснено масштабування для нормалізації даних до діапазону $[0, 1]$ з метою порівняння різних метрик; проведено кореляційний аналіз з метою оцінки взаємозв'язків між показниками активності (кількість лайків, поширень, коментарів) та ціною біткоіна. На завершальному етапі побудовано моделі лінійної, поліноміальної регресії, дерев рішень, випадкового лісу, а також інших алгоритмів машинного навчання з метою визначення залежності між соціальними показниками та ринковою ціною біткоіна. Оцінка моделей здійснена шляхом тестування за допомогою метрик MSE і MAE.

Виходячи з аналізу літератури, виявлено, що взаємозв'язок між активністю в соціальних мережах і ціною біткоіна є важливою, але недостатньо дослідженою темою, що вимагає більш глибокого вивчення. Наявні роботи вказують на потенційний вплив соціальних факторів на цінні коливання, проте чіткі закономірності, що поєднують ці фактори, не завжди виявлені або ж обмежені застосуванням лінійних моделей.

На першому кроці було проведено кореляційний аналіз змінних, зокрема ціни біткоіна та кількості публікацій у соціальних мережах. Для проведення кореляційного аналізу було обрано лише числові змінні, які можуть впливати на цільову змінну – ціну біткоіна. До змінних, що мають сенс для кореляційного аналізу, належать:

- ✦ *digg_count* – кількість лайків (інтерес до публікацій);
- ✦ *share_count* – кількість поширень (індикатор розповсюдження інформації);
- ✦ *comment_count* – кількість коментарів (рівень залученості аудиторії);
- ✦ *play_count* – кількість переглядів (індикатор популярності контенту);
- ✦ *created_time* – для синхронізації значень соціальних змінних з ціною біткоіна у часі;
- ✦ *Close* (ціна біткоіна) – цільова змінна для оцінки кореляції.

Змінні, які не були включені у кореляційний аналіз:

- ✦ id, text, lang, hashtags, author_id, author_username – текстові та категоріальні змінні, які не мають числової форми, що обмежує їхню застосовність для кореляції з числовими показниками;
- ✦ is_original_item, is_official_item, is_for_friend, is_duet_enabled, is_stitch_enabled, is_share_enabled, is_ad – булеві змінні, які мають значення True/False і, на відміну від кількісних показників, не здатні надати інформативну кореляцію з цільовою змінною.

Урахування лише релевантних числових змінних для кореляційного аналізу забезпечує точніший і більш ефективний аналіз взаємозв'язків між показниками активності в соціальних мережах та ціною біткоїна (рис. 1).

кореляцію 0,69 (популярні пости, які отримують більше лайків, частіше поширюються); comment_count має слабку кореляцію з іншими показниками (коментарі не завжди є показником популярності контенту порівняно з лайками та поширеннями).

Таким чином, соціальні показники TikTok мають сильні зв'язки між собою та слабкі зв'язки із ціною біткоїна.

На другому кроці дослідження було побудовано лінійну регресійну модель з урахуванням поведінкових метрик стейкхолдерів ринку криптовалют за даними соціальних мереж (рис. 2). Отримані дані було поділено на навчальні (80%) і тестові (20%) вибірки. Моделі «вчилися» на навчальних даних, щоб знайти зв'язок між незалежними змінними (активність у соцмережах) та залежною змінною (ціна біткоїна). Після того, як моделі були навчені, їх використовували для про-

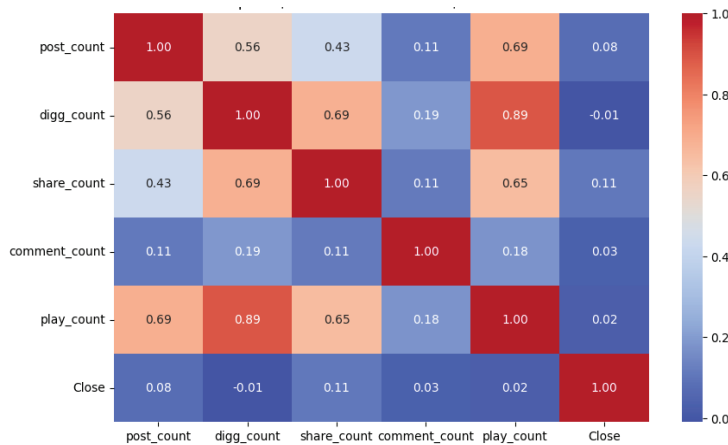


Рис. 1. Матриця кореляцій між метриками TikTok і ціною біткоїна

Джерело: авторська розробка.

Аналіз даних, наведених на рис. 1, дозволяє зробити висновок про те, що активність у соціальній мережі TikTok, така як кількість публікацій, лайків, поширень, коментарів і переглядів, не має значного впливу на ціну біткоїна або цей вплив є мінімальним.

За даними рис. 1 спостерігається сильна кореляція між соціальними показниками, такими як digg_count і play_count, що мають кореляцію 0,89 (більше лайків часто супроводжується більшою кількістю переглядів); digg_count і share_count, що мають кореляцію 0,69 (позитивний зв'язок між кількістю лайків і поширень); post_count і play_count, що мають кореляцію 0,69, (більше публікацій зазвичай супроводжується більшою кількістю переглядів).

Зв'язок між іншими соціальними показниками такий: share_count і digg_count мають помітну

гнозування ціни біткоїна на основі тестових даних. Для оцінки точності моделей використовували метрики MSE і MAE.

Показники коефіцієнта детермінації R^2 для всіх моделей дуже низькі (переважно близько 0), що вказує на слабку відповідність моделі лінійної регресії для прогнозування ціни біткоїна на основі кожного із соціальних показників TikTok (див. рис. 2). Отримані дані вказують на те, що лінійна модель майже не пояснює варіативність ціни біткоїна за допомогою цих показників. Середньоквадратичні помилки (MSE) є високими для всіх моделей, що додатково підтверджує низьку якість прогнозу. Проте для отримання глибшого розуміння й оцінки можливої комбінації показників доцільно розглянути підхід багатofакторної лінійної регресії (рис. 3). Такий підхід дозволяє врахувати сукупний вплив різних соціальних метрик і може виявити

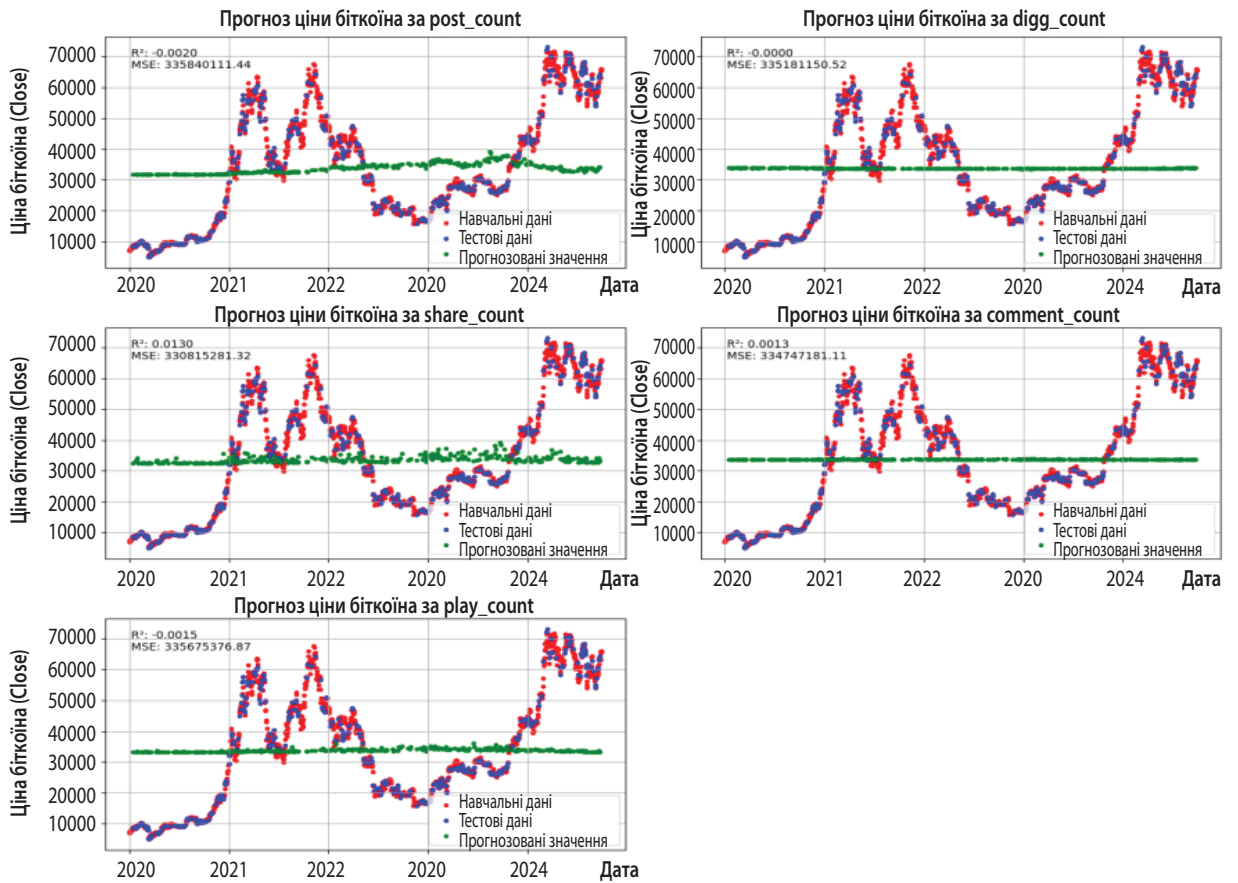


Рис. 2. Лінійні моделі прогнозування ціни біткоїна з урахуванням активності в соціальній мережі TikTok
 Джерело: авторська розробка.

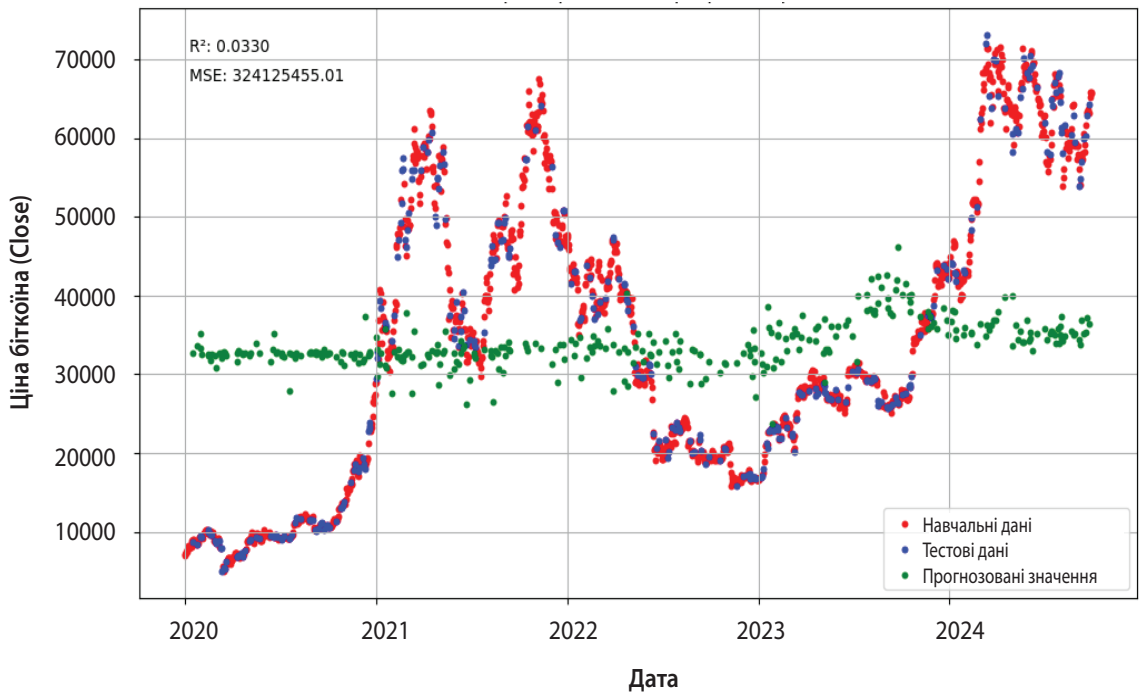


Рис. 3. Багатофакторна лінійна регресія для аналізу динаміки ціни біткоїна
 Джерело: авторська розробка.

потенційні взаємозв'язки, не помітні при розгляді кожного показника окремо.

Коефіцієнт детермінації R^2 дорівнює 0,0330, що означає, що модель пояснює лише близько 3,3% варіативності в ціні біткоїна за допомогою таких соціальних показників TikTok, як `post_count`, `digg_count`, `share_count`, `comment_count` і `play_count`. Середньоквадратична помилка (MSE) становить 324125455,01, що також свідчить про суттєві відхилення між реальними та прогнозованими значеннями. Прогнозовані значення показують, що модель здебільшого дозволяє прогнозувати середню ціну біткоїна (близько 30,000–35,000 USD), не відображаючи волатильності ринку, яку демонструють реальні дані.

Таким чином, результати прогнозування за допомогою багатофакторної лінійної регресії свідчать про обмеження цього підходу при аналізі волатильного ринку криптовалют, зокрема в умовах, коли соціальні показники TikTok мають слабку лінійну кореляцію з ціною біткоїна.

На третьому кроці дослідження було побудовано поліноміальну модель (рис. 4), яка здатна краще врахувати можливі нелінійні взаємозв'язки між показниками активності в соціальних мережах та ціною біткоїна.

Для побудови моделі поліноміальної регресії (див. рис. 4) було використано кількість публікацій `post_count`, але додаємо до моделі ще `post_count^2` і `post_count^3`. Коефіцієнт детермінації $R^2 = 0,4643$,

що значно краще порівняно з лінійною регресією. R^2 показує, що поліноміальна модель пояснює близько 46,43% варіативності в ціні біткоїна. Середньоквадратична помилка (MSE) дорівнює 179561942,45. Значення MSE стало меншим порівняно з лінійною регресією, що вказує на те, що поліноміальна модель краще відображає тенденції в даних. Прогнозовані значення знаходяться ближче до реальних тестових даних, ніж у випадку лінійної регресії, особливо в періоді з 2021 по 2022 рр. Це показує, що поліноміальна регресія краще «вловлює» коливання ціни біткоїна. Однак є періоди (наприклад, на початку 2021 р.), де модель все ще погано відображає стрибки цін.

Четвертим кроком було застосування методу дерева рішень, який здатен адаптуватися до структурних змін у даних і враховувати нелінійні взаємозв'язки між різними факторами. На графіку (рис. 5) представлено результати прогнозування цін на біткоїн за допомогою моделі дерева рішень. Синя лінія представляє згладжені фактичні ціни біткоїна, тоді як червона показує прогнозовані ціни. Обидві лінії демонструють схожі тренди, що вказує на те, що модель змогла «вловити» загальний напрямок руху цін біткоїна. У період з 2020 по 2021 рр. моделі вдалося точно передбачити тенденцію зростання цін на біткоїн. З 2022 р. прогнози стали менш точними, і прогнозовані ціни показують більші коливання порівняно з реальними.

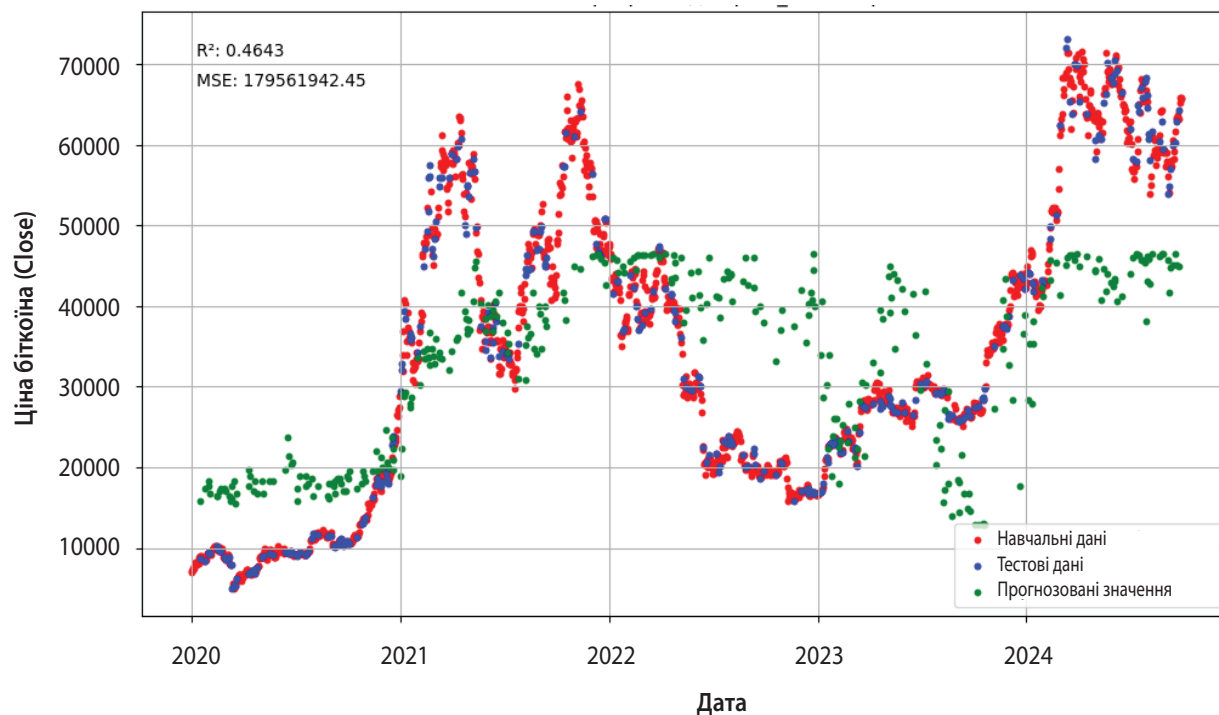


Рис. 4. Поліноміальна регресія для аналізу динаміки ціни біткоїна

Джерело: авторська розробка.

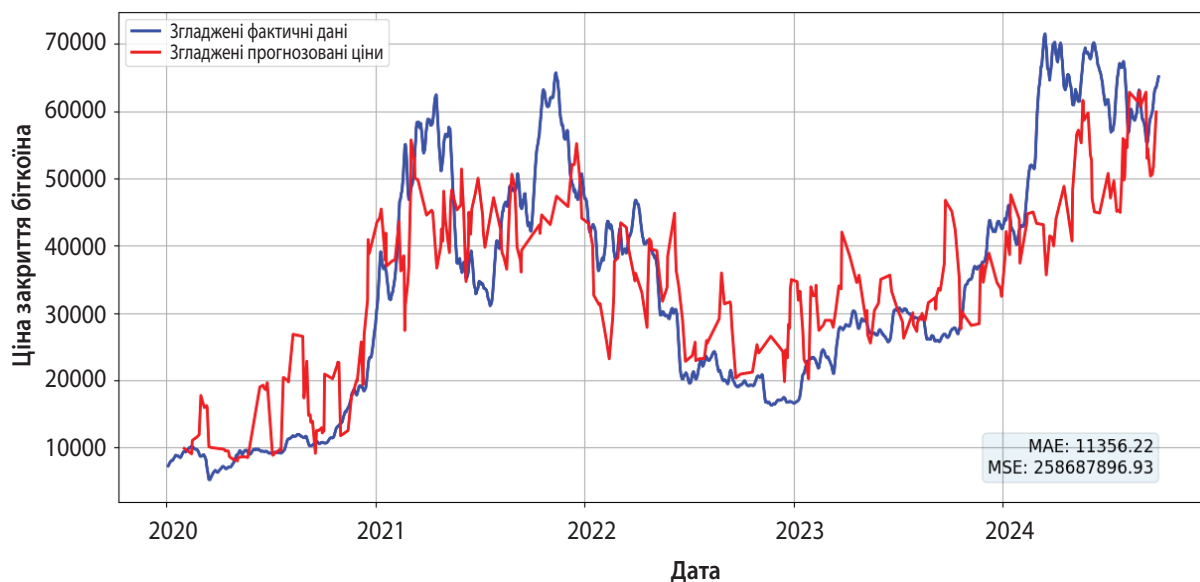


Рис. 5. Дерево рішень для аналізу динаміки ціни біткоїна

Джерело: авторська розробка.

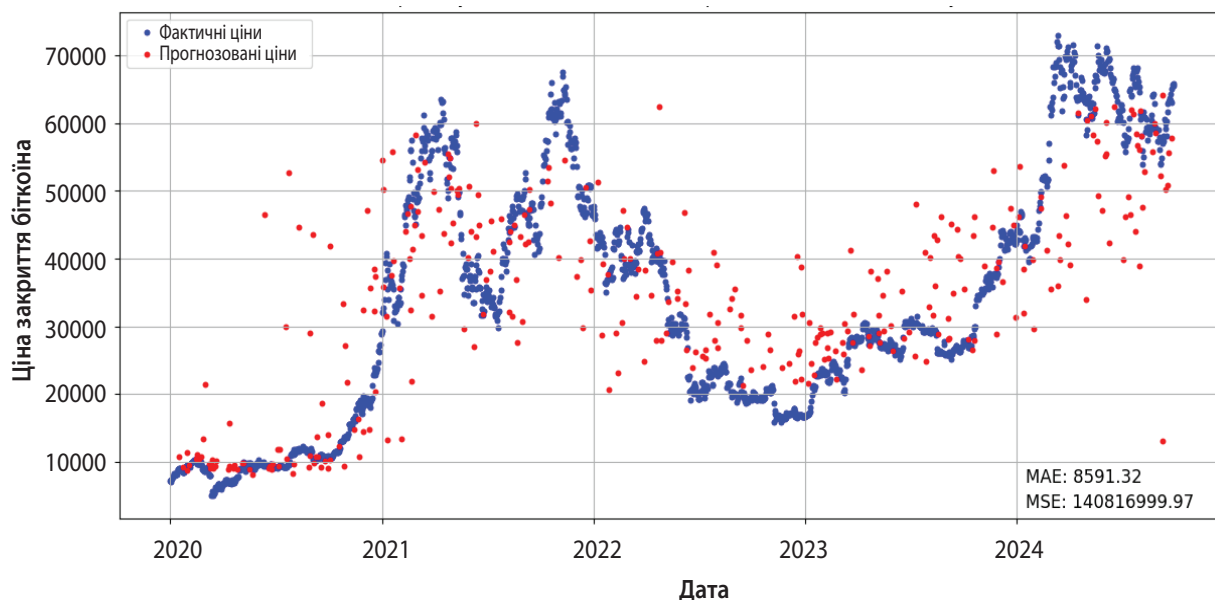


Рис. 6. Випадковий ліс для аналізу динаміки ціни біткоїна

Джерело: авторська розробка.

Побудована модель дерева рішень продемонструвала певний успіх у прогнозуванні цін на біткоїн, проте зросли відхилення у прогнозах з часом. Тому на п'ятому кроці дослідження розроблено модель випадкового лісу, яка може забезпечити кращу стабільність і точність прогнозів завдяки своїй ансамблевій структурі та здатності враховувати більш складні взаємозв'язки між змінними (рис. 6).

На графіку (див. рис. 6) представлено фактичні та прогнозовані ціни біткоїна за період з 2020 по 2024 рр. Фактичні ціни показують реальний розви-

ток ринку, тоді як прогнозовані ціни відображають результати моделі випадкового лісу. Загальна тенденція фактичних цін біткоїна демонструє періодичні коливання, з помітним зростанням у 2021 р. і зниженням у 2022–2023 рр. Прогнозовані ціни демонструють певні коливання, які збігаються з тенденціями фактичних цін, але прогнозовані ціни відхиляються від фактичних. На основі порівняння показників MSE і MAE випадковий ліс виявився найефективнішою моделлю для прогнозування цін біткоїна серед розглянутих.

ВИСНОВКИ

Результати проведених досліджень вказують на необхідність урахування поведінкових метрик і факторів стейкхолдерів ринку криптовалюти при прогнозуванні цін віртуальних активів. Проведений комплексний аналіз моделей динаміки ринку криптовалюти вказав на слабку лінійну кореляцію між активністю в соціальних мережах та змінами ціни біткоїна. Лінійна регресія виявилася мало-ефективною для прогнозування ціни біткоїна, оскільки ринкові зміни криптовалюти мають нелінійну природу. Поліноміальна регресія показала кращі результати порівняно з лінійною моделлю, дозволяючи враховувати нелінійні зв'язки між змінними. Моделі дерева рішень та випадковий ліс продемонстрували вищу точність у прогнозуванні динаміки ціни біткоїна.

Подальші дослідження будуть спрямовані на застосування додаткових індикаторів соціальних мереж з метою оцінки впливу поведінкових детермінант на динаміку ринку віртуальних активів. ■

БІБЛІОГРАФІЯ

1. Raheman A., Kolonin A., Fridkins I. et al. Social Media Sentiment Analysis for Cryptocurrency Market Prediction. 2022. *arXiv preprint*. arXiv:2204.10185. DOI: <https://doi.org/10.48550/arXiv.2204.10185>
2. Wołk K. Advanced social media sentiment analysis for short-term cryptocurrency price prediction. *Expert Systems*. 2019. Vol. 37. Iss. 2. DOI: <https://doi.org/10.1111/exsy.12493>
3. Кононова К. Ю., Дек А. О. Дослідження поведінкових особливостей трейдерів: поєднання методів експериментальної економіки та технологій машинного навчання. *Нейро-нечіткі технології моделювання в економіці*. 2018. № 7. С. 148–167. DOI: <https://doi.org/10.33111/nfmte.2018.148>
4. Chokor A., Alfieri E. Long and Short-Term Impacts of Regulation in the Cryptocurrency Market. *The Quarterly Review of Economics and Finance*. 2021. Vol. 81. P. 157–173. DOI: <https://doi.org/10.1016/j.qref.2021.05.005>
5. Erdoğan M. C., Canayaz M. Crypto-Currency Sentiment Analysis on Social Media. In *Proceedings of the International Conference on Data Analysis and Processing (IDAP)*. Malatya, Turkey. 28–30 September 2018. DOI: <https://doi.org/10.1109/IDAP.2018.8620724>
6. Guidi B., Michienzi A. Bitcoin Price Variation: An Analysis of the Correlations. In: *Euro-Par 2019 Workshops*. Springer, 2020. P. 429–441. DOI: https://doi.org/10.1007/978-3-030-48340-1_33
7. Derbentsev V., Matviychuk A., Soloviev V. Forecasting of Cryptocurrency Prices Using Machine Learning. In: *Advanced Studies of Financial Technologies and Cryptocurrency Markets*. Springer. 2020. P. 211–231. DOI: https://doi.org/10.1007/978-981-15-4498-9_12

8. Zhang C., Li W., Zhang H., Zhan T. Recent Advances in Intelligent Data Analysis and Its Applications. *Electronics*. 2024. Vol. 13. No. 1. Art. 226. DOI: <https://doi.org/10.3390/electronics13010226>
9. Mudassir M., Bennbaia S., Unal D., Hammoudeh M. Time-series forecasting of Bitcoin prices using high-dimensional features: a machine learning approach. *Neural Computing and Applications*. 2020. Vol. 32. Iss. 20. P. 15869–15888. DOI: <https://doi.org/10.1007/s00521-020-05129-6>
10. Могильна М. В., Дубровін В. І. Інтелектуальний аналіз тексту: застосування та безкоштовні програмні засоби. *Прикладні питання математичного моделювання*. 2022. Т. 5. № 2. С.41–49. DOI: <https://doi.org/10.32782/mathematical-modelling/2022-5-2-5>
11. Chaudhary D., Saroj S. K. Cryptocurrency Price Prediction Using Machine Learning Algorithms. *Advances in Distributed Computing and Artificial Intelligence Journal*. 2023. Vol. 12. No. 1. DOI: <https://doi.org/10.14201/adcaij.31490>
12. Valencia F., Gomez-Espinosa A., Valdes-Aguirre B. Cryptocurrency Price Movement Prediction Using Sentiment Analysis and Machine Learning. *Entropy*. 2019. Vol. 21. Iss. 6. Art. 589. DOI: <https://doi.org/10.3390/e21060589>
13. Tanwar S., Patel N. P., Patel S. N. et al. Deep Learning-Based Cryptocurrency Price Prediction Scheme with Inter-Dependent Relations. *IEEE Access*. 2021. Vol. 9. P. 138633–138646. DOI: <https://doi.org/10.1109/ACCESS.2021.3117848>
14. Murray K., Rossi A., Carraro D., Visentin A. Cryptocurrency Price Prediction: A Comparison of Machine Learning, Deep Learning, and Ensemble Methods. *Forecasting*. 2023. Vol. 5. Iss. 1. P. 196–209. DOI: <https://doi.org/10.3390/forecast5010010>
15. Луценко Р. Р. Прикладний програмний інтерфейс як метод моніторингу даних соціальних мереж для досліджень у поведінковій економіці. *Бізнес Інформ*. 2024. № 8. С. 133–141. DOI: <https://doi.org/10.32983/2222-4459-2024-8-133-141>

REFERENCES

- Chaudhary, D., and Saroj, S. K. "Cryptocurrency Price Prediction Using Machine Learning Algorithms". *Advances in Distributed Computing and Artificial Intelligence Journal*, vol. 12, no. 1 (2023). DOI: <https://doi.org/10.14201/adcaij.31490>
- Chokor, A., and Alfieri, E. "Long and Short-Term Impacts of Regulation in the Cryptocurrency Market". *The Quarterly Review of Economics and Finance*, vol. 81 (2021): 157-173. DOI: <https://doi.org/10.1016/j.qref.2021.05.005>
- Derbentsev, V., Matviychuk, A., and Soloviev, V. "Forecasting of Cryptocurrency Prices Using Machine Learning". In *Advanced Studies of Financial Technologies and Cryptocurrency Markets*, 211-231. Springer, 2020. DOI: https://doi.org/10.1007/978-981-15-4498-9_12

- Erdogan, M. C., and Canayaz, M. "Crypto-Currency Sentiment Analysis on Social Media". *Proceedings of the International Conference on Data Analysis and Processing (IDAP)*. Malatya, Turkey, 2018.
DOI: <https://doi.org/10.1109/IDAP.2018.8620724>
- Guidi, B., and Michienzi, A. "Bitcoin Price Variation: An Analysis of the Correlations". In *Euro-Par 2019 Workshops*, 429-441. Springer, 2020.
DOI: https://doi.org/10.1007/978-3-030-48340-1_33
- Kononova, K. Yu., and Dek, A. O. "Doslidzhennia povedinkovykh osoblyvostei treideriv: poiednannia metodiv eksperymentalnoi ekonomiky ta tekhnolohii mashynnoho navchannia" [Investigation of Traders' Behavioral Characteristics: Experimental Economics Methods and Machine Learning Technologies]. *Neiro-nechitki tekhnolohii modeliuвання v ekonomitsi*, no. 7 (2018): 148-167.
DOI: <https://doi.org/10.33111/nfmte.2018.148>
- Lutsenko, R. R. "Prykladnyi prohramnyi interfeis yak metod monitorynhu danykh sotsialnykh merezh dlia doslidzhen u povedinkovii ekonomitsi" [The Application Programming Interface as a Method for Monitoring Social Network Data for Research in Behavioral Economics]. *Biznes Inform*, no. 8 (2024): 133-141.
DOI: <https://doi.org/10.32983/2222-4459-2024-8-133-141>
- Mohylna, M. V., and Dubrovin, V. I. "Intelektualnyi analiz tekstu: zastosuvanniata bezkoshtovni prohramni zasoby" [Text Mining: Applications and Free Software Tools]. *Prykladni pytannia matematychnoho modeliuвання*, vol. 5, no. 2 (2022): 41-49.
DOI: <https://doi.org/10.32782/mathematical-modelling/2022-5-2-5>
- Mudassir, M. et al. "Time-series forecasting of Bitcoin prices using high-dimensional features: a machine learning approach". *Neural Computing and Applications*, vol. 32, no. 20 (2020): 15869-15888.
DOI: <https://doi.org/10.1007/s00521-020-05129-6>
- Murray, K. et al. "Cryptocurrency Price Prediction: A Comparison of Machine Learning, Deep Learning, and Ensemble Methods". *Forecasting*, vol. 5, no. 1 (2023): 196-209.
DOI: <https://doi.org/10.3390/forecast5010010>
- Raheman, A. et al. *Social Media Sentiment Analysis for Cryptocurrency Market Prediction*. arXiv preprint. arXiv:2204.10185, 2022.
DOI: <https://doi.org/10.48550/arXiv.2204.10185>
- Tanwar, S. et al. "Deep Learning-Based Cryptocurrency Price Prediction Scheme with Inter-Dependent Relations". *IEEE Access*, vol. 9 (2021): 138633-138646.
DOI: <https://doi.org/10.1109/ACCESS.2021.3117848>
- Valencia, F., Gomez-Espinosa, A., and Valdes-Aguirre, B. "Cryptocurrency Price Movement Prediction Using Sentiment Analysis and Machine Learning". *Entropy*, art. 589, vol. 21, no. 6 (2019).
DOI: <https://doi.org/10.3390/e21060589>
- Wolk, K. "Advanced social media sentiment analysis for short-term cryptocurrency price prediction". *Expert Systems*, vol. 37, no. 2 (2019).
DOI: <https://doi.org/10.1111/exsy.12493>
- Zhang, C. et al. "Recent Advances in Intelligent Data Analysis and Its Applications". *Electronics*, art. 226, vol. 13, no. 1 (2024).
DOI: <https://doi.org/10.3390/electronics13010226>