

УДК 336.76:519.7:004.8

JEL C22, C53, G11, G17

DOI: <https://doi.org/10.32782/2617-5940.1.2026.20>**Світлана Науменкова**

доктор економічних наук, професор,
професор кафедри фінансів,
Київський національний університет імені Тараса Шевченка
ORCID: <https://orcid.org/0000-0001-8582-6044>
E-mail: naumenkova@knu.ua

Світлана Міщенко

доктор економічних наук, професор,
професор кафедри фінансів,
Державний торговельно-економічний університет
ORCID: <https://orcid.org/0000-0002-1840-8579>
E-mail: s.mishchenko@knute.edu.ua

Наталія Ярич

магістр,
Київський національний університет імені Тараса Шевченка
ORCID: <https://orcid.org/0009-0001-6747-7124>
E-mail: nataliia.yarych@knu.ua

ПРОГНОЗУВАННЯ ДИНАМІКИ ЦІН АКЦІЙ НА ФІНАНСОВОМУ РИНКУ З ВИКОРИСТАННЯМ МАШИННОГО НАВЧАННЯ ТА ГІБРИДНИХ МОДЕЛЕЙ

Анотація. В умовах стрімкого зростання обсягів фінансової інформації та активного впровадження технологій штучного інтелекту прогнозування динаміки цін фінансових активів набуває особливої актуальності для прийняття обґрунтованих інвестиційних рішень за умов підвищеної волатильності та можливого прояву шоків на фінансових ринках. Це зумовлює зростання інтересу до застосування методів машинного навчання, здатних виявляти складні нелінійні залежності та швидко адаптуватися до змін нестационарних часових рядів. Метою статті є крос-секторальний емпіричний аналіз прогностичної ефективності економетричних моделей і алгоритмів машинного навчання порівняно з класичними статистичними підходами, а також оцінювання гібридної моделі ARIMA–XGBoost для прогнозування напряму руху цін акцій. Об'єктом дослідження є фінансові часові ряди шести транснаціональних корпорацій різних секторів економіки з урахуванням макроекономічних індикаторів (S&P 500, VIX, NASDAQ), що відображають ключові інформаційні канали взаємодії на фінансових ринках. Для забезпечення коректності та відтворюваності результатів застосовано метод валідації з розширюваним навчальним вікном (Expanding Window Walk-Forward Validation), який враховує зміну ринкових умов у часі. Дослідження ґрунтується на використанні моделей експоненційного згладжування (ETS), ARIMA, GARCH та алгоритмів градієнтного бустингу (XGBoost, LightGBM). Запропонована двокомпонентна гібридна модель ARIMA–XGBoost поєднує декомпозицію часового ряду та навчання на залишках, забезпечуючи інтеграцію лінійних і нелінійних ефектів. Емпіричні результати свідчать, що гібридний підхід демонструє вищу прогностичну ефективність і більшу стійкість до ринкових шоків порівняно з окремими економетричними та машинними моделями. Практична цінність отриманих результатів полягає у можливості їх використання для формування ефективної політики управління ризиками, підвищення результативності управління портфелями фінансових активів, а також впровадження інструментів штучного інтелекту в умовах цифровізації фінансових ринків.

Ключові слова: фінансовий ринок; прогнозування цін акцій; машинне навчання; гібридні моделі; управління інвестиціями; ризик-менеджмент; фінансові часові ряди

Svitlana Naumenkova

Doctor of Economics, Professor, Department of Finance,
Taras Shevchenko National University of Kyiv
ORCID: <https://orcid.org/0000-0001-8582-6044>
E-mail: naumenkova@knu.ua

Svitlana Mishchenko

Doctor of Economics, Professor, Department of Finance,
State University of Trade and Economics
ORCID: <https://orcid.org/0000-0002-1840-8579>
E-mail: s.mishchenko@knute.edu.ua

Nataliia Yarych

Master Student,
Taras Shevchenko National University of Kyiv
ORCID: <https://orcid.org/0009-0001-6747-7124>
E-mail: nataliia.yarych@knu.ua



PREDICTING STOCK PRICE DIRECTION IN FINANCIAL MARKETS USING MACHINE LEARNING AND HYBRID MODELS

Abstract. In the context of the rapid growth of financial information and the active adoption of artificial intelligence technologies, forecasting financial asset price dynamics is becoming increasingly important for making informed investment decisions under conditions of heightened volatility and potential market shocks. This drives growing interest in the application of machine learning methods capable of identifying complex nonlinear relationships and rapidly adapting to changes in non-stationary time series. The aim of the article is to conduct a cross-sectoral empirical analysis of the predictive performance of econometric models and machine learning algorithms in comparison with classical statistical approaches, as well as to evaluate a hybrid ARIMA–XGBoost model for predicting stock price direction. The object of the study is financial time series of six multinational corporations from different sectors of the economy, taking into account macroeconomic indicators (S&P 500, VIX, NASDAQ) that reflect key information interaction channels in financial markets. To ensure the correctness and reproducibility of the results, an expanding window walk-forward validation approach is applied, which accounts for changes in market conditions over time. The study is based on the use of exponential smoothing models (ETS), ARIMA, GARCH, and gradient boosting algorithms (XGBoost, LightGBM). The proposed two-component hybrid ARIMA–XGBoost model combines time series decomposition with residual learning, ensuring the integration of linear and nonlinear effects. The empirical results indicate that the hybrid approach demonstrates higher predictive performance and greater robustness to market shocks compared to standalone econometric and machine learning models, demonstrating improved stability across different market conditions. The practical value of the results lies in their applicability for developing effective risk management policies, improving portfolio management efficiency, and implementing artificial intelligence tools in the context of financial market digitalization.

Keywords: financial market; stock price prediction; machine learning; hybrid models; investment management; risk management; financial time series.

Вступ. У сучасних умовах алгоритмізації фінансових ринків та стрімкого зростання обсягів доступних даних прогнозування динаміки цін акцій набуває ключового значення для інвесторів, портфельних менеджерів і фінансових аналітиків. Традиційна парадигма прогнозування абсолютних цінових значень стикається з проблемою нестационарності фінансових часових рядів, які характеризуються високим рівнем шуму, наявністю структурних зламів та асиметрією розподілів. У зв'язку з цим сучасний кількісний аналіз зміщує фокус із прогнозування рівня ціни на задачу бінарної класифікації – визначення напрямку руху активу (доходності).

Застосування виключно класичних статистичних методів, таких як ARIMA чи експоненційне згладжування, є обмеженим, оскільки вони ґрунтуються на припущенні лінійної автокореляційної структури та не враховують впливу екзогенних макроекономічних факторів. Натомість алгоритми машинного навчання здатні виявляти складні нелінійні залежності між ціною акції та ринковими індикаторами, однак їх ефективність суттєво варіює залежно від сектору економіки та ринкового режиму.

Недостатньо дослідженими залишаються питання крос-секторальної стабільності моделей, їх чутливості до зміни часових горизонтів прогнозування, а також ефективності інтеграції економетричних підходів і методів штучного інтелекту в межах гібридних (ансамблевих) архітектур.

Це зумовлює необхідність емпіричного порівняння прогностичної ефективності багатовимірних моделей машинного навчання та класичних статистичних підходів із застосуванням методології динамічного перенавчання, що дозволяє мінімізувати ризик витоку даних та наблизити експериментальні умови до реальних ринкових сценаріїв.

Літературний огляд. Активне впровадження технологій штучного інтелекту зумовлює потребу в переосмисленні концептуальних засад і практичних механізмів функціонування сучасних фінансових ринків, зокрема, в контексті взаємодії централізованих і

децентралізованих систем, що підтверджується сучасними дослідженнями [1; 2; 3, с. 117; 4, с. 33]. Штучний інтелект виступає ключовим драйвером інновацій у сфері оброблення даних та управління інформаційними потоками, у тому числі у фінансовому секторі, що підкреслюється у звіті Ради з фінансової стабільності [5, с. 14–16].

Прогнозування тенденцій і напрямку руху ціни акцій спрямоване на підвищення обґрунтованості прийняття інвестиційних рішень. У фундаментальних дослідженнях Ф. Фабозці [6] та А. Сандерс, М. Корнетт [7], підкреслюється, що класичні моделі забезпечують високий рівень інтерпретованості, проте не завжди здатні адекватно відображати складні нелінійні взаємозв'язки. Удосконалення портфельного аналізу зумовило необхідність більш точного пояснення аномалій ціноутворення фінансових активів під впливом різноманітних факторів [8, с. 450–455; 9, с. 149–150].

Традиційні економетричні підходи, зокрема, моделі ARIMA та GARCH, залишаються базовими інструментами аналізу лінійної структури та волатильності фінансових часових рядів, що відображено у роботі Р. Гайндмана, Г. Атанасопулоса [10], однак їх ефективність обмежується за умов високої нестационарності та наявності нелінійних залежностей.

Зростання складності фінансових даних та рівня шуму зумовлює активне використання методів машинного навчання (ML) та глибокого навчання (DL), а також розвиток гібридних підходів. У сучасних дослідженнях широко застосовуються алгоритми опорних векторів (SVM), випадкового лісу (RF), k-найближчих сусідів (KNN), а також ансамблеві методи для виявлення складних нелінійних закономірностей у фінансових часових рядах, що підтверджено у роботах С. Свами, С. Раджголі, Т. Хегде [11, с. 268–269], З. Го [12, с. 105–108], А. Кайит, М. Ісмаїл [13, с. 10–13, 19], І. Нті, Ф. Адекоя, Б. Вейорі [14, с. 15–17] та З. Тан, З. Янь, Г. Чжу [15].

Окрему увагу приділено методам глибокого навчання, зокрема рекурентним нейронним мережам, таким як LSTM та GRU, які демонструють високу

ефективність у задачах прогнозування часових рядів. Зокрема, результати досліджень Ц. Чен [16, с. 213] та І. Се [17, с. 743–744] свідчать про переваги таких моделей над традиційними підходами за метриками RMSE, MAE та MAPE.

Важливим напрямом є використання алгоритмів градієнтного бустингу, зокрема XGBoost, переваги якого обґрунтовано у роботі Т. Чен, К. Гуестрін [18, с. 786–790], де підкреслюється висока точність і стійкість моделі до шуму в задачах класифікації фінансових даних.

З метою подолання обмежень окремих підходів активно розвиваються гібридні моделі, що поєднують економічні та машинні методи. Зокрема, поєднання моделей GARCH із нейронними мережами (LSTM, GRU, Transformer) дозволяє одночасно враховувати волатильність і нелінійні залежності, що підтверджено у дослідженнях Дж. Мутінда, А. Лангат [19, с. 10–12] та Х. Хан, З. Лю, М. Барріос та ін. [20, с. 10–13]. Подібні підходи передбачають побудову багатокомпонентних гібридних моделей, що поєднують методи декомпозиції сигналів (EMD), процедури попередньої обробки (зокрема, інтерполяцію Akima) та нейронні мережі (LSTM), що дозволяє підвищити точність прогнозування в умовах високої волатильності, що доведено у роботі М. Лі, Д. Хан, Х. Альшанбарі, А. Ель-Багурі [21, с. 7–10].

Окремим сучасним напрямом є застосування методів машинного навчання на основі аналізу настроїв. Використання структурного тематичного моделювання дозволяє виокремити п'ять ключових кластерів досліджень: прогнозування на основі шаблонів, торгівлю на основі правил, алгоритмічну торгівлю, техно-фундаментальну торгівлю, а також машинне навчання та аналіз настроїв. При цьому домінуючими є напрями торгівлі на основі правил (26,79%), машинного навчання та аналізу настроїв (23,52%) і алгоритмічної торгівлі (21,11%), що продемонстровано у дослідженні С. Інані, Х. Прадхан, С. Кумар, Б. Бісвас [22, с. 275].

Перспективним напрямом подальших досліджень є інтеграція великих мовних моделей (LLM) у процесі аналізу фінансових ринків, зокрема, розширення контекстних вікон (context windows), що виконують функцію короткотривалої пам'яті, дозволяє моделям обробляти значні обсяги інформації та враховувати складні взаємозв'язки у фінансових даних. Це відкриває нові можливості для поєднання методів машинного навчання з інструментами обробки природної мови та цифровими фінансовими технологіями, включаючи блокчейн [23, с.56].

Попри значний науковий прогрес, у літературі зберігається низка невіршених питань, а саме: недостатня увага до крос-секторальної стійкості моделей; відсутність уніфікованих підходів до використання walk-forward validation у порівняльних дослідженнях; недостатня інтеграція економічних і машинних методів у межах єдиних гібридних архітектур.

Таким чином, аналіз сучасних досліджень свідчить про відсутність комплексного підходу, який би поєднував класичні економічні моделі (ARIMA, GARCH), алгоритми машинного навчання (XGBoost, LightGBM) та гібридні архітектури в межах єдиної крос-секторальної експериментальної методології. Це формує наукову нішу для подальших досліджень.

Метою статті є крос-секторальний емпіричний аналіз прогностичної точності лінійних статистичних моделей, методів машинного навчання та запропонованої гібридної моделі ARIMA–XGBoost для прогнозування напрямку руху цін акцій із застосуванням методології динамічної валідації в умовах різних ринкових режимів.

Методологія. Для забезпечення комплексності дослідження та перевірки гіпотези про перевагу нелінійних алгоритмів над лінійними, прогностичний інструментарій було сформовано за наступним принципом: від класичної економіки до сучасних ансамблевих методів машинного навчання та їх гібридизації.

З огляду на нестационарність та нелінійність фінансових часових рядів, задачу прогнозування сформульовано як задачу бінарної класифікації – передбачення напрямку руху цін акцій, що дозволяє підвищити стабільність оцінювання моделей у порівнянні з прогнозуванням абсолютних значень.

У дослідженні використано три групи моделей: класичні статистичні та економічні моделі (ETS, ARIMA, GARCH), алгоритми машинного навчання (XGBoost, LightGBM), гібридну модель ARIMA–XGBoost, що поєднує лінійне моделювання та нелінійне навчання на залишках.

Емпіричною базою дослідження є фінансові часові ряди шести транснаціональних компаній різних секторів економіки, сформовані на основі скоригованих цін закриття акцій (Adjusted Close) за період 2019–2025 рр. Дані агреговано до тижневої частоти з метою зменшення впливу мікроструктурного шуму. До моделі включено технічні індикатори (RSI, ковзні середні), а також макроекономічні змінні (S&P 500, VIX, NASDAQ), що відображають загальноринкові умови. Усі пояснювальні змінні сформовано із лагом в один період, що забезпечує відсутність витoku майбутньої інформації.

Для забезпечення коректності порівняння застосовано метод валідації з розширюваним навчальним вікном (Expanding Window Walk-Forward Validation), що дозволяє враховувати зміну ринкових умов у часі та забезпечує відтворюваність результатів. Оцінювання якості моделей здійснюється за метрикою частки правильних прогнозів (Hit Rate), що є релевантною для задач класифікації напрямку руху цін.

Основна частина. Оскільки фінансові часові ряди характеризуються нелінійністю, структурними зламами та змінною волатильністю, класичні підходи до прогнозування абсолютних цін значень часто демонструють обмежену прогностичну здатність через нестационарність даних.

В процесі дослідження класичні статистичні та економічні моделі (ETS, ARIMA, GARCH) постають як базова лінія порівняння, ґрунтуються на аналізі внутрішньої структури часового ряду без явного врахування зовнішніх факторів, що відповідає традиційному підходу до моделювання фінансових часових рядів.

Експоненційне згладжування (ETS) надає більшу вагу останнім спостереженням, що дозволяє моделювати короткострокові тенденції та може бути інтерпретоване як індикатор короткострокового імпульсу (моментуму) ціни, однак є чутливим до змін ринкового режиму.

Базова модель простого експоненційного представлена у вигляді:

$$y_t = \alpha y_{t-1} + (1 - \alpha) \cdot \hat{y}_{t-1} \quad (1)$$

де \hat{y}_t – прогноз на момент часу t ,

y_{t-1} – фактичне значення часового ряду в момент часу $t - 1$,

\hat{y}_{t-1} – попередній прогноз,

$\alpha \in (0, 1)$ – параметр згладжування, що визначає вагу нової інформації.

У дослідженні використано спрощену форму моделі без урахування трендової та сезонної компоненти.

Авторегресійна модель інтегрованого ковзного середнього (ARIMA) описує часовий ряд як лінійну комбінацію його попередніх значень та випадкових збурень і представлена у вигляді:

$$\varphi(B)(1 - B)^d y_t = \theta(B)\varepsilon_t \quad (2)$$

де B – оператор зсуву, ($B y_t = y_{t-1}$),

$\varphi(B)$ – поліном авторегресійної частини,

та $\theta(B)$ – поліном ковзного середнього,

d – порядок диференціювання,

ε_t – випадкова похибка (білий шум).

Модель ефективно відображає лінійну автокореляційну структуру, однак обмежена у здатності враховувати нелінійні ефекти та екзогенні шоки.

Модель умовної гетероскедастичності (GARCH) застосовується для моделювання часово-змінної дисперсії фінансових рядів. З огляду на прояв ефекту «кластеризації волатильності», коли періоди відносної стабільності змінюються фазами турбулентності. Застосування GARCH дозволило враховувати динамічну зміну волатильності у часі:

$$\sigma_t^2 = \omega + \alpha \varepsilon_{t-1}^2 + \beta \sigma_{t-1}^2 \quad (3)$$

де σ_t^2 – умовна дисперсія в момент часу t ,

ε_{t-1} – попередній шок (помилка),

ω – константа,

α – параметр реакції на нову інформацію,

β – параметр інерційності волатильності.

Урахування змінної волатильності є важливим для аналізу кризових періодів, зокрема, при тестуванні моделей на активах енергетичного та промислового секторів, які демонструють підвищену чутливість до макроекономічних шоків.

Ансамблеві алгоритми машинного навчання (*LightGBM*, *XGBoost*) є важливою складовою експере-

ментального дослідження. На відміну від традиційних нейронних мереж, зокрема LSTM, які є схильними до перенавчання на зашумлених фінансових рядах середнього обсягу, використання алгоритмів градієнтного бустингу над деревами рішень постає як сучасний індустріальний стандарт для моделювання табличних даних.

Так, модель *LightGBM* є модифікацією градієнтного бустингу, що використовує гістограмний підхід до побудови дерев рішень. Її застосування забезпечує високу обчислювальну ефективність при збереженні точності прогнозування. Додатково, вбудовані механізми балансування класів підвищують ефективність моделі у задачах класифікації напрямку ціни за умов асиметричного розподілу спостережень, зокрема, коли домінують спадні рухи.

Модель *XGBoost* (*eXtreme Gradient Boosting*) формує ансамбль дерев рішень послідовно, де кожне наступне дерево мінімізує помилки попередніх ітерацій. Завдяки використанню L1/L2-регуляризації та здатності працювати з багатовимірними ознаками, модель ефективно апроксимує нелінійні залежності між макроекономічними індикаторами (зокрема, VIX та S&P 500) і динамікою цін фінансових активів.

Таким чином, поєднання переваг лінійних і нелінійних підходів здійснено на основі використання двокомпонентної гібридної архітектури (рис. 1).

Моделювання часового ряду через декомпозицію та навчання на залишках охоплює такі основні етапи:

I етап: класична модель ARIMA формує базовий прогноз (Y_{ARIMA}), який відображає лінійну компоненту динаміки часового ряду.

II етап: на основі навчальної вибірки обчислюються залишки (residuals) як різниця між фактичними значеннями та прогнозом ARIMA:

$$E_t = Y_t - Y_{ARIMA} \quad (4)$$

III етап: алгоритм *XGBoost* використовується для моделювання нелінійної компоненти залишків. Залишки виступають цільовою змінною, а модель навчається апроксимувати нелінійну складову помилки $E_{XGBoost}$, використовуючи матрицю екзогенних макроекономічних ознак, що дозволяє враховувати складні нелінійні залежності між ринковими індикаторами та відхиленнями цін.

IV етап: фінальний прогноз гібридної архітектури формується як адитивна композиція двох компонентів:

$$Y_{final} = Y_{ARIMA} + E_{XGBoost} \quad (5)$$

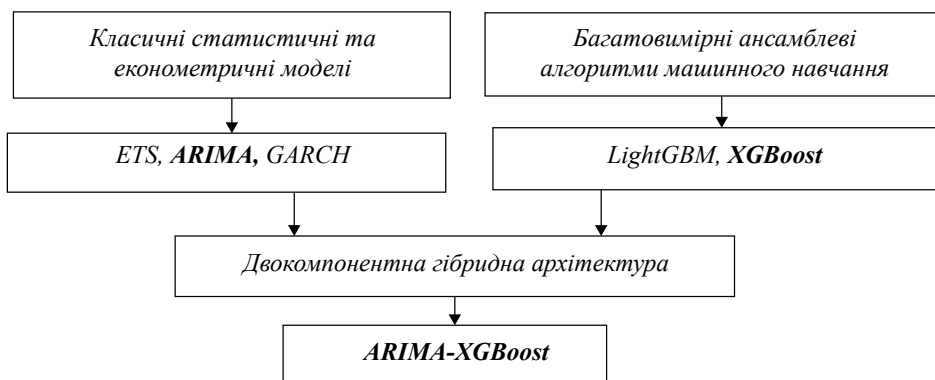


Рис. 1. Особливості формування двокомпонентної гібридної архітектури

Джерело: сформовано авторами

Такий підхід дозволив поєднати стабільність лінійного моделювання ARIMA з адаптивністю нелінійних алгоритмів машинного навчання, забезпечуючи високу точність прогнозування як у періоди ринкової стабільності, так і в умовах підвищеної волатильності та структурних зламів.

З метою забезпечення крос-секторальної репрезентативності дослідження, сформовано вибірку з шести транснаціональних корпорацій відповідно до світового стандарту класифікації галузей GICS (Global Industry Classification Standard):

- Advanced Micro Devices (AMD) – технологічний сектор;
- Coca-Cola (KO) – сектор товарів повсякденного попиту;
- ExxonMobil (XOM) – енергетичний сектор;
- JPMorgan Chase (JPM) – фінансовий сектор;
- Caterpillar (CAT) – промисловий сектор;
- Pfizer (PFE) – фармацевтичний сектор.

Емпіричною базою дослідження слугували історичні скориговані ціни закриття акцій (Adjusted Close) за період 2019–2025 рр. [24]. Для згладжування мікроструктурного шуму дані агреговано до тижневої частоти. Додатково до кожного часового ряду було включено технічні індикатори (зокрема, RSI та ковзні середні), а також макроекономічні змінні: індекс S&P 500, індекс волатильності VIX та технологічний індекс NASDAQ. Усі пояснювальні змінні сформовано із застосуванням зсуву на один період (shift(1)), що забезпечує відсутність витоку майбутньої інформації.

Прогноз здійснюється на один тиждень уперед з подальшим оновленням навчальної вибірки. Тестовим періодом обрано четвертий квартал 2025 року, який характеризується підвищеною волатильністю та концентрацією ринкового шуму, де класичні лінійні підходи є найменш ефективними. Нормалізована динаміка скоригованих цін акцій окремих транснаціональних корпорацій різних секторів економіки у передтестовий та тестовий періоди свідчить про перебування активів у різних фазах ринкового

циклу: від інтенсивного зростання в технологічному секторі до спадної динаміки у фармацевтичному та підвищеної волатильності у промисловому секторі (рис. 2).

Отже, сформовані часові ряди демонструють значну структурну неоднорідність та характеризуються змінною волатильністю, що обґрунтовує доцільність використання крос-секторального підходу та застосування ансамблевих та гібридних алгоритмів.

Оцінювання моделей базується на методології динамічного перенавчання – Expanding Window Walk-Forward Validation, яка дозволяє наблизити процес моделювання до реальних умов прогнозування та усуває look-ahead bias.

На кожному кроці модель навчається на розширеній вибірці:

$$D_{train}^{(t+1)} = D_{train}^{(t)} \cup \{t\}, D_{test}^{(t)} = \{t+1\} \quad (6)$$

де $D_{train}^{(t+1)}$ – навчальна вибірка на наступному кроці після оновлення,

$D_{train}^{(t)}$ – навчальна вибірка на кроці t ,

$D_{test}^{(t)}$ – тестова вибірка на кроці t ,

t – поточний момент часу (індекс спостереження),

$t+1$ – наступний часовий крок прогнозування,

\cup – операція об'єднання множин (union), що означає додавання нового спостереження до навчальної вибірки.

Оцінювання якості моделей здійснювалося за метрикою частки правильних прогнозів (Hit Rate, %):

$$HR = \frac{100}{N} \cdot \sum_{t=1}^N I(\hat{y}_t = y_t) \quad (7)$$

де $I(\cdot)$ – індикаторна функція.

Експеримент дозволив оцінити адаптивність алгоритмів як до стійких трендів, так і до структурних зламів та підвищеної волатильності. Детальні результати крос-секторального моделювання для статистичних методів, ансамблевих алгоритмів та гібридної архітектури наведено на тепловій карті (рис. 3).



Рис. 2. Нормалізована динаміка ціни акцій окремих транснаціональних корпорацій різних секторів економіки

Джерело: сформовано авторами на основі [24]



Рис. 3. Теплова карта частки правильних прогнозів напрямку ціни акцій

Джерело: сформовано авторами на основі [24]

Аналіз теплової карти демонструє виражену залежність ефективності моделей від секторальної специфіки. Так, у промисловому (Caterpillar) та фінансовому (JPMorgan) секторах, які демонструють виражену залежність від фаз макроекономічного циклу, ансамблеві методи (XGBoost) показали найвищу ефективність із точністю 61,54% та 76,92%, відповідно, перевищивши класичні статистичні моделі (46,15%–53,85%). Це підтверджує здатність деревоподібних ансамблів ефективно інкорпорувати екзогенні макроекономічні фактори.

Натомість класична модель ARIMA продемонструвала повну деградацію якості прогнозу (0,00%) для енергетичного сектору (ExxonMobil), що вказує на її обмежену здатність адаптації до структурних зламів. Гібридна архітектура в цьому випадку компенсувала дане обмеження, підвищивши точність до 61,54%.

Для фармацевтичного сектору (Pfizer) виявлено межі застосовності виключно кількісних підходів: жодна з моделей не перевищила рівень 53,85%, а складні моделі машинного навчання показали знижену ефективність (38,46%). Це пояснюється високою залежністю даного сектору від екзогенних неструктурованих факторів (зокрема, результатів клінічних досліджень та новинного фонду), які не відображаються у цінових рядах.

Багатовимірний профіль прогностичної ефективності моделей представлено на радарній діаграмі (рис. 4), де гібридна модель демонструє найбільш збалансовану продуктивність у різних секторах.

Крім середньої точності, важливим є показник стабільності прогнозів. Розподіл похибок, представлений на діаграмі розмаху (рис. 5), свідчить про те, що класичні статистичні моделі характеризуються значною варіативністю результатів.

У періоди структурних зламів їх точність знижується до рівня випадкового прогнозування (~50%) або нижче.

Наведені дані свідчать про те, що класичні статистичні моделі мають занадто широкий міжквартильний розмах. Незважаючи на високі результати в умовах ста-

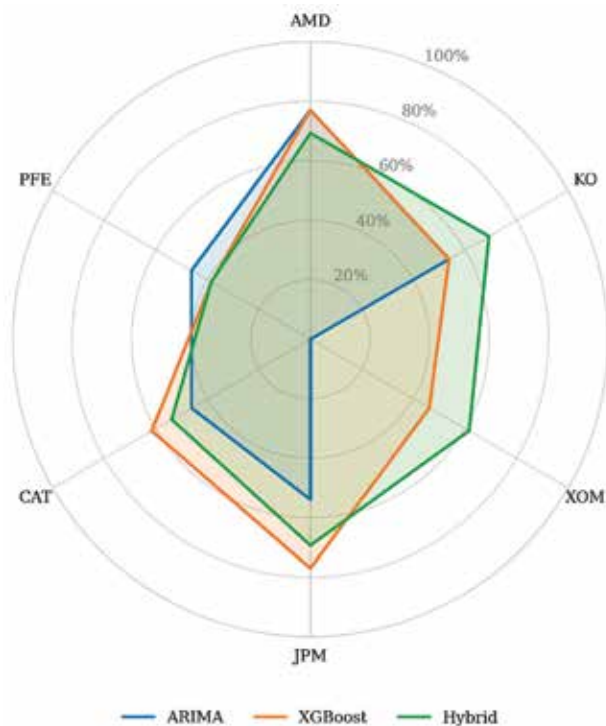


Рис. 4. Багатовимірний профіль прогностичної ефективності алгоритмів

Джерело: сформовано авторами на основі [24]

більного тренду, вони несуть неприпустимі ризики для інвестора в періоди структурних зламів (де їхня точність падає до 0% або 46%). Натомість гібридна архітектура демонструє більш стабільний розподіл результатів і рідше опускається нижче порогу випадковості у 50%.

Підсумковий рейтинг середньої ефективності моделей наведено на рис. 6.

Емпіричні результати свідчать про те, що запропонована гібридна архітектура досягає найвищого середнього

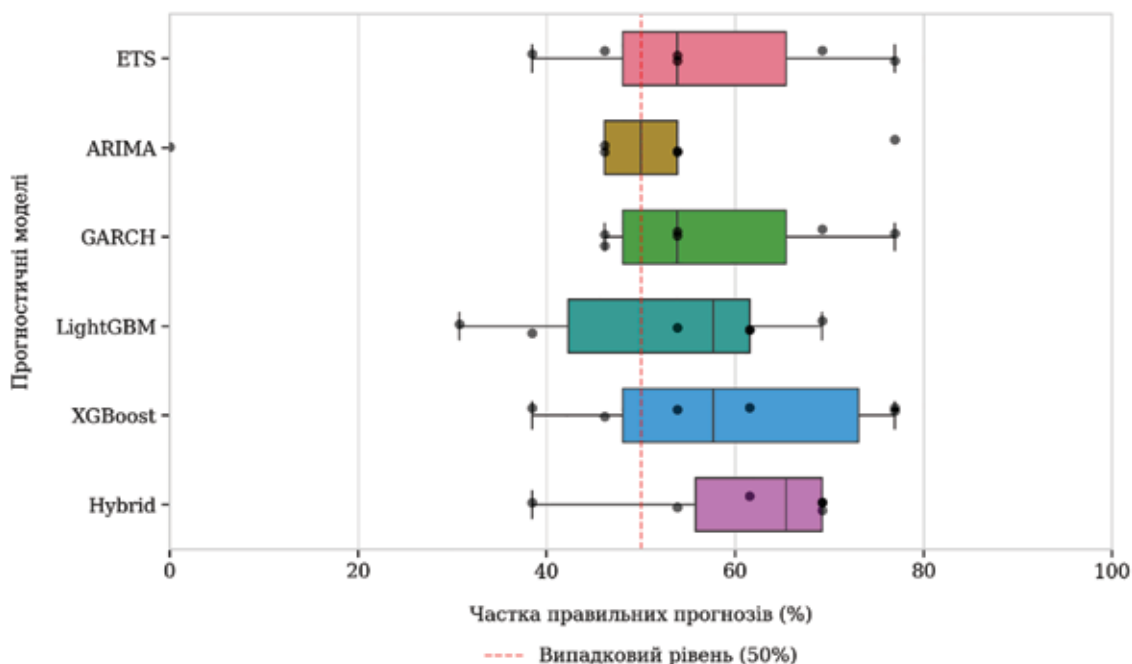


Рис. 5. Діаграма розмаху стабільності моделей за секторами економіки

Джерело: сформовано авторами

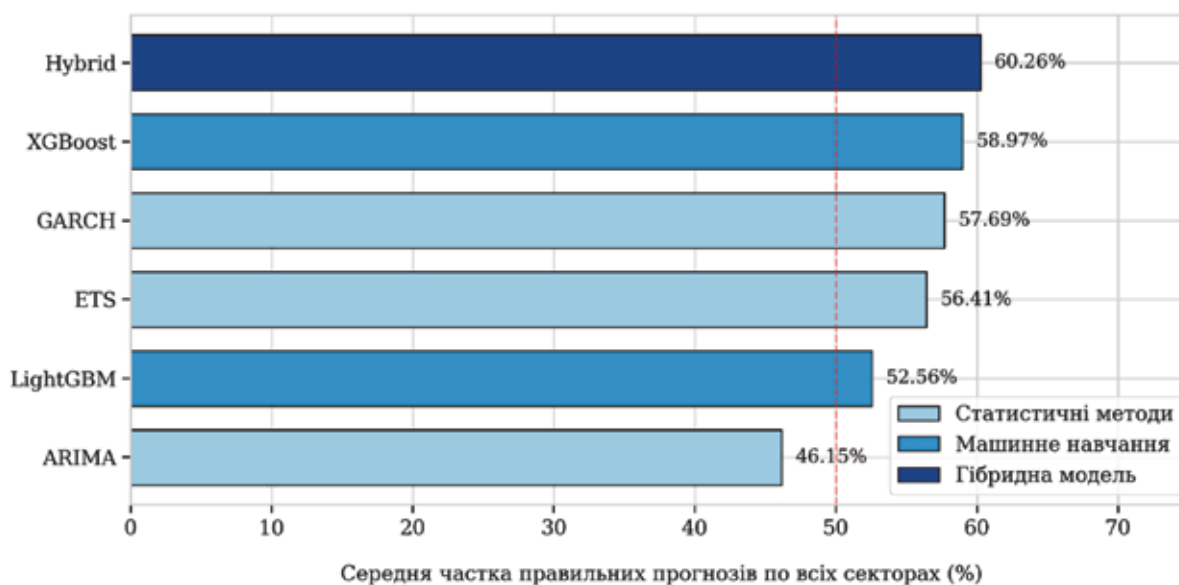


Рис. 6. Рейтинг середньої частки правильних прогнозів по всіх секторах

Джерело: сформовано авторами

значення точності (60,26%), перевищуючи як класичні статистичні моделі, так і окремі ансамблеві алгоритми.

Висновки. В статті представлено результати крос-секторальної емпіричної оцінки прогностичної ефективності лінійних статистичних моделей, методів машинного навчання та гібридної моделі (ARIMA–XGBoost) для задачі прогнозування напрямку руху цін фінансових активів. Аналіз проведено на вибірці з шести транснаціональних корпорацій, що охоплюють різні галузеві режими, на короткостроковому (квартальному) горизонті.

Застосування методології динамічного перенавчання Expanding Window Walk-Forward Validation забезпечило коректну оцінку моделей в умовах часової нестационарності та виключило ризик витоку даних.

Отримані результати свідчать про суттєві обмеження класичних економетричних підходів (ARIMA, ETS, GARCH) у періоди структурних змін та високої волатильності. Особливо низьку ефективність вони демонструють у секторах із високою чутливістю до макроекономічних шоків (енергетичному та промисловому).

Натомість ансамблеві моделі машинного навчання (XGBoost, LightGBM) забезпечують вищу прогностичну здатність завдяки моделюванню нелінійних залежностей між цінами активів та екзогенними макроекономічними індикаторами (S&P 500, VIX, NASDAQ).

Водночас дослідження підтверджує відсутність універсального алгоритму, точність якого стабільно перевершувала інші у всіх секторах. Зокрема, у фармацевтич-

ному секторі (Pfizer) ефективність кількісних моделей є обмеженою через домінування неструктурованих інформаційних факторів.

Найкращі результати за середньою точністю (60,26%) та стабільністю показників продемонструвала запропонована гібридна архітектура (ARIMA–XGBoost). Її ефективність пояснюється поєднанням лінійної екстраполяції

тренду та нелінійного моделювання залишкової компоненти за допомогою алгоритмів градієнтного бустингу.

Подальші дослідження можуть бути спрямовані розширення набору макроекономічних ознак, а також застосування більш складних архітектур для виявлення довгострокових нелінійних залежностей у фінансових часових рядах.

Список використаних джерел:

1. Chen H.-Y., Jenweeranon P., Alam N. Global Perspectives in the Metaverse. Palgrave Macmillan, Cham. 2024. DOI: https://doi.org/10.1007/978-3-031-54802-4_1
2. Garg K. Digital identities in the metaverse: Privacy, security, and user authentication in virtual financial systems. *International Journal of Financial Engineering*. 2024. № 11(4). DOI: <https://doi.org/10.1142/S242478632442009X>
3. Mishchenko V., Naumenkova S., Mishchenko S., Tishchenko I. Formation and Functioning of Financial Metaverse Platforms. *Financial and Credit Activity Problems of Theory and Practice*. 2025. № 1(60). P. 111–122. DOI: <https://doi.org/10.55643/fcaptop.1.60.2025.4689>
4. Міщенко В.І., Науменкова С.В. Механізми державної підтримки використання штучного інтелекту для забезпечення стійкості економічного розвитку. *Економіка України*. 2024. № 5(750). С. 30–56. DOI: <https://doi.org/10.15407/economyukr.2024.05.030>
5. Financial Stability Board. The Financial Stability Implications of Artificial Intelligence. 2024. URL: <https://www.fsb.org/2024/11/the-financial-stability-implications-of-artificial-intelligence/> (дата звернення: 10.03.2026)
6. Fabozzi F. J. The Handbook of Financial Instruments. John Wiley & Sons, Inc. 2018. 864 p.
7. Sanders A., Cornett M. Financial Markets and Institution. The McGraw-Hill, 2012. 754 p.
8. Fama E., French K. International Tests of a Five-Factor Asset Pricing Model. *Journal of Financial Economics*. 2017. № 123(3). P. 441–463. DOI: <https://doi.org/10.1016/j.jfineco.2016.11.004>
9. Paliienko O., Naumenkova S., Mishchenko S. An empirical investigation of the Fama-French five-factor model. *Investment Management and Financial Innovations*. 2020. № 17(1). P. 143–155. DOI: [https://doi.org/10.21511/imfi.17\(1\).2020.13](https://doi.org/10.21511/imfi.17(1).2020.13)
10. Hyndman R. J., Athanasopoulos G. Forecasting: Principles and Practice. 2021. URL: <https://otexts.com/fpp3/> (дата звернення: 15.01.2026)
11. Swamy S. R., Rajgoli S. R., Hegde T. Stock Market Prediction with Machine Learning: A Comprehensive Review. *Indiana Journal of Multidisciplinary Research*. 2024. № 4(3). P. 265–271. DOI: <https://doi.org/10.5281/zenodo.12685839>
12. Guo Z. Stock Market Price Prediction Using Machine Learning Models. *2nd International Conference on Financial Technology and Business Analysis*. 2023. № 45. P. 102–111. DOI: <https://doi.org/10.54254/2754-1169/45/20230266>
13. Kayit A., Ismail M. Advancing stock price prediction through the development of hybrid ensembles: a comprehensive comparative analysis of machine learning approaches. *Journal of Big Data*. 2025. № 12. 232. DOI: <https://doi.org/10.1186/s40537-025-01185-8>
14. Nti I., Adekoya F., Weyori B. A comprehensive evaluation of ensemble learning for stock market prediction. *Journal of Big Data*. 2020. № 7. 20. URL: <https://link.springer.com/article/10.1186/s40537-020-00299-5>
15. Tan Z., Yan Z., Zhu G. Stock selection with random forest: An exploitation of excess return in the Chinese stock market. *Heliyon*. 2019. № 5(8). DOI: <https://doi.org/10.1016/j.heliyon.2019.e02310>
16. Chen Q. Stock price forecasting using machine-learning methods. *Applied and Computational Engineering*. 2024. № 52. P. 208–214. DOI: <https://doi.org/10.54254/2755-2721/52/20241570>
17. Xie Y. Stock Price Forecasting: Traditional Statistical Methods and Deep Learning Methods. *Highlights in Business, Economics and Management*. 2023. № 21. P. 740–745. DOI: <https://doi.org/10.54097/hbem.v21i.14754>
18. Chen T., Guestrin C. XGBoost: A Scalable Tree Boosting System. *The 22nd ACM SIGKDD International Conference*. 2016. P. 785–794. DOI: <https://doi.org/10.1145/2939672.2939785>
19. Mutinda J. K., Langat A. K. Stock price prediction using combined GARCH-AI models. *Scientific African*. 2024. № 26. P. 1–14. DOI: <https://doi.org/10.1016/j.sciaf.2024.e02374>
20. Han H., Liu Z., Barrios M. et al. Time series forecasting model for non-stationary series pattern extraction using deep learning and GARCH modeling. *Journal of Cloud Computing*. 2024. № 13. 2. DOI: <https://doi.org/10.1186/s13677-023-00576-7>
21. Li M., Khan D., Alshambari H., El-Bagoury A. Prediction of complex stock market data using an improved hybrid EMD-LSTM model. *Applied Sciences*. 2023. № 13(3). 1429. DOI: <https://doi.org/10.3390/app13031429>
22. Inani S., Pradhan H., Kumar S., Biswas B. Navigating the technical analysis in stock markets: Insights from bibliometric and topic modeling approaches. *Investment Management and Financial Innovations*. 2024. № 21(1). P. 275–288. DOI: [https://doi.org/10.21511/imfi.21\(1\).2024.21](https://doi.org/10.21511/imfi.21(1).2024.21)
23. Міщенко С.В., Науменкова С.В., Тищенко Є.О. Синергія штучного інтелекту та блокчейн у фінансовому секторі. *Вісник Київського національного університету імені Тараса Шевченка. Економіка*. 2025. № 2(227). С. 54–64. DOI: <https://doi.org/10.17721/1728-2667.2025/227-2/7>
24. Yahoo Finance. URL: <https://finance.yahoo.com/> (дата звернення: 20.01.2026)

References:

1. Chen H.-Y., Jenweeranon P., Alam N. (2024) Global Perspectives in the Metaverse. Palgrave Macmillan, Cham. DOI: https://doi.org/10.1007/978-3-031-54802-4_1
2. Garg K. (2024) Digital identities in the metaverse: Privacy, security, and user authentication in virtual financial systems. *International Journal of Financial Engineering*, vol. 11(4). DOI: <https://doi.org/10.1142/S242478632442009X>
3. Mishchenko V., Naumenkova S., Mishchenko S., Tishchenko I. (2025) Formation and Functioning of Financial Metaverse Platforms. *Financial and Credit Activity Problems of Theory and Practice*, vol. 1(60), pp. 111–122. DOI: <https://doi.org/10.55643/fcaptop.1.60.2025.4689>
4. Mishchenko V., Naumenkova S. (2024) Mekhanizmy derzhavnoi pidtrymky vykorystannia shtuchnoho intelektu dlia zabezpechennia stiikosti ekonomichnoho rozvytku [State Support Mechanisms for the Use of Artificial Intelligence to Ensure Resilience of Economic Development]. *Economy of Ukraine*, vol. 67(5), pp. 30–56. DOI: <https://doi.org/10.15407/economyukr.2024.05.030>

5. Financial Stability Board. (2024) The Financial Stability Implications of Artificial Intelligence. Available at: <https://fsb.org/2024/11/the-financial-stability-implications-of-artificial-intelligence/> (accessed March 10, 2026)
6. Fabozzi F. J. (2018) *The Handbook of Financial Instruments*. John Wiley & Sons, Inc. 864 p.
7. Sanders A., Cornett M. (2012) *Financial Markets and Institution*. The McGraw-Hill. 754 p.
8. Fama E., French K. (2017) International Tests of a Five-Factor Asset Pricing Model. *Journal of Financial Economics*, vol. 123(3), pp. 441–463. DOI: <https://doi.org/10.1016/j.jfineco.2016.11.004>
9. Paliienko O., Naumenkova S., Mishchenko, S. (2020) An empirical investigation of the Fama-French five-factor model. *Investment Management and Financial Innovations*, vol. 17(1), pp. 143–155. DOI: [https://doi.org/10.21511/imfi.17\(1\).2020.13](https://doi.org/10.21511/imfi.17(1).2020.13)
10. Hyndman R. J., Athanasopoulos G. (2021) *Forecasting: Principles and Practice*. Monash University, Australia. Available at: <https://otexts.com/fpp3/> (accessed January 15, 2026)
11. Swamy S. R., Rajgoli S. R., Hegde T. (2024) Stock Market Prediction with Machine Learning: A Comprehensive Review. *Indiana Journal of Multidisciplinary Research*, vol. 4(3), pp. 265–271. DOI: <https://doi.org/10.5281/zenodo.12685839>
12. Guo, Z. (2023) Stock Market Price Prediction Using Machine Learning Models. *2nd International Conference on Financial Technology and Business Analysis*, no. (45), pp. 102–111. DOI: <https://doi.org/10.54254/2754-1169/45/20230266>
13. Kayit A., Ismail M. (2025). Advancing stock price prediction through the development of hybrid ensembles: a comprehensive comparative analysis of machine learning approaches. *Journal of Big Data*, vol. 12. 232. DOI: <https://doi.org/10.1186/s40537-025-01185-8>
14. Nti I., Adekoya F., Weyori B. (2020). A comprehensive evaluation of ensemble learning for stock market prediction. *Journal of Big Data*, vol. 7. 20. DOI: <https://doi.org/10.1186/s40537-020-00299-5>
15. Tan Z., Yan Z., Zhu G. (2019) Stock selection with random forest: An exploitation of excess return in the Chinese stock market. *Heliyon*, vol. 5(8). DOI: <https://doi.org/10.1016/j.heliyon.2019.e02310>
16. Chen Q. (2024) Stock price forecasting using machine-learning methods. *Applied and Computational Engineering*, vol. 52, pp. 208–214. DOI: <https://doi.org/10.54254/2755-2721/52/20241570>
17. Xie Y. (2023) Stock Price Forecasting: Traditional Statistical Methods and Deep Learning Methods. *Highlights in Business, Economics and Management*, vol. 21, pp. 740–745. DOI: <https://doi.org/10.54097/hbem.v21i.14754>
18. Chen T., Guestrin, C. (2016) XGBoost: A Scalable Tree Boosting System. *The 22nd ACM SIGKDD International Conference*, pp. 785–794. DOI: <https://doi.org/10.1145/2939672.2939785>
19. Mutinda J. K., & Langat A. K. (2024) Stock price prediction using combined GARCH-AI models. *Scientific African*, vol. 26, pp. 1–14. DOI: <https://doi.org/10.1016/j.sciaf.2024.e02374>
20. Han, H., Liu, Z., Barrios Barrios, M. et al. (2024). Time series forecasting model for non-stationary series pattern extraction using deep learning and GARCH modeling. *Journal of Cloud Computing*, vol. 13. 2. DOI: <https://doi.org/10.1186/s13677-023-00576-7>
21. Li M, Khan D, Alshanbari H, El-Bagoury A. (2023). Prediction of complex stock market data using an improved hybrid EMD-LSTM model. *Applied Sciences*, vol. 13(3), 1429. DOI: <https://doi.org/10.3390/app13031429>.
22. Inani S., Pradhan H., Kumar S., Biswas B. (2024) Navigating the technical analysis in stock markets: Insights from bibliometric and topic modeling approaches. *Investment Management and Financial Innovations*, vol. 21(1), pp. 275–288. DOI: [https://doi.org/10.21511/imfi.21\(1\).2024.21](https://doi.org/10.21511/imfi.21(1).2024.21)
23. Mishchenko S., Naumenkova S., Tishchenko I. (2025) Synerhiia shtuchnoho intelektu ta blokchein u finansovomu sektori [The Synergy of Artificial Intelligence and Blockchain Technologies in the Financial Sector]. *Bulletin of Taras Shevchenko National University of Kyiv. Economic*, vol. 2(227), pp. 54–64. DOI: <https://doi.org/10.17721/1728-2667.2025/227-2/7>
24. Yahoo Finance. Available at: <https://finance.yahoo.com/> (accessed January 20, 2026).

Дата надходження статті: 19.03.2026

Дата прийняття статті: 09.04.2026

Дата публікації статті: 24.06.2026