

УДК 004.8:336.76:519.866

DOI <https://doi.org/10.32782/tnv-tech.2024.5.4>

ВПРОВАДЖЕННЯ НЕЙРОМЕРЕЖЕВИХ АЛГОРИТМІВ ПРИ ПРОГНОЗУВАННІ ДИНАМІКИ ЦІН ФІНАНСОВИХ АКТИВІВ

Іваненко В. А. – аспірант

Київського національного університету імені Тараса Шевченка

ORCID ID: 0000-0002-1212-898X

Проведено аналіз особливостей впровадження нейромережеских алгоритмів при прогнозуванні динаміки цін фінансових активів. Зазначено, що впровадження методів автоматизації процедури прогнозування цін фінансових активів включає у себе постановку технічного завдання, пошук критеріїв, що вказують на вирішення поставленої задачі, оцінку статистичних показників, що вказують на обсяги даних фінансової аналітики, а також вибір і оптимізацію актуальних засобів системи машинного аналізу. Для визначення ефективності системи машинного аналізу даних фінансової аналітики було проведено її порівняння з лінійними моделями на основі методики оцінки капітальних активів, що базується на однопіриодичній інвестиційній схемі, концепції уникнення ризику з боку більшості інвесторів, а також припущенні про нульові транзакційні витрати і відсутність асиметрії інформації. Зазначено переваги застосування програмних пакетів ймовірнісного прогнозування, «NGBoost», як то визначення розподілу ймовірності для результатів, що дозволяє отримувати повноцінний розподіл можливих значень, вибір користувачем типу ймовірнісного розподілу для побудови моделі, що дає додаткову гнучкість для застосування в різних задачах та оптимізація ймовірнісних моделей методом натурального градієнта. Для оцінки продуктивності нейромережеских алгоритмів при автоматизації фінансової аналітики та прогнозування до аналізу були включені моделі «NGBoost», «XGBoost», «CatBoost», «LightGBM», неглибока багатошарова нейронна мережа прямого поширення та глибока багатошарова нейронна мережа прямого поширення, причому для кожної з нейромережеских архітектур було проведено оптимізацію гіперпараметрів. Зазначено, що вибір наведених моделей є показовим для застосування у фінансовій аналітиці. Поєднання градієнтного бустингу та нейромереж забезпечує широкий спектр методів для вирішення завдань прогнозування цін активів, оцінки ризиків і точного передбачення майбутніх фінансових показників. Водночас, використання гіперпараметричної оптимізації дає змогу налаштувати нейромережесві моделі для досягнення максимальної точності, що є критично важливим для забезпечення конкурентних переваг у фінансових ринках. Розглянуто особливості впровадження зазначеної системи оцінки з метою оптимізації нейромережеских алгоритмів для сучасних нейромережеских архітектур, як то моделі GRU, GAT і «Informer».

Ключові слова: машинний аналіз, нейромережева архітектура, фінансові активи, оцінка капітальних активів, кероване навчання, ймовірнісне прогнозування, гіперпараметрична оптимізація.

Ivanenko V. A. Implementation of neural network algorithms in forecasting the price dynamics of financial assets

The analysis of the implementation of neural network algorithms in forecasting the price dynamics of financial assets has been conducted. It was noted that the implementation of automated methods for predicting financial asset prices includes the formulation of technical tasks, the search for criteria that indicate the solution to the task, the evaluation of statistical indicators pointing to the volume of financial data, as well as the selection and optimization of relevant machine analysis tools. To determine the effectiveness of the financial data machine analysis system, it was compared with linear models based on the capital asset pricing model (CAPM), which is founded on a single-period investment scheme, the risk-averse nature of most investors, and the assumption of zero transaction costs and no information asymmetry. The advantages of the probabilistic forecasting software package «NGBoost» were highlighted, including the determination of probability distributions for outcomes, which allows for a full distribution of possible values, the flexibility for users to choose the type of probabilistic distribution for model building, and the optimization of probabilistic models using natural gradient descent, ensuring a more stable training process when working with complex distributions. To evaluate

the performance of neural network algorithms in automating financial analytics and forecasting, models such as «NGBoost», «XGBoost», «CatBoost», «LightGBM», shallow feed-forward neural networks, and deep feed-forward neural networks were included in the analysis, with hyperparameter optimization conducted for each neural network architecture. It was noted that the selection of these models is indicative for application in financial analytics. The combination of gradient boosting and neural networks provides a wide range of methods for solving tasks related to asset price forecasting, risk assessment, and accurate prediction of future financial indicators. At the same time, the use of hyperparameter optimization allows neural network models to be tuned to achieve maximum accuracy, which is critically important for ensuring competitive advantages in financial markets. The specific features of implementing this evaluation system have been examined to optimize neural network algorithms for modern neural network architectures, such as GRU, GAT, and Informer models.

Key words: machine analysis, neural network architecture, financial assets, capital asset pricing, supervised learning, probabilistic forecasting, hyperparameter optimization.

Вступ. Прогнозування динаміки цін фінансових активів на міжнародних фінансових ринках залишається однією з найбільш складних та актуальних задач, вирішення якої складає інтерес як при формуванні та розширенні інструментарію у галузі кількісного фінансування, так і з точки зору фундаментальної науки [1, 2]. Впровадження методів автоматизації процедури прогнозування цін фінансових активів при цьому включає у себе визначення таких аспектів як постановка завдання дослідження, пошук критеріїв, що вказують на вирішення поставленої задачі, оцінка статистичних показників, що вказують на обсяг масиву даних фінансової аналітики, а також вибір і оптимізація засобів системи машинного аналізу. Варто підкреслити, що перевага методів машинного аналізу даних фінансової аналітики значною мірою полягає саме у здатності працювати з обробки великих масивів неструктурованих даних [3, 4], зокрема, мультимедійних даних з широкого набору мережевих джерел, що відповідає концепції «альтернативних даних» (Alternative Data; AD). Дослідники вказують на широкий спектр джерел альтернативних даних, які використовуються в фінансовій аналітиці. Зазначається, що альтернативні дані виходять за межі традиційних фінансових показників, як то доходи чи звіти про прибутки, і включають такі фактори, як поведінка споживачів представлена у соціальних медіа і мережевих додатках електронної комерції [5, 6], динаміку забудови і дорожній трафік, представлені у супутниковими знімками [7, 8], тощо.

Як показав *аналіз наукових досліджень* присвячених проблемам автоматизації процедур фінансової аналітики [1–8] річна вартість засобів машинного навчання у банківській та фінансовій сферах вже наприкінці 2010-х років досягла рівня 5,2% світових доходів [9, 10]. Продуктивність алгоритму машинного навчання визначається здатністю здатним виявляти нелінійні взаємозв'язки з широкого спектра джерел даних [11], а отже, розробка відповідної системи прогнозування має включати у себе наступні аналіз наступних категорій:

1. Оцінка складності і нелінійності фінансової аналітики. Фінансові ринки характеризуються особливою динамікою, яка не може бути описана на рівні лінійної моделі і залежить від численних факторів [11].

2. Визначення особливостей Застосування нейронних мереж глибинного навчання (Deep Neural Networks, DNN). Алгоритми на основі відповідного класу нейромережевої архітектури характеризуються здатністю виявляти складні патерни і залежності великих масивах неструктурованих даних незалежно від формату їх представлення [12].

3. Порівняння продуктивності нейромережевих алгоритмів автоматизації процедури прогнозування цін фінансових активів у порівнянні з традиційними

методами оцінки капітальних активів (Capital Asset Pricing Model; CAPM), що базуються на математичних моделях, у яких використовуються лінійні зв'язки між цільовими функціями і аргументами цільових функцій [13].

Дослідження останніх десяти років демонструють високу ефективність алгоритмів машинного навчання для прогнозування цін на активи в порівнянні з моделями, розробленими в рамках традиційної статистики та фінансів [14–17], зокрема, при роботі з екзотичними деривативами, як то енергетичні [18] і крипто-валютні [19, 20].

Слід зазначити, тим не менш, що для підготовки цілісної методології автоматизації фінансової аналітики та прогнозування необхідно узагальнити підходи, що використовуються у окремих галузях фінансової аналітики, що розглядається як *невирішена частина загального дослідження*. Таким чином, *метою роботи* стала побудова комплексної методики організації алгоритмів машинного аналізу великих масивів даних із подальшим порівнянням їх із наявним інструментарієм CAPM на основі цільових показників.

Постановка завдання автоматизації процедури прогнозування цін фінансових активів. У відповідності до проведеного аналізу можна вказати, що постановка завдання дослідження включає у себе визначення наступних аспектів, що пов'язані з вирішенням завдання автоматизації процедури прогнозування цін фінансових активів:

1. Пошук критеріїв, що однозначно вказують на вирішення поставленої задачі та оптимізацію запропонованого підходу. Важливість зазначеного пункту пов'язана з тим, що прогнозування цін на фінансові активи як інструмент фінансової аналітики, у загальному випадку не має однозначних рішень на аналітичному рівні через високу волатильність ринків та вплив великої кількості факторів, що не можуть бути цілковито включені у аналіз.

2. Визначення на рівні статистичних показників обсягу поточних даних, що підлягають аналізу. Використання великих об'ємів даних (концепція «Big Data») суттєво розширює можливості для кращого аналізу фінансових ринків, але водночас ставить виклик щодо навантаження на обчислювальний ресурс апаратно-програмної платформи, що виконує їх обробку.

3. Вибір програмних і нейромережових алгоритмів системи машинного аналізу, на основі яких проводиться автоматизація процедури прогнозування цін фінансових активів. Актуалізація засобів машинного навчання у галузі фінансової аналітики пов'язана з можливістю обробки великих масивів даними, що вказують на нелінійні залежності, що призводить до кардинальних змін у практиці інвестиційних стратегій.

Узагальнена схема постановки завдання автоматизації процедури прогнозування цін фінансових активів у відповідності до наведеного переліку актуальних категорій представлена на рис. 1.

Для визначення ефективності системи машинного аналізу даних фінансової аналітики необхідно також провести узагальнення лінійної моделі, що лежить у основі CAPM, що базується на наступних підходах:

– одноперіодична інвестиційна модель, яка включає у себе припущення, що інвестиційні рішення проводяться здебільше на короткостроковій основі, без врахування довгострокової стратегії інвесторів, що дозволяє суттєво спростити процедуру прогнозування;

– концепція уникнення ризику з боку більшості інвесторів, що у ряді випадків призводить до великої кількості викидів при проведенні фінансової аналітики, що зменшує адекватність прогнозування;

- припущення про нульові трансакційні витрати, яке у ряді випадків не відповідає актуальним даним фінансової аналітики, що частково знецінює актуальний інструментарій системи прогнозування;
- припущення про відсутність асиметрії інформації, що відповідає представленню про рівний доступ до актуальних даних у режимі реального часу всіх інвесторів, що вважається ідеалістичним, оскільки в реальності інформація може бути розподілена нерівномірно між учасниками ринку, що призводить до дисбалансу.



Рис. 1. Узагальнена схема проектування та оптимізації системи прогнозування цін фінансових активів

Рівноважне співвідношення CAPM між очікуваною прибутковістю інвестиції i набору $i \in [1; I]$ та її ринковим бета-коефіцієнтом β_i визначається наступним чином:

$$ER(r_i) = r_F + \beta_i \cdot (ER(r_M) - r_F), \quad (1)$$

де функції $ER(r_i)$ і $ER(r_M)$ визначають очікувану прибутковість активу i та очікувана прибутковість ринкового портфеля, відповідно, причому r_F становить безризикову ставку прибутковості [21]. Таким чином, різниця $(ER(r_M) - r_F)$ визначає так звану ринкову премію за ризик (Market Risk Premium; MRP). У свою чергу, бета-коефіцієнт β_i є мірою того, наскільки актив i чутливий до змін на ринку:

$$\begin{cases} \beta_i > 1 \text{ визначає, що актив більш волатильний, ніж ринок} \\ \beta_i < 1 \text{ визначає, що актив менш волатильний, ніж ринок} \end{cases} \quad (2)$$

Відповідний показник визначається через коваріацію між дохідністю активу i та дохідністю ринкового портфеля:

$$\beta_i = \frac{\text{cov}(r_i - r_M)}{\sigma_M^2}, \quad (3)$$

де σ_M^2 є дисперсією дохідності ринкового портфеля. Коваріація $cov(r_i - r_M)$ показує, як дохідність активу i змінюється разом із дохідністю ринку. Якщо коваріація позитивна, це означає, що актив схильний зростати разом із ринком; якщо негативна – актив зростає, коли ринок падає, і навпаки. Дисперсія σ_M^2 є мірою волатильності дохідності ринку, вона показує, наскільки сильно дохідність ринкового портфеля відхиляється від свого середнього значення.

Опитування близько 4.500 міжнародних компаній продемонструвало, що CAPM є однією з найпопулярніших моделей, яку використовують фінансові директори компаній (Chief Financial Officers, CFO) для розрахунку вартості власного капіталу [21]. На її основі визначаються такі ключові показники як середньозважена вартість капіталу (Weighted Average Cost of Capital, WACC) як ставки дисконтування, що дозволяє визначити теперішню вартість майбутніх грошових потоків компанії та модель DCF (Discounted Cash Flow; DCF) є стандартним інструментом для оцінки вартості компанії на основі прогнозованих грошових потоків.

Вибір та оптимізація гіперпараметрів нейромережевої моделі для автоматизації фінансової аналітики і прогнозування. Впровадження нейромережевої архітектури при аналізі великих масивів даних у рамках автоматизації процедур фінансової аналітики і прогнозування надає можливість вказати на переваги алгоритмів керованого навчання (Supervised Learning, SL), що базуються на моделі дерева ухвалення рішень (Decision Tree, DT) або ансамблів дерев ухвалення рішень, як то випадкові ліси та градієнтні бустингові дерева [22]. При цьому для оцінки продуктивності нейромережевих алгоритмів у порівнянні з лінійними моделями CAPM використовувались програмні пакети ймовірнісного прогнозування [23], зокрема «NGBoost», що характеризується рядом переваг у порівнянні з аналогами (XGBoost, LightGBM та CatBoost):

- на відміну від традиційних реалізацій градієнтного бустингу, NGBoost прогнозує не лише середнє значення, а й дає розподіл імовірності для результатів, що дозволяє отримувати не просто «точкові» прогнози, а повноцінний розподіл можливих значень;

- при проведенні статистичного аналізу інструментарій NGBoost дозволяє обирати тип ймовірнісного розподілу для побудови моделі, що дає додаткову гнучкість для застосування в різних задачах;

- пакет використовує метод натурального градієнта, що є більш ефективним при оптимізації ймовірнісних моделей, забезпечує більш стабільний процес навчання при роботі зі складними розподілами.

Таким чином, програмний пакет «NGBoost» є корисним для задач, де важливі не лише точність прогнозу, а й оцінка рівня невизначеності, як то у фінансовому моделюванні, де прогнозування розподілу ризиків може допомогти приймати кращі рішення в умовах невизначеності.

Для кожної моделі, що підлягала аналізу, алгоритм машинного навчання характеризувався унікальним набором початкових параметрів (гіперпараметрів), що визначали його продуктивність. Автоматизація процесів оптимізації гіперпараметрів проводилась з застосуванням байесових алгоритмів з попереднім визначенням простору пошуку для кожного гіперпараметру моделі [24]. Порівняння продуктивності нейромережевих алгоритмів та програмних алгоритмів на основі лінійної моделі CAPM, таким чином, проводилось після застосування байесової оптимізації гіперпараметрів у відповідності до даних отриманих Центром передових обчислювальних досліджень Оксфордського університету (Advance Research Computing, ARC), що представлені у рамках відповідного дослідження [21]. Аналіз проводився для наступних моделей:

1. Модель «NGBoost» [23]. Алгоритм оптимізації: пошук по сітці (Grid Search; GS). Оптимізовані гіперпараметри: кількість оцінювачів дерев.

2. Модель «XGBoost» [25]. Алгоритм оптимізації: модель «HyperOpt» на основі дерева Парцена для 50 ітерацій. Оптимізовані гіперпараметри: кількість оцінювачів дерев, максимальна глибина кожного дерева, швидкість навчання, параметри регуляризації і параметри вибірки даних.

3. Модель «CatBoost» [26]. Алгоритм оптимізації: модель «HyperOpt» на основі дерева Парцена для 50 ітерацій. Оптимізовані гіперпараметри: кількість оцінювачів дерев, максимальна глибина кожного дерева, швидкість навчання, параметри регуляризації і параметри вибірки даних.

4. Модель «LightGBM» [27]. Алгоритм оптимізації: модель «HyperOpt» на основі дерева Парцена для 50 ітерацій. Оптимізовані гіперпараметри: кількість оцінювачів дерев, максимальна глибина кожного дерева, швидкість навчання, параметри регуляризації і параметри вибірки даних.

Неглибока багатошарова нейронна мережа прямого поширення (Shallow Feed-Forward Neural Network, Shallow FFNN) [25]. Алгоритм оптимізації: модель «HyperOpt» на основі дерева Парцена для 100 ітерацій. Оптимізовані гіперпараметри: кількість прихованих шарів $n \in [1; 2]$, кількість вузлів на кожному прихованому шарі $m \in [256; 1024]$, конфігурації нормалізації пакетів для кожного прихованого шару, параметри регуляризації та конфігурації активаційної функції для кожного прихованого шару.

Глибока багатошарова нейронна мережа (Deep Feed-Forward Neural Network, Deep FFNN) [25]. Алгоритм оптимізації: модель «HyperOpt» на основі дерева Парцена для 100 ітерацій. Оптимізовані гіперпараметри: кількість прихованих шарів, де $n \in [3; 5]$, кількість вузлів на кожному прихованому шарі, де $m \in [256; 1024]$, конфігурації нормалізації пакетів для кожного прихованого шару, параметри регуляризації та конфігурації активаційної функції для кожного прихованого шару.

Вибір наведених моделей є показовим для застосування у фінансовій аналітиці та прогнозуванні, оскільки вони представляють найсучасніші підходи в машинному навчанні, що дозволяють працювати з великими обсягами даних і складними залежностями. Поєднання градієнтного бустингу («NGBoost», «XGBoost», «CatBoost», «LightGBM») та нейронних мереж («Shallow FNN», «Deep FNN») забезпечує широкий спектр методів для вирішення завдань прогнозування цін активів, оцінки ризиків і точного передбачення майбутніх фінансових показників. Використання гіперпараметричної оптимізації, такої як «Grid Search» та «HyperOpt», дає змогу налаштовувати моделі для досягнення максимальної точності, що є критично важливим для забезпечення конкурентних переваг у фінансових ринках.

Оцінка ефективності алгоритмів машинного навчання при прогнозуванні цін фінансових активів. Оцінка продуктивності алгоритмів машинного аналізу проводилась на основі даних відкритих торгів акцій США протягом 1983–2019 рр., за даними хмарного сервісу «Wharton Research Data Services». Задачею алгоритмів машинного аналізу та програмних алгоритмів, що базуються на лінійній моделі CAPM було прогнозування річної прибутковості акцій на основі вхідних масивів даних щодо щомісячних та річних цін активів, фінансових звітів компаній, а також найбільш вагомих макроекономічних факторів.

Показником ефективності для всіх представлених моделей алгоритмів машинного аналізу та програмних алгоритмів, що базуються на лінійній моделі CAPM виступало значення середньоквадратичного відхилення (Mean Squared Deviation,

MSD), що є найбільш популярною метрикою для оцінки точності прогнозування. MSD обчислює середнє значення квадратів різниць між прогнозованими та фактичними значеннями, що дозволяє кількісно оцінити, наскільки точно модель передбачає значення у порівнянні з реальними даними. Моделі з меншим значенням MSD забезпечують кращу відповідність прогнозів до фактичних даних, що робить цю метрику критично важливою для оцінки моделей, які прогнозують фінансові показники. На математичному рівні MSD розраховується наступним чином:

$$MSD(PAR-AAR) = \frac{\sum_{k=1}^K (PAR_k - AAR_k)^2}{K}, \quad (4)$$

де множини $\{PAR_i\}$ та $\{AAR_i\}$ представляють собою прогнозовану річну прибутковість та фактичну річну прибутковість, відповідно, а $k \in [1; K]$ – повний набір спостережень, що виконувались у рамках прогнозування.

Результати дослідження представлені на рис. 2 демонструють, що методи машинного навчання значно покращують продуктивність прогнозування цін фінансових активів у порівнянні з класичною моделлю CAPM. Окремо можна відзначити переваги у прогнозуванні на основі архітектури глибокої нейромережі прямого поширення у порівнянні з неглибокою, що вказує на актуальні підходи для подальшої оптимізації системи прогнозування.

Також згідно з представленими результатами серед усіх моделей найкращі показники точності прогнозування мають «Catboost» і «XGBoost», які показали найменше значення MSE. Переваги зазначених моделей полягають у специфіці роботи з категорійними ознаками (Catboost оптимізований для обробки категорійних даних без попередньої трансформації), масштабованості для роботи з великими масивами даних, гнучкістю у налаштуванні параметрів, що надає можливість уникати перенавчання завдяки використанню регуляризаційних параметрів, а також ефективній роботі з розрідженими і даними, що може бракувати простішим моделям, як то «NGBoost».

Впровадження системи оцінки ефективності алгоритмів машинного навчання для сучасних моделей нейромережевої архітектури. Представлені моделі алгоритмів машинного навчання у рамках дослідження розглядаються як цілісний набір модельних об'єктів, що були використані для впровадження та адаптації системи оцінки засобів автоматизації прогнозування цін фінансових активів. У той же час необхідно вказати на актуальність впровадження та оцінки ефективності сучасних нейромережевих архітектур аналізу великих масивів фінансових даних, що здатні ефективно обробляти складні часові ряди і виявляти нелінійні зв'язки в даних. Таким чином, на даному етапі пропонується включити у аналіз наступні моделі штучних нейронних мереж (Artificial Neural Network; ANN):

- рекурентні нейронні мережі (Recurrent Neural Networks, RNN) з вентилярними рекурентними вузлами (Gated Recurrent Units; GRU), що можна розглядати як спрощену схему довгої-короткочасної пам'яті (Long Short-Term Memory, LSTM) без вихідного вентиля [28, 29];

- графові нейронні мережі (Graph Attention Networks; GAT), які здатні обробляти дані, представлені у формі графів, що надає можливість визначити взаємозв'язки між різними активами [30, 31];

- модель «Informer», побудована на основі архітектури трансформера та адаптована для обробки великих часових рядів [32, 33].

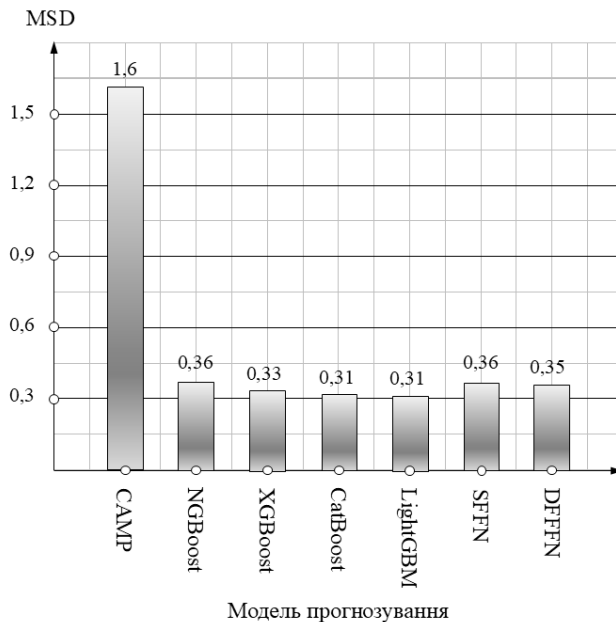


Рис. 2. Середнє квадратичне відхилення при прогнозуванні річної прибутковості акцій алгоритмами машинного аналізу

У основі побудови GRU лежить зменшення обчислювальної складності та вирішення проблеми зникання градієнту (Vanishing Gradient Problem; VGP), що є характерною для класичної архітектури RNN [28, 29]. Завдяки механізмам оновлення стану вентиля (Update Gate; UG) та скидання стану вентиля (Reset Gate; RG) відповідна модель здатна ефективно обробляти довгі часові ряди з утриманням для подальшого аналізу масивів даних, які можуть бути використані при прогнозуванні через виконання на базі UG та RG наступних функцій:

1. UG визначає масив даних, що підлягає збереженню, та масив даних, що має бути оновлено. Це дозволяє утримувати довгострокові залежності та, водночас, адаптуватися до нових даних при аналізі великих часових рядів з мінливими тенденціями.

2. RG визначає масив даних, що може ігноруватись при поточному аналізі. Це допомагає уникнути перенасичення контекстними даними, що сприяє більш точному аналізу короткострокових змін у часових рядах.

Відповідні механізми використовується при впровадженні нейромережевої архітектури для прогнозування фінансових показників, де залежності між попередніми та поточними значеннями можуть бути критичними, що, зокрема, надає можливість для ефективного застосування GRU при моделюванні щомісячних та річних цін акцій на основі послідовних змін макроекономічних факторів. Використання GRU у прогнозуванні річної прибутковості акцій є перспективним завдяки здатності цієї моделі відслідковувати тривалі тренди і, водночас, реагувати на локальні коливання. Механізми UG та RG дозволяють адаптуватися до місячних коливань цін, змін у фінансових звітах і макроекономічних індикаторах, зберігаючи при цьому важливі залежності, які можуть визначати майбутню прибутковість акцій.

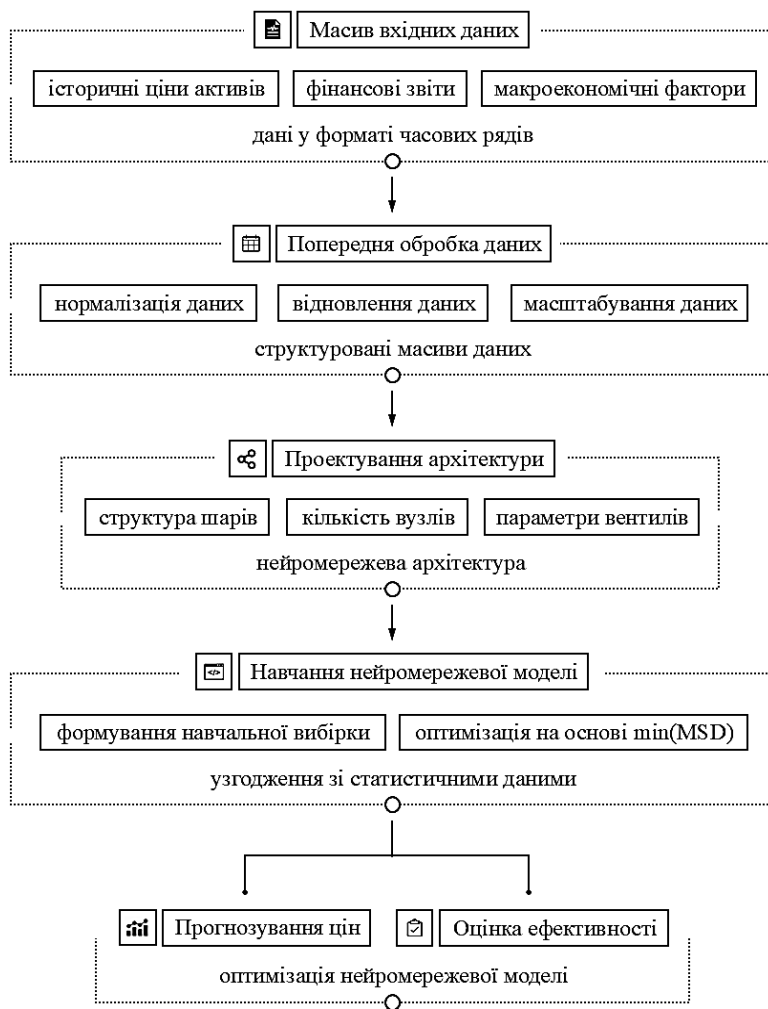


Рис. 3. Діаграма адаптації нейромережевої моделі на основі GRU для прогнозування цін фінансових активів

При цьому модель навчання GRU оптимізується на основі метрики MSD, що лежить у основі даного дослідження. Відповідний показник дозволяє кількісно оцінити, наскільки точно модель передбачає значення у порівнянні з реальними даними. На діаграмі (рис. 3) показана базова схема проектування, налаштування та оптимізації нейромережевої моделі на основі GRU для прогнозування цін фінансових активів.

Аналогічним чином, нейромережеві алгоритми на основі архітектури «Informer» як класу архітектури трансформера також обробляють вхідні дані представлені на основі часових рядів. Алгоритми «Informer» використовують механізм розрідженої самоуваги (Sparse Self-Attention; SSA), що дозволяє швидко обробляти довгі послідовності, знижуючи обчислювальну складність і, відповідно, навантаження на обчислювальний ресурс. Для фінансових даних це означає

можливість аналізувати великі масиви історичних цін масивів з високою точністю. Базова схема адаптації нейромережевої моделі на основі «Informer» для прогнозування цін фінансових активів включає у себе виконання наступних етапів:

1. Формування масиву вхідних даних, що представлені у вигляді часових рядів історичних даних про ціни активів, макроекономічні показники, фінансові звіти.

2. Впровадження механізму SSA шляхом виділення ключових елементів у часових рядах. SSA дозволяє фокусуватися на найважливіших даних, зменшуючи обчислювальну складність.

3. Перетворення набору вхідних даних через застосування кодувальних шарів (Coding Layers; CL) у більш компактний формат, який зберігає важливу інформацію для подальшого прогнозування.

4. Відновлення агрегованих даних та їх трансформація у прогнозовані значення річної прибутковості через застосування декодувальних шарів (Decoding Layers; DL).

5. Прогнозування річної прибутковості через формування прогнозованих значень прибутковості активів з використанням результатів аналізу часових рядів.

Після цього проводиться оцінка ефективності моделі «Informer» на основі метрики MSD з її подальшим коригуванням у разі необхідності.

Нейромережеві алгоритми на основі архітектури GAT надають можливість для обробки масивів даних, що представлені у формі графів. У фінансовому прогнозуванні модель GAT допомагає визначити взаємозв'язки між активами, наприклад, виявляючи вплив ринкових змін на пов'язані компанії чи сектори економіки. Завдяки механізму уваги GAT дозволяє виділяти найбільш важливі зв'язки у графі, присвоюючи їм більшу вагу при розрахунках. Базова схема адаптації нейромережевої моделі на основі GAT для прогнозування цін фінансових активів включає у себе виконання наступних етапів:

1. Створення графу активів (Asset Graph Creation; AGC), вузли якого представляють окремі активи, а зв'язки між вузлами – залежності між активами, як то схожість секторів, кореляція цін або спільні макроекономічні фактори.

2. Впровадження механізму уваги (Attention Mechanism; AM) через розрахунок ваг зв'язків для визначення важливості сусідніх вузлів. Механізм уваги виділяє найбільш значущі зв'язки, посилюючи вплив більш важливих сусідніх вузлів у графі.

3. Агрегація ознак сусідніх вузлів (Neighbor Feature Aggregation; NFA) з урахуванням ваг визначених у відповідності до AM, щоб кожен вузол отримав інформацію від своїх сусідів.

4. Прогнозування річної прибутковості (Prediction Output; PO) на основі агрегованих ознак для кожного вузла моделі GAT.

Після цього проводиться оцінка ефективності моделі GAT на основі метрики MSD з подальшим коригуванням моделі у разі необхідності.

Висновки. У результаті проведеного дослідження було проаналізовано особливості впровадження нейромережевих алгоритмів при прогнозуванні динаміки цін фінансових активів. Зазначено, що впровадження методів автоматизації процедури прогнозування цін фінансових активів включає у себе пошук критеріїв, що вказують на вирішення поставленої задачі, оцінку статистичних показників, що вказують на обсяги даних фінансової аналітики, а також вибір і оптимізацію актуальних засобів системи машинного аналізу. Для визначення ефективності системи машинного аналізу даних фінансової аналітики було проведено її порівняння з лінійними моделями на основі оцінки капітальних активів, що базується на одноперіодичній

інвестиційній схемі, концепції уникнення ризику з боку більшості інвесторів, а також припущенні про нульові транзакційні витрати і відсутність асиметрії інформації. Результати дослідження показують, що методи машинного навчання дозволяють кратно збільшити точність прогнозування цін фінансових активів у порівнянні з лінійними моделями SARIM. Розглянуто особливості впровадження зазначеної системи оцінки з метою оптимізації нейромережових алгоритмів для сучасних нейромережових архітектур, як то рекурентні нейронні мережі з вентильними рекурентними вузлами, графові нейронні мережі та модель «Informer», побудована на основі архітектури трансформера для обробки великих часових рядів.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ:

1. Henrique B. M., Sobreiro V. A., Kimura H. Practical machine learning: Forecasting daily financial markets directions. *Expert Systems with Applications*. 2023. Vol. 233. DOI: <https://doi.org/10.1016/j.eswa.2023.120840>.
2. Ashtiani M. N., Raahemi B. News-based intelligent prediction of financial markets using text mining and machine learning: A systematic literature review. *Expert Systems with Applications*. 2023. Vol. 217. DOI: <https://doi.org/10.1016/j.eswa.2023.119509>.
3. Monk A., Prins M., Rook D. Rethinking alternative data in institutional investment. *The Journal of Financial Data Science*. 2019. Vol. 1. No. 1. P. 14–31. DOI: <https://doi.org/10.3905/jfds.2019.1.1.014>.
4. Barker A., Monk A., Rook D. Technological disruption and long-term investors: Managing risk and opportunities. *SSRN Electronic Journal*. 2024. DOI: <https://doi.org/10.2139/ssrn.4741908>.
5. Risk and risk management in the credit card industry / Butaru F., Chen Q., Clark B., et al. *Journal of Banking & Finance*. 2016. Vol. 72. DOI: <https://doi.org/10.1016/j.jbankfin.2016.07.015>.
6. Gottfried M. How credit-card data might be distorting retail stocks. *The Wall Street Journal*. 2017. URL: <https://www.wsj.com/articles/how-credit-card-data-might-be-distorting-retail-stocks-1483468912> (last accessed: 29.10.2024).
7. A deep neural network for oil spill semantic segmentation in SAR images / Orfanidis J., Ioannidis K., Avgerinakis K., et al. *ICIP 2018*. 2018. P. 3773–3777. DOI: <https://doi.org/10.1109/ICIP.2018.8451113>.
8. Beyond spatial auto-regressive models: Predicting housing prices with satellite imagery / Bency A., Rallapalli S., Ganti R., et al. *WACV 2017*. 2019. P. 320–329. DOI: <https://doi.org/10.1109/WACV.2017.42>.
9. Buchanan. Artificial intelligence in finance. *The Alan Turing Institute. Apr.* 2019. DOI: <https://doi.org/10.5281/zenodo.2612537>.
10. McKinsey on Payments: Special Edition on Advanced Analytics in Banking / Buehler D'Silva, Fitzpatrick, et al. *Tech. rep.* Aug. 2018. DOI: <https://doi.org/10.48550/arXiv.2004.01504>.
11. Heaton J. B., Polson N. Deep learning for finance: Deep portfolios. *SSRN Electronic Journal*. 2016. DOI: <https://doi.org/10.2139/ssrn.2838013>.
12. Warin X. Deep learning for efficient frontier calculation in finance. *Journal of Computational Finance*. 2022. DOI: <https://doi.org/10.21314/jcf.2021.017>.
13. Chen J. M. The intertemporal capital asset pricing model. *Econophysics and Capital Asset Pricing*. 2017. P. 127–138. DOI: https://doi.org/10.1007/978-3-319-63465-4_7.
14. Kyriakou I., Mousavi P., Nielsen J. P., et al. Machine learning for forecasting excess stock returns – The five-year-view. *Graz Economics Papers 2019–06*. 2019. URL: <https://ideas.repec.org/p/grz/wpaper/2019-06.html> (last accessed: 29.10.2024).
15. Equity2vec / Wu Q., Brinton C. G., Zhang Z., Pizzoferrato A., Liu Z., Cucuringu M. *Proceedings of the Second ACM International Conference on AI in Finance*. 2021. Vol. 79. P. 1–9. DOI: <https://doi.org/10.1145/3490354.3494409>.

16. A deep learning framework for pricing financial instruments / Wu Q., Zhang Z., Pizzoferrato A., et al. *arXiv.org*. Sept. 2019. Article ID 1909.04497.
 17. Gu S., Kelly B. T., Xiu D. Empirical asset pricing via machine learning. *SSRN Electronic Journal*. Jan. 2018. DOI: <https://doi.org/10.2139/ssrn.3281018>.
 18. An extensive evaluation of seven machine learning methods for rainfall prediction in weather derivatives / Cramer S., Kampouridis M., Freitas A. A., Alexandridis A., et al. *Expert Systems with Applications*. 2017. Vol. 85. DOI: <https://doi.org/10.1016/j.eswa.2017.05.029>.
 19. Machine learning the cryptocurrency market / Alessandretti L., Bahrawy A., Aiello L. M., Baronchelli A. *SSRN Electronic Journal*. May 2018. DOI: <https://doi.org/10.2139/ssrn.3183792>.
 20. Lahmiri S., Bekiros S. Cryptocurrency forecasting with deep learning chaotic neural networks. *Chaos, Solitons and Fractals*. 2019. Vol. 118. P. 35–40. DOI: <https://doi.org/10.1016/j.chaos.2018.11.014>.
 21. Ndikum P. Machine learning algorithms for financial asset price forecasting. *ArXiv* 2004.01504v1. 2020. P. 1–16. DOI: <https://doi.org/10.48550/arXiv.2004.01504>.
 22. Kelly B. T., Pruitt S., Su Y. Some characteristics are risk exposures, and the rest are irrelevant. *SSRN Electronic Journal*. Jan. 2017. DOI: <https://doi.org/10.2139/ssrn.3032013>.
 23. Ghahramani Z. Probabilistic machine learning and artificial intelligence. *Nature*. 2015. Vol. 521. No. 7553. P. 452–459. DOI: <https://doi.org/10.1038/nature14541>.
 24. Zöllner M. – A., Huber M. F. Benchmark and survey of automated machine learning frameworks. *Journal of Artificial Intelligence Research*. 2021. Vol. 70. P. 409–472. DOI: <https://doi.org/10.1613/jair.1.11854>.
 25. Bergstra J., Yamins D., Cox D. Hyperopt: A python library for optimizing the hyperparameters of machine learning algorithms. *Proceedings of the Python in Science Conference*. 2013. P. 13–19. DOI: <https://doi.org/10.25080/majora-8b375195-003>.
 26. Ershov V. High performance insights from GPU version CatBoost. *Computer Tools in Education*. 2022. No. 2. P. 59–73. DOI: <https://doi.org/10.32603/2071-2340-2022-2-59-73>.
 27. Sukarsa I. M., Pandika Pinata N. N., Dwi Rusjyanthi N. K., Wisswani N. W. Estimation of gourami supplies using gradient boosting decision tree method of XGBoost. *TEM Journal*. 2021. P. 144–151. DOI: <https://doi.org/10.18421/tem101-17>.
 28. Kwak N. W., Lim D. H. Financial time series forecasting using AdaBoost-GRU ensemble model. *Journal of the Korean Data and Information Science Society*. 2021. Vol. 32. No. 2. P. 267–281. DOI: <https://doi.org/10.7465/jkdi.2021.32.2.267>.
 29. Liu B., Lai M. Advanced machine learning for financial markets: A PCA-GRU-LSTM approach. *Journal of the Knowledge Economy*. 2024. DOI: <https://doi.org/10.1007/s13132-024-02108-3>.
 30. Cheng H., Wang K., Tan X. A link prediction method for Chinese financial event knowledge graph based on graph attention networks and convolutional neural networks. *Engineering Applications of Artificial Intelligence*. 2024. Vol. 138. DOI: <https://doi.org/10.1016/j.engappai.2024.109361>.
 31. Graph attention site prediction (GRASP): Identifying druggable binding sites using graph neural networks with attention / Smith Z., Strobel M., Vani B. P., Tiwary P. *bioRxiv*: website. 2023. DOI: <https://doi.org/10.1101/2023.07.25.550565>.
 32. Wu Y. Comparison between Transformer, Informer, Autoformer and non-stationary Transformer in financial market. *Applied and Computational Engineering*. 2023. Vol. 29. No. 1. P. 68–78. DOI: <https://doi.org/10.54254/2755-2721/29/20230874>.
 33. Ojeda C., Artal C., Tejera F. Informer, an information organization Transformer architecture. *Proceedings of the 13th International Conference on Agents and Artificial Intelligence*. 2021. P. 381–389. DOI: <https://doi.org/10.5220/0010372703810389>.
-

REFERENCES:

1. Henrique, B. M., Sobreiro, V. A., & Kimura, H. (2023). Practical machine learning: Forecasting daily financial markets directions. *Expert Systems with Applications*, 233. DOI: <https://doi.org/10.1016/j.eswa.2023.120840> [In English].
2. Ashtiani, M. N., & Raahemi, B. (2023). News-based intelligent prediction of financial markets using text mining and machine learning: A systematic literature review. *Expert Systems with Applications*, 217. DOI: <https://doi.org/10.1016/j.eswa.2023.119509> [In English].
3. Monk, A., Prins, M., & Rook, D. (2019). Rethinking alternative data in institutional investment. *The Journal of Financial Data Science*, 1 (1), 14–31. DOI: <https://doi.org/10.3905/jfds.2019.1.1.014> [In English].
4. Barker, A., Monk, A., & Rook, D. (2024). Technological disruption and long-term investors: Managing risk and opportunities. *SSRN Electronic Journal*. DOI: <https://doi.org/10.2139/ssrn.4741908> [In English].
5. Butaru, F., Chen, Q., Clark, B., et al. (2016). Risk and risk management in the credit card industry. *Journal of Banking & Finance*, 72. DOI: <https://doi.org/10.1016/j.jbankfin.2016.07.015> [In English].
6. Gottfried, M. (2017). How credit-card data might be distorting retail stocks. *The Wall Street Journal*. Retrieved from <https://www.wsj.com/articles/how-credit-card-data-might-be-distorting-retail-stocks-1483468912> (last accessed: 29.10.2024) [In English].
7. Orfanidis, J., Ioannidis, K., Avgerinakis, K., et al. (2018). A deep neural network for oil spill semantic segmentation in SAR images. *ICIP 2018 – ICIP 2018*, 3773–3777. DOI: <https://doi.org/10.1109/ICIP.2018.8451113> [In English].
8. Bency, A., Rallapalli, S., Ganti, R., et al. (2019). Beyond spatial auto-regressive models: Predicting housing prices with satellite imagery. *WACV 2017 – WACV 2017*, 320–329. DOI: <https://doi.org/10.1109/WACV.2017.42> [In English].
9. Buchanan. (2019, April). Artificial intelligence in finance. *The Alan Turing Institute*. DOI: <https://doi.org/10.5281/zenodo.2612537> [In English].
10. Buehler, D’Silva, Fitzpatrick, et al. (2018, August). McKinsey on Payments: Special Edition on Advanced Analytics in Banking. *Tech. rep.* DOI: <https://doi.org/10.48550/arXiv.2004.01504> [In English].
11. Heaton, J. B., & Polson, N. (2016). Deep learning for finance: Deep portfolios. *SSRN Electronic Journal*. DOI: <https://doi.org/10.2139/ssrn.2838013> [In English].
12. Warin, X. (2022). Deep learning for efficient frontier calculation in finance. *Journal of Computational Finance*. DOI: <https://doi.org/10.21314/jcf.2021.017> [In English].
13. Chen, J. M. (2017). The intertemporal capital asset pricing model. In *Econophysics and Capital Asset Pricing*, 127–138. DOI: https://doi.org/10.1007/978-3-319-63465-4_7 [In English].
14. Kyriakou, I., Mousavi, P., Nielsen, J. P., et al. (2019). Machine learning for forecasting excess stock returns – The five-year-view. *Graz Economics Papers 2019–06*. Retrieved from <https://ideas.repec.org/p/grz/wpaper/2019-06.html> (last accessed: 29.10.2024) [In English].
15. Wu, Q., Brinton, C. G., Zhang, Z., Pizzoferrato, A., Liu, Z., & Cucuringu, M. (2021). Equity2vec. *Proceedings of the Second ACM International Conference on AI in Finance*, 79, 1–9. DOI: <https://doi.org/10.1145/3490354.3494409> [In English].
16. Wu, Q., Zhang, Z., Pizzoferrato, A., et al. (2019, September). A deep learning framework for pricing financial instruments. *arXiv.org*, Article ID 1909.04497 [In English].
17. Gu, S., Kelly, B. T., & Xiu, D. (2018, January). Empirical asset pricing via machine learning. *SSRN Electronic Journal*. DOI: <https://doi.org/10.2139/ssrn.3281018> [In English].
18. Cramer, S., Kampouridis, M., Freitas, A. A., & Alexandridis, A. (2017). An extensive evaluation of seven machine learning methods for rainfall prediction in weather

derivatives. *Expert Systems with Applications*, 85. DOI: <https://doi.org/10.1016/j.eswa.2017.05.029> [In English].

19. Alessandretti, L., Bahrawy, A., Aiello, L. M., & Baronchelli, A. (2018, May). Machine learning the cryptocurrency market. *SSRN Electronic Journal*. DOI: <https://doi.org/10.2139/ssrn.3183792> [In English].

20. Lahmiri, S., & Bekiros, S. (2019). Cryptocurrency forecasting with deep learning chaotic neural networks. *Chaos, Solitons and Fractals*, 118, 35–40. DOI: <https://doi.org/10.1016/j.chaos.2018.11.014> [In English].

21. Ndikum, P. (2020). Machine learning algorithms for financial asset price forecasting. *ArXiv 2004.01504v1*, 1–16. DOI: <https://doi.org/10.48550/arXiv.2004.01504> [In English].

22. Kelly, B. T., Pruitt, S., & Su, Y. (2017, January). Some characteristics are risk exposures, and the rest are irrelevant. *SSRN Electronic Journal*. DOI: <https://doi.org/10.2139/ssrn.3032013> [In English].

23. Ghahramani, Z. (2015). Probabilistic machine learning and artificial intelligence. *Nature*, 521 (7553), 452–459. DOI: <https://doi.org/10.1038/nature14541> [In English].

24. Zöllner, M. – A., & Huber, M. F. (2021). Benchmark and survey of automated machine learning frameworks. *Journal of Artificial Intelligence Research*, 70, 409–472. DOI: <https://doi.org/10.1613/jair.1.11854> [In English].

25. Bergstra, J., Yamins, D., & Cox, D. (2013). Hyperopt: A python library for optimizing the hyperparameters of machine learning algorithms. *Proceedings of the Python in Science Conference*, 13–19. DOI: <https://doi.org/10.25080/majora-8b375195-003> [In English].

26. Ershov, V. (2022). High performance insights from GPU version CatBoost. *Computer Tools in Education*, (2), 59–73. DOI: <https://doi.org/10.32603/2071-2340-2022-2-59-73> [In English].

27. Sukarsa, I. M., Pandika Pinata, N. N., Dwi Rusjayanthi, N. K., & Wisswani, N. W. (2021). Estimation of gourami supplies using gradient boosting decision tree method of XGBoost. *TEM Journal*, 144–151. DOI: <https://doi.org/10.18421/tem101-17> [In English].

28. Kwak, N. W., & Lim, D. H. (2021). Financial time series forecasting using AdaBoost-GRU ensemble model. *Journal of the Korean Data and Information Science Society*, 32 (2), 267–281. DOI: <https://doi.org/10.7465/jkdi.2021.32.2.267> [In English].

29. Liu, B., & Lai, M. (2024). Advanced machine learning for financial markets: A PCA-GRU-LSTM approach. *Journal of the Knowledge Economy*. DOI: <https://doi.org/10.1007/s13132-024-02108-3> [In English].

30. Cheng, H., Wang, K., & Tan, X. (2024). A link prediction method for Chinese financial event knowledge graph based on graph attention networks and convolutional neural networks. *Engineering Applications of Artificial Intelligence*, 138. DOI: <https://doi.org/10.1016/j.engappai.2024.109361> [In English].

31. Smith, Z., Strobel, M., Vani, B. P., & Tiwary, P. (2023). Graph attention site prediction (GRASP): Identifying druggable binding sites using graph neural networks with attention. *bioRxiv*. DOI: <https://doi.org/10.1101/2023.07.25.550565> [In English].

32. Wu, Y. (2023). Comparison between Transformer, Informer, Autoformer and non-stationary Transformer in financial market. *Applied and Computational Engineering*, 29(1), 68–78. DOI: <https://doi.org/10.54254/2755-2721/29/20230874> [In English].

33. Ojeda, C., Artal, C., & Tejera, F. (2021). Informer, an information organization Transformer architecture. *Proceedings of the 13th International Conference on Agents and Artificial Intelligence*, 381–389. DOI: <https://doi.org/10.5220/0010372703810389> [In English].