

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
ЧЕРНІВЕЦЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ
ІМЕНІ ЮРІЯ ФЕДЬКОВИЧА

Кваліфікаційна наукова праця
на правах рукопису

Токарєва Катерина Анатоліївна

УДК 004.738.1:332.12

ДИСЕРТАЦІЯ

**ГІБРИДНІ ТА АНСАМБЛЕВІ МЕТОДИ ТА МОДЕЛІ
МАШИННОГО НАВЧАННЯ ПРОГНОЗУВАННЯ ФІНАНСОВИХ
ЧАСОВИХ РЯДІВ**

113 – Прикладна математика

11 – Математика та статистика

Подається на здобуття наукового ступеня доктора філософії.

Дисертація містить результати власних досліджень. Використання ідей, результатів і текстів авторів мають посилання на відповідне джерело



К.А. Токарєва

Науковий керівник **Кушнір Микола Ярославович**, кандидат фізико-математичних наук, доцент

Чернівці – 2024

АНОТАЦІЯ

Токарева К.А. Гібридні та ансамблеві методи та моделі машинного навчання прогнозування фінансових часових рядів. – Кваліфікаційна наукова праця на правах рукопису.

Дисертація на здобуття наукового ступеня доктора філософії за спеціальністю 113 – Прикладна математика. – Чернівецький національний університет імені Юрія Федьковича МОН України, Чернівці, 2024.

Дисертаційна робота присвячена побудові гібридних алгоритмів на основі методів та моделей машинного навчання та класичних алгоритмів часових рядів.

Результати дисертаційної роботи є підґрунтям для подальших теоретичних і практичних наукових розробок у дослідженні гібридних алгоритмів побудови прогнозу для фінансових часових рядів.

Дисертація складається із вступу, чотирьох розділів, висновків, переліку використаних джерел та додатку.

У вступі обґрунтовано актуальність теми дослідження, сформульовано мету, завдання, предмет, об'єкт та методи дослідження, вказано наукову новизну, теоретичне та практичне значення отриманих результатів, проаналізовано зв'язок роботи з науковими дослідженнями та особистий внесок здобувача, а також наведено відомості про апробацію та публікації основних результатів дисертації. Описано структуру та обсяг дисертаційної роботи.

У першому розділі здійснено огляд наукової літератури, присвяченої основним напрямкам досліджень часових рядів в цілому та фінансових часових рядів зокрема, розглянуто ключові відомості з теорії часових рядів та методів машинного навчання, наведено опис основних напрямів досліджень та визначено завдання, якими займається теорія часових рядів включаючи основні проблеми побудови прогнозу фінансових часових рядів. Проведено детальний огляд основних класичних алгоритмів побудови прогнозу для часового ряду, зокрема, а також їх сфери застосування. Детально

проаналізовано хронологію розвитку теорії часових рядів та умовні періоди ускладнення моделей. У другому розділі даного розділу розглянуто основні моделі штучного інтелекту та машинного навчання для аналізу часових рядів, описано топології нейронних мереж, які будуть використовуватися при побудові гібридних алгоритмів.

У другому розділі проведено дослідження гібридних моделей, в яких основний процес оцінювався на основі моделі часового ряду, а залишки моделі оцінювалися на основі нейронних мереж. Використання такого підходу дало змогу розширення використання класичних моделей часових рядів на випадок відсутності гомоскедастичності у залишках та відсутності незалежності залишків часового ряду. А саме показано, що

- часові ряди фондових індексів містять як лінійну, так і не лінійну складову, а отже, окремо лінійна модель ARIMA та нелінійна модель ANN не можуть дати точну оцінку таких часових рядів;
- розроблено гібридну модель моделі авторегресії-ковзної середньої та штучної нейронної мережі прогнозування часових рядів фондових індексів;
- доведено, якщо часовий ряд є строго стаціонарним, то ARIMA модель є більш адекватною для гауссівських рядів в розробленій гібридній моделі.
- здійснено порівняльний аналіз запропонованого алгоритму із класичними алгоритмами оцінки прогнозу фінансових часових рядів, в результаті чого розроблений метод продемонстрував більшу точність порівняно з іншими гібридними моделями, оскільки ряд даних розкладається на компоненти з низькою та високою волатильністю (лінійну та нелінійну), а отже сумарна похибка моделі є низькою.

У третьому розділі розглянуто зв'язок хаотичних динамічних систем та часових. Проаналізовано основні підходи до визначення детермінованого хаосу та основні методи врахування наявності детермінованого хаосу при

оцінці прогнозу фінансового часового ряду. Для визначення існування хаосу в системі були представлені чотири методи:

- перетворення Фур'є;
- показник Ляпунова;
- показник Герста;
- фрактальна розмірність.

Основним підходом для застосування теорії хаосу, а саме детермінованого хаосу, є використання теореми вкладення Такенса, згідно якої структуру хаотичного детермінованого процесу можна спростити використовуючи розбиття векторного простору $(x(t), \dots, x(t + T))$. Метод реконструкції фазового простору використовується для реконструкції точок фазового простору з хаотичних часових рядів з використанням розмірності вкладення D та часової затримки T . Відповідно до методу реконструкції фазового простору, точки фазового простору відтворюють ключові характеристики вихідного часового ряду. Для підтвердження ефективності запропонованої методики в даному розділі було досліджено три різні класичні хаотичні часових ряди: Маккея-Глесса, логістичного та Ено. Запропонований метод об'єднувався з генетичними алгоритмами для покращення результатів прогнозування. Було показано, що новий метод прогнозування, покращений за допомогою GA, демонструє визначні результати. Порівняння результатів запропонованої методики з результатами, наведеними в літературі, обґрунтовує ефективність запропонованого методу.

У розділі 4 досліджено дві гібридні моделі, які ґрунтуються на моделях часових рядів. Перша модель являє собою неперервний та нелінійний аналог класичної лінійної ARMA моделі, друга модель являє собою узагальнення класичної LSTM нейронної мережі із врахуванням довільної скінченної передісторії. Перша модель базується на новому методі оцінки прогнозу процесів фінансових ринків в основі якого лежать моделі стохастичних диференціальних рівнянь із запізненням. Для оцінки параметрів в розробленій моделі використано генетичний алгоритм, а саме

метод рою частинок, оскільки цей метод виявився найбільш оптимальним в нашій задачі. Також вказано один із підходів переходу від оптимізаційної задачі по класу функцій до параметричної оптимізаційної задачі, що дозволяє значно ефективніше використовувати генетичні алгоритми пошуку оптимальних параметрів. Отримані теоретичні результати протестовано на реальних даних із використанням динаміки трьох реальних фінансових процесів. Результати тестування показали, що нова модель краще апроксимує реальні фінансові процеси, оскільки враховує як нелінійності моделей, так і нерегулярність між замірами реальних процесів. Друга модель розділу 4 даної роботи присвячена одному узагальненню LSTM нейронних мереж із врахуванням довжини передісторії нейронної мережі. Основна відмінність від класичної LSTM моделі або багатовимірного варіанту MV-LSTM моделі полягає у наявності гіперпараметру p , який відображає глибину залежності вхідного сигналу (x_t, \dots, x_{t+1-p}) та вихідного сигналу $g(x_{t+1})$ або в більш загальному випадку $g(x_{t+h})$ для $h > 0$.

У висновках підсумовано основні результати дисертаційного дослідження. **У додатках** подано наукові публікації, в яких відображено основні наукові результати роботи.

Теоретичне значення. Результати роботи, отримані в ході наукового дослідження, є удосконаленням теорії часових рядів та теорії машинного навчання. Підходи, запропоновані у дисертації, можуть використовуватися для подальших досліджень у цій галузі, у навчальних курсах кафедри прикладної математики та інформаційних технологій Чернівецького національного університету імені Юрія Федьковича (та інших ЗВО), пов'язаних з інтелектуальним аналізом даних, машинним навчанням, у методичних розробках, навчальних посібниках для освітнього процесу та науково-дослідної роботи студентів та аспірантів.

Практичне значення. Розроблені у дисертаційній роботі алгоритми можуть в подальшому використовуватися для практичного дослідження фінансових часових рядів. Основні результати роботи ґрунтуються на

твердженнях теорії ймовірності, статистики, хаотичних систем, теорії машинного навчання, випадкових процесів та теорії часових рядів.

Практичну цінність роботи проілюстровано на оцінюванні прогнозів реальних фінансових. Результати роботи знайдуть застосування у подальших дослідженнях із даної тематики.

Ключові слова: стохастичні диференціально-функціональні рівняння, генетичний алгоритм, прогнозування фінансових процесів, стохастична оптимізація, нейронні мережі, оцінка параметрів, метод зворотного поширення, часові ряди, штучний інтелект, машинне навчання, моделювання фондових індексів, часовий ряд фінансових даних, організаційна карта Кохонена, адаптивна нейро-нечітка система виводу, оцінка похибки.

ABSTRACT

Tokarieva K.A. Hybrid and ensemble machine learning methods and models for forecasting financial time series. – Qualification scientific work in the form of a manuscript.

Dissertation for the degree of Doctor of Philosophy in speciality 113 – Applied Mathematics. – Yuriy Fedkovych Chernivtsi National University, Ministry of Education and Science of Ukraine, Chernivtsi, 2024.

The dissertation is devoted to the construction of hybrid algorithms based on methods and models of machine learning and classical time series algorithms.

The results of the thesis are the basis for further theoretical and practical research developments in the study of hybrid forecasting algorithms for financial time series.

The thesis consists of an introduction, four chapters, conclusions, a list of references and an appendix.

The introduction substantiates the relevance of the research topic, formulates the purpose, objectives, subject, object and methods of the research, indicates the scientific novelty, theoretical and practical significance of the results obtained, analyses the relationship of the work with scientific research and the

personal contribution of the applicant, and provides information on the testing and publication of the main results of the dissertation. The structure and scope of the dissertation are described.

The first chapter reviews the scientific literature in the main areas of research on time series in general and financial time series in particular, considers key information on time series theory and machine learning methods, describes the main research areas and identifies the tasks that time series theory deals with, including the main problems of building a forecast for financial time series. A detailed overview of the main classical algorithms for building a forecast for a time series is provided, including their scope of application. The chronology of the development of time series theory and conditional periods of model complexity are analysed in detail. The second section of this chapter discusses the main models of artificial intelligence and machine learning for time series analysis, describes the topologies of neural networks that will be used in the construction of hybrid algorithms.

In the second chapter, the author investigate hybrid models in which the main process is estimated on the basis of a time series model. Model residuals are estimated on the basis of neural networks. The use of this approach allowed us to extend the use of classical time series models to the case of the absence of homoscedasticity in the residuals and the absence of non-smoothness of the time series residuals. The main results are:

- time series of stock indices contain both linear and non-linear components, and therefore, a linear ARIMA model and a non-linear ANN model alone cannot provide an accurate estimate of such time series;
- a hybrid model of an autoregressive moving average model and an artificial neural network for forecasting time series of stock indices is developed;
- it is proved that if the time series is strictly stationary, the ARIMA model is more adequate for Gaussian series in the developed hybrid model;
- a comparative analysis of the proposed algorithm with classical algorithms for estimating the forecast of financial time series is carried out, as a result

of which the developed method demonstrates greater accuracy compared to other hybrid models, since the data series is decomposed into components with low and high volatility (linear and nonlinear), and therefore the total error of the model is low.

The third chapter considers the relationship between chaotic dynamical systems and time series systems. The main approaches to the definition of deterministic chaos and the main methods of taking into account the presence of deterministic chaos when assessing the forecast of a financial time series are analysed. Four methods are presented to determine the existence of chaos in the system:

- Fourier transform;
- Lyapunov exponent;
- Hurst exponent;
- fractal dimension.

The main approach to the application of chaos theory, namely deterministic chaos, is to use the Tuckens embedding theorem, according to which the structure of a chaotic deterministic process can be simplified by using a partition of the vector space $(x(t), \dots, x(t + T))$. The phase space reconstruction method is used to reconstruct phase space points from chaotic time series using the embedding dimension D and time delay T . According to the phase space reconstruction method, the phase space points reproduce the key characteristics of the original time series. In order to confirm the effectiveness of the proposed methodology, three different classical chaotic time series were studied in this section: McKay-Glass, logistic and Eno. The proposed method was combined with genetic algorithms to improve the forecasting results. It was shown that the new forecasting method improved with GA demonstrates outstanding results. Comparison of the results of the proposed method with the results reported in the literature substantiates the effectiveness of the proposed method.

In chapter four, the author examines two hybrid models based on time series models. The first model is a continuous and nonlinear analogue of the

classical linear ARMA model, while the second model is a generalisation of the classical LSTM neural network with an arbitrary finite backardation. The first model is based on a new method of estimating the forecast of financial market processes based on stochastic differential equation models with a delay. A genetic algorithm, namely the particle swarm method, was used to estimate the parameters in the developed model, as this method proved to be the most optimal for our problem. The article also presents one of the approaches to the transition from a function class optimisation problem to a parametric optimisation problem, which allows a much more efficient use of genetic algorithms to find optimal parameters. The theoretical results obtained are tested on real data using the dynamics of three real financial processes. The test results show that the new model better approximates real financial processes, as it takes into account both the non-linearity of the models and the irregularity between measurements of real processes. The second model in chapter 4 of this paper is dedicated to a generalisation of LSTM neural networks, taking into account the length of the back-propagation of the neural network. The main difference with the classical LSTM model or the multivariate variant of the MV-LSTM model is the presence of the hyperparameter p , which reflects the depth of dependence of the input signal (x_t, \dots, x_{t+1-p}) and output signal $g(x_{t+1})$ or in a more general case $g(x_{t+h})$ для $h > 0$.

The **conclusions** summarise the main findings of the thesis research, and the **appendices** contain scientific publications that reflect the main scientific findings of the thesis.

Theoretical significance. The results obtained in the course of the research improve the theory of time series and machine learning. The approaches proposed in the thesis can be used for further research in this area, in the courses of the Department of Applied Mathematics and Information Technologies of the Yuriy Fedkovych Chernivtsi National University (and other universities) related to data mining, machine learning, in methodological developments, textbooks for the educational process and research work of students and postgraduates.

Practical significance. The algorithms developed in this thesis can be further used for practical research on financial time series. The main results of the thesis are based on probability theory, statistics, chaotic systems, machine learning theory, random processes and time series theory.

The practical value of the work is illustrated by evaluating forecasts of real financial markets. The results of the work will be used for further research on this topic.

Keywords: stochastic differential functional equations, genetic algorithm, financial forecasting, stochastic optimisation, neural networks, parameter estimation, back-propagation, time series, artificial intelligence, machine learning, stock index modelling, financial time series, Kohonen chart, adaptive neuro-fuzzy inference system, error estimation.

СПИСОК ПУБЛІКАЦІЙ ЗА ТЕМОЮ ДИСЕРТАЦІЇ

Наукові праці у періодичних наукових виданнях, проіндексованих у наукометричній базі даних Scopus:

1. Kushnir, M., Tokarieva, K. A Generalization of the Arima Model to the Nonlinear and Continuous Cases. *Cybernetics and Systems Analysis*. 2023. Vol. 59. Iss. 6. P. 900–909 (Scopus), (Q3 – <https://link.springer.com/article/10.1007/s10559-023-00625-8>)
2. Кушнір М.Я., Токарева К.А. Використання Систем Штучного Інтелекту У Задачах Прогнозування Фінансових Індексів: Огляд Наукових Джерел / Artificial Intelligence Systems In The Financial Market Predictions: Literature Review. *Radioelectronic and Computer Systems*. 2020. Is. 3(95). С. 108-117 (Scopus), (URL: <https://www.scopus.com/record/display.uri?eid=2-s2.0-85101514809&origin=resultslist>)

Наукові праці у виданнях, включених до переліку наукових фахових видань України:

3. Кушнір М.Я., Токарева К.А. Одне узагальнення ARIMA-моделі на нелінійний та неперервний випадки / A Generalization of the Arima Model to the Nonlinear and Continuous Cases. *Cybernetics and Systems Analysis*. 2023. Vol. 59. Iss. 6. P. 900 – 909. (Scopus), (Q3 - URL: <https://www.scimagojr.com/journalsearch.php?q=12933&tip=sid&clean=0>)
4. Кушнір М.Я., Токарева К.А. Гібридна модель самоорганізаційної карти Кохонена та адаптивної нейро-нечіткої системи виводу у задачах прогнозування цін фондових індексів. *Буковинський математичний журнал*. 2021. Т. 9. № 2. С. 70–80. (URL: <https://doi.org/10.31861/bmj2021.02.05>)
5. Кушнір, М. Я., Токарева, К. А. Одне узагальнення LSTM-нейронних мереж. *International Scientific Technical Journal "Problems of Control and Informatics"*. 2023. 68(4). С. 111–118. (URL: <https://jais.net.ua/index.php/files/article/view/118>)

Наукові праці, які засвідчують апробацію матеріалів дисертації:

6. M. Kushnir, A. Komarnytskyi, K. Tokarieva, N. Savchyn, P. Kroialo and V. Toronchuk, Technological and Legal Aspects of the Use of Machine Learning Elements in Chaotic Information Processing Systems. *2020 IEEE International Conference on Problems of Infocommunications. Science and Technology (PIC S&T)*, Kharkiv, Ukraine, 6-9 October 2020, P. 713-716. ISBN: 978-172819177-5 (Scopus) (URL: <https://www.scopus.com/record/display.uri?eid=2-s2.0-85114410889&origin=resultslist>)
7. V. Hryhorkiv, L. Buiak, A. Verstiak, M. Hryhorkiv, O. Verstiak and K. Tokarieva, "Forecasting Financial Time Series Using Combined ARIMA-ANN Algorithm," *2020 10th International Conference on Advanced Computer Information Technologies (ACIT)*, Deggendorf, Germany, 16-18 September 2020, P. 455-458. ISBN: 978-172816760-2 (Scopus), (URL: <https://www.scopus.com/record/display.uri?eid=2-s2.0-85094125855&origin=resultslist>)
8. Токарева К.А. Використання систем штучного інтелекту у прогнозуванні фінансових ринків. *Міжнародний науковий симпозиум «Інтелектуальні рішення»*. Обчислювальний інтелект (результати, проблеми, перспективи): праці міжнар. наук.-практ. конф., 15-20 квітня 2019 р., Ужгород / М-во освіти і науки України, ДВНЗ «Ужгородський національний університет», та [ін.]; наук. ред. В.Є. Снитюк. С. 54-55. (URL: <https://er.chdtu.edu.ua/bitstream/ChSTU/3192/1/OI-2019.pdf>)
9. Токарева К.А. Лог-рентабельність активів у дослідженні фінансових часових рядів. *Математичні методи, моделі та інформаційні технології в економіці*: Матеріали VI Міжнародної науково-методичної конференції. Чернівці: Друк Арт, 6 вересня 2019. С. 154-155.

ЗМІСТ

РОЗДІЛ 1. ТЕОРЕТИКО-МЕТОДОЛОГІЧНІ ОСНОВИ ВИКОРИСТАННЯ СИСТЕМ ШТУЧНОГО ІНТЕЛЕКТУ У АНАЛІЗІ ФІНАНСОВИХ ЧАСОВИХ РЯДІВ	21
1.1. Поняття фінансових часових рядів та статистичних властивостей їх розподілу.....	21
1.2. Системи штучного інтелекту у фундаментальному і технічному аналізі та прогнозуванні фінансових ринків: критичний огляд наукових джерел	28
1.3. Існуючі методи та моделі штучного інтелекту у прогнозуванні фінансових часових рядів.....	34
Висновки до розділу 1.....	47
РОЗДІЛ 2. ГІБРИДНІ МЕТОДИ ТА МОДЕЛІ: ПОРІВНЯЛЬНИЙ АНАЛІЗ	49
2.1. Алгоритм прогнозування часових рядів фондових індексів на основі гібридної моделі ARMA моделі та штучної нейронної мережі	49
2.2. Алгоритм прогнозування часових рядів фондових індексів на основі гібридної моделі авторегресії-ковзної середньої та рекурентної нейронної мережі.....	60
Висновки до розділу 2.....	68
РОЗДІЛ 3. ВИКОРИСТАННЯ МЕТОДІВ МАШИННОГО НАВЧАННЯ ДЛЯ ДОСЛІДЖЕННЯ ХАОТИЧНИХ ФІНАНСОВИХ ЧАСОВИХ РЯДІВ	70
3.1. Огляд теореми про вкладення та штучний інтелект	70

3.2. Основні показники хаотичності системи	73
3.2.1 Хаотичні часові ряди	74
3.2.2 Визначення хаосу в системах	74
3.2.3. Теорема вкладення.....	77
3.2.4. Визначення затримки часу	78
3.2.5. Визначення розмірності вкладення.....	79
3.3. Генетичні алгоритми.....	80
3.4. Використання машинного навчання для прогнозування фінансових хаотичних часових рядів.....	82
3.5. Числовий аналіз за допомогою нейронної мережі прямого поширення	88
Висновки до розділу 3.....	100
РОЗДІЛ 4. УЗАГАЛЬНЕНІ ГІБРИДНІ МОДЕЛІ ЧАСОВИХ РЯДІВ	103
4.1. Неперервний аналог часових рядів та відповідна оптимізаційна задача	103
4.1.1. Мотивація використання часових рядів	103
4.1.2. Основна модель та допоміжний результат.....	107
4.1.3. Метод оцінки невідомих параметрів	110
4.1.4. Аналіз реальних даних	112
4.2. Одне узагальнення LSTM нейронних мереж	117
4.2.1. Історія питання, мотивація.....	117
4.2.2. Узагальнена модель	122
4.2.3. Оцінка коефіцієнтів	123
Висновки до розділу 4.....	130

ВИСНОВКИ	133
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ	137
ДОДАТОК.....	156

ВСТУП

Актуальність теми дослідження. Ускладнення систем є природнім шляхом розвитку будь якої науки включаючи математику в цілому, так і її прикладні аспекти зокрема. З точки зору теорії часових рядів, найпростіші моделі (AR та MA моделі) ґрунтувалися на припущеннях про гомоскедастичність на незалежність залишків моделі. Дані класичні моделі були зручними з точки зору теоретичних викладок та простоти використання на реальних прикладах через простоту оцінок параметрів. Проте як було помічено багатьма дослідниками дана простота з дуже низькою ймовірністю зустрічається в реальному світі, оскільки неможливо позбутися впливу зовнішніх факторів. У зв'язку із цим, теорія часових рядів зосередила свою увагу на комбінованих моделях, в яких основний процес описується з використанням класичних моделей, а “допоміжні процеси” описуються за допомогою інших математичних моделей. Найпростішим прикладом такої комбінації є ARCH, GARCH моделі та їх модифікації, в яких робиться припущення про гетероскедастичність дисперсії σ_t^2 основного процесу $x(t)$, причому σ_t^2 також підпорядковується моделі часового ряду із врахуванням впливу екзогенних змінних та процесу $x(t)$. Оцінка параметрів таких моделей стала можливою як за рахунок швидкого розвитку теорії оптимізації з одного боку, так і за рахунок збільшення обчислювальних потужностей середньостатистичного користувача.

Другом кроком для узагальнення класичних моделей являється використання гібридних моделей, в яких сновний процес описується за допомогою деякої моделі часового ряду, проте залишки можуть описуватися більш складнішими математичними моделями. В даній роботі основна увага звернена на випадок використання моделей машинного навчання в якості моделі опису залишків. Як буде показано нижче, вибраний підхід добре себе зарекомендував з точки зору задачі прогнозування часових рядів. В якості нейронних мереж було вибрано декілька основних підкласів, які найчастіше

зустрічаються в літературі для прогнозування часових рядів, а саме RNN, LSTM, нейронна мережа прямого поширення помилки, нейронна мережа Елмана, багат шаровий перцептрон. Крім того, для візуалізації багатовимірних даних шляхом конфігурації нейронів з метою квантування або кластеризації вхідного простору в топологічну структуру використувувалась самоорганізаційна картка Кохонена. Оскільки при використанні відсутнє припущення про те, що одна модель здатна «охопити динаміку» всього часового ряду, то здійснено її інтеграцію з адаптивною нейро-нечіткою системою виводу. Дані моделі дали змогу достатньо повно описати загальний підхід до побудови гібридних моделей часових рядів на основі нейронних мереж.

З іншого боку, достатньо часто із наявністю випадкових чинників, динамічні системи підпорядковані впливу детермінованого хаосу, за наявності якого детермінований процес буде мати поведінку до випадкового. У зв'язку з цим у дисертаційній роботі розроблено підхід до визначення хаотичної складової у часовому ряді для покращення оцінки прогнозу.

Отже, дисертаційна робота спрямована на побудову гібридних моделей на основі нейронних мереж для опису динаміки залишків та використання методів визначення наявності детермінованого хаосу в системі для зменшення впливу детермінованих чинників у системі.

Об'єктом дослідження є задача знаходження прогнозу часового ряду для реальних фінансових процесів, що описують динаміку випадкових фінансових активів (акцій, облігацій, дериватив).

Предметом дослідження є побудова оптимального прогнозу на основі використання гібридних систем.

Мета дослідження полягає у розробці гібридних систем з використання нейронних мереж як складової гібридної системи; у проведенні порівняльного аналізу між існуючими та розробленими алгоритмами; у розробці теоретичного підґрунтя для розробки узагальнень деяких класичних нейронних мереж з використанням теорії часових рядів.

Методи дослідження. В основу дисертаційної роботи покладені методи часових рядів; методи машинного навчання та штучного інтелекту, які включають в себе нейронні мережі, метод опорних векторів; статистичні метод оцінок параметрів; теорія стохастичних диференціальних рівнянь та теорія оптимального керування; генетичні алгоритми.

Зв'язок роботи з науковими програмами, планами, темами. Тема дисертаційної роботи відповідає науковому напрямку кафедри прикладної математики та інформаційних технологій Чернівецького національного університету імені Юрія Федьковича. Дослідження дисертаційної роботи проведено в рамках науково-дослідної теми кафедри «Математичне моделювання і числово-аналітичні методи дослідження динамічних та інформаційних процесів» (номер державної реєстрації 0102U006591).

Наукова новизна отриманих у роботі результатів полягає в тому, що для покращення оцінки прогнозу часового ряду фінансових активів уперше:

- розглянуто гібридну систему, опис динаміки залишків якої описується на основі моделей машинного навчання, а саме нейронних мереж;
- розглянуто різні моделі нейронних мереж як кандидати для побудови оптимального прогнозу;
- проведено порівняльний аналіз існуючих класичних алгоритмів оцінки прогнозу та розроблених гібридних алгоритмів на реальних даних фінансового ринку;
- проведено аналіз гібридних систем із врахуванням наявності детермінованого хаосу в часовому ряді та розроблено алгоритм оцінки детермінованого хаосу в часових рядах
- розроблено алгоритм зменшення складності оцінки прогнозу часового ряду із використанням генетичного алгоритму для пошуку початкових значень ваг нейронної мережі;
- розроблено розширену LSTM модель нейронної мережі із врахуванням залежності типу AR;

- розроблено алгоритми оцінки параметрів часового ряду на основі стохастичних диференціальних рівнянь із запізненням та здійснено перехід від непараметричної оптимізаційної задачі до параметричної оптимізаційної задачі пошуку коефіцієнтів часового ряду або відповідного стохастичного диференціального рівняння.

Практичне значення отриманих результатів. Результати дисертаційної роботи є вагомим внеском у розвиток теорії часових рядів та побудови гібридних алгоритмів. Розроблені та апробовані запропоновані алгоритми можуть використовуватися для подальших досліджень у цій галузі, адже на сьогодні немає універсального способу побудови прогнозу часових рядів, зокрема фінансових. Цю проблему поставлено в дисертаційній роботі, простежено, що розроблені генетичні алгоритми оцінки прогнозу часового ряду дають кращі результати порівняно з існуючими класичними алгоритмами як для реальних задач, так і для класичних теоретичних динамічних систем.

Практичне значення дисертаційної роботи полягає в тому, що розроблений гібридні системи можуть бути використані в прогностичних системах для аналізу фінансових ринків з метою побудови більш чітких прогнозів.

Особистий внесок здобувача. У дисертаційній роботі зроблено значний внесок у розв'язання конкретних наукових завдань. Зокрема:

- побудовано ряд гібридних моделей із використанням нейронних мереж, як складової моделі для аналізу динаміки залишків часового ряду;
- розроблено підхід оцінки коефіцієнтів стохастичних диференціальних рівнянь для часових рядів із нерівномірними вимірюваннями;
- розроблено новий підхід для оцінки наявності детермінованого хаосу в часовому ряді на основі одного із чотирьох показників (показника Ляпунова, розкладу Фур'є, показника Герста та фрактальної розмірності процесу);

- проведено порівняльний аналіз розроблених алгоритмів із уже існуючими алгоритмами та доведено їх ефективність при побудові прогнозу фінансових часових рядів;

- проведено аналіз реальних ризикових активів на основі аналізу в Python та Matlab.

Апробацію матеріалів дисертації здійснено на таких конференціях:

1. IEEE International Conference on Problems of Infocommunications. Science and Technology (PIC S&T), Kharkiv, Ukraine, 6-9 October 2020.
2. 10th International Conference on Advanced Computer Information Technologies (ACIT), Deggendorf, Germany, 16-18 September 2020.
3. Міжнародний науковий симпозіум «Інтелектуальні рішення». Обчислювальний інтелект (результати, проблеми, перспективи), м. Ужгород, Україна, 15-20 квітня 2019 р.,
4. VI Міжнародна науково-методична конференції, м. Чернівці, Україна, 6 вересня 2019 р.

Структура й обсяг дисертації. Дисертаційна робота є завершеним науковим дослідженням, загальним обсягом 157 сторінок (з них: 117 сторінки основного тексту, 18 сторінок – література). Дисертація складається зі вступу, чотирьох розділів з підрозділами, висновків, списку використаних джерел (158 позицій) та додатку (список публікацій здобувачки за темою дисертації).

РОЗДІЛ 1. ТЕОРЕТИКО-МЕТОДОЛОГІЧНІ ОСНОВИ ВИКОРИСТАННЯ СИСТЕМ ШТУЧНОГО ІНТЕЛЕКТУ У АНАЛІЗІ ФІНАНСОВИХ ЧАСОВИХ РЯДІВ

1.1. Поняття фінансових часових рядів та статистичних властивостей їх розподілу

На сьогодні фінансові ринки виступають визначальним елементом економічних та соціальних систем сучасного суспільства. Крім того, фінансова діяльність відіграє важливу роль у світовій економіці, впливаючи на економічний розвиток більшості країн світу [3]. На фінансових ринках успіх інвестора залежить від якості інформації, яку він використовує для підтримки прийняття рішень, а також від того, як швидко він здатний приймати рішення. Тому, завдяки своєму практичному значенню, аналіз фінансових ринків в останні десятиріччя почав широко вивчатися науковцями з галузі математики, комп'ютерних наук та інженерії [4]. Прогнозування фінансових часових рядів можна вважати однією з основних задач у науковій літературі, метою яких є вивчення часових рядів [1, 5]. Нагадаємо, що часовий ряд – це набір спостережень x_t , визначених в моменти часу $t = t_1, t_2, \dots, t_N$, тобто це ряд точок даних, проіндексованих в хронологічному порядку.

Аналіз часових рядів передбачає використання численних методів їх аналізу з метою визначення значимих статистик та інших характеристик даних. Прогнозування часових рядів – це використання широкого класу моделей з метою передбачення майбутніх їх значень на основі історичних даних.

Зауважимо, що дослідження фінансових ринків передбачає використання, так званої, рентабельності або прибутковості активів (англ. Return on assets), а не виключно їх цін. В публікації [8] наводяться дві

головних причини використання прибутковості, перша з яких полягає в тому, що для більшості інвесторів рентабельність є головною оцінкою інвестиційних можливостей, а друга – в тому, що рентабельність має більш «привабливі» статистичні характеристики, ніж ціни активів. Існує декілька визначень рентабельності фінансових активів. Визначимо їх детально.

Нехай, P_t - ціна активу в момент часу t . Тоді можна виділити так типи рентабельності (R_t) фінансових активів.

1. *Одноперіодна проста рентабельність*. При визначенні активу для одного період від часу $t-1$ до часу t отримаємо просту валову рентабельність:

$$1 + R_t = \frac{P_t}{P_{t-1}} \text{ або } P_t = P_{t-1}(1 + R_t).$$

Відповідно, одноперіодна проста чиста рентабельність матиме вигляд:

$$R_t = \frac{P_t}{P_{t-1}} - 1 = \frac{P_t - P_{t-1}}{P_{t-1}} \approx \ln\left(\frac{P_t}{P_{t-1}}\right).$$

2. *Багатоперіодна проста рентабельність*. При визначенні активу для k періодів між датами $t-k$ та t отримаємо k -періодну просту валову рентабельність:

$$\begin{aligned} 1 + R_t[k] &= \frac{P_t}{P_{t-k}} = \frac{P_t}{P_{t-1}} \times \frac{P_{t-1}}{P_{t-2}} \times \dots \times \frac{P_{t-k+1}}{P_{t-k}} = (1 + R_t)(1 + R_{t-1}) \dots (1 + R_{t-k-1}) = \\ &= \prod_{j=0}^{k-1} (1 + R_{t-j}). \end{aligned}$$

Відповідна k -періодна проста чиста рентабельність має вигляд $R_t[k] = (P_t - P_{t-k}) / P_{t-k}$.

На практиці, фактичний часовий інтервал є важливим для аналізу рентабельності (наприклад, щомісячна або річна рентабельність) з метою прийняття відповідних ефективних управлінських рішень. Якщо часовий інтервал не задано, то неявно припускається, що він відповідає одному року. Якщо актив аналізується за k років, то розрахунок рентабельності на річній основі (середньої) виглядатиме як (його ще називають аннуїлізованим):

$$\{R_t[k]\} = \left[\prod_{j=0}^{k-1} (1 + R_{t-j}) \right]^{1/k} - 1.$$

Очевидно, що це геометрична середня k -одноперіодних простих валових рентабельностей і може бути розрахована як:

$$\{R_t[k]\} = \exp \left[\frac{1}{k} \sum_{j=0}^{k-1} \ln(1 + R_{t-j}) \right] - 1.$$

Оскільки, арифметичне середнє розрахувати легше, аніж геометричне, а також одноперіодна рентабельність є малою, то можна використати ряд Тейлора першого порядку для апроксимації аннулізованої рентабельності:

$$\{R_t[k]\} \approx \frac{1}{k} \sum_{j=0}^{k-1} R_{t-j}.$$

2. *Рентабельність неперервного нарахування складних відсотків* (компаундинг або логарифм рентабельності, лог-рентабельність) виражається як:

$$r_t = \ln(1 + R_t) = \ln \frac{P_t}{P_{t-1}} = p_t - p_{t-1},$$

де $p_t = \ln(P_t)$. Рентабельність неперервного нарахування складних відсотків r_t має певні переваги над простою чистою рентабельністю R_t . Так, розглянемо багатоперіодну рентабельність:

$$\begin{aligned} r_t |k| &= \ln(1 + R_t |k|) = \ln \left[(1 + R_t)(1 + R_{t-1}) \dots (1 + R_{t-k+1}) \right] = \\ &= \ln(1 + R_t) + \ln(1 + R_{t-1}) + \dots + \ln(1 + R_{t-k+1}) = r_t + r_{t-1} + \dots + r_{t-k+1} = \\ &= r_t + r_{t-1} + \dots + r_{t-k+1}. \end{aligned}$$

Таки чином, багатоперіодна рентабельність неперервного нарахування складних відсотків є сумою одноперіодних рентабельностей неперервного нарахування складних відсотків. Також, знаходження статистичних характеристик логарифмів рентабельностей більш прості у розрахунках.

3. *Рентабельність портфеля інвестицій*. Проста чиста рентабельність портфеля інвестицій, що складається з N активів, є середньозваженою величиною простої чистої рентабельності активів, в якій вага кожного активу

становить відсоткову частку портфеля інвестицій, інвестованих в цей актив. Позначимо через p портфель інвестицій, який має вагу w_i в активі i . Тоді проста рентабельність портфеля p в час t матиме вигляд:

$$R_{p,t} = \sum_{i=1}^N w_i R_{it},$$

де R_{it} - проста рентабельність активу i .

4. *Виплата дивідендів.* Якщо актив передбачає періодичну виплату дивідендів, їх рентабельність такого активу потрібно модифікувати. Позначимо через D_t виплату дивідендів по активу в період з $t-1$ по t , а через P_t - ціну активу в кінці періоду t . В такому випадку дивіденди не включаються в P_t . Тоді проста чиста рентабельність та рентабельність неперервного нарахування складних відсотків в час t матиме відповідний вигляд:

$$R_t = \frac{P_t + D_t}{P_{t-1}} - 1, \quad r_t = \ln(P_t + D_t) - \ln(P_{t-1}).$$

5. *Надлишкова дохідність (рентабельність)* активу в момент часу t виражається різницею між дохідністю даного активу та прибутковістю деякого базового активу. Проста чиста проста надлишкова рентабельність та надлишкова рентабельність неперервного нарахування складних відсотків в час t матиме відповідний вигляд:

$$Z_t = R_t - R_{0t}, \quad z_t = r_t - r_{0t},$$

де R_{0t} та r_{0t} , відповідно, проста та рентабельність неперервного нарахування складних відсотків базового активу (начастіше безризикового активу).

Очевидно, що взаємозв'язок між простою рентабельністю R_t та лог-рентабельністю r_t задається через співвідношення:

$$r_t = \ln(1 + R_t), \quad R_t = e^{r_t} - 1.$$

Якщо рентабельності R_t та r_t виражені у відсотках, тоді

$$r_t = 100 \ln \left(1 + \frac{R_t}{100} \right), \quad R_t = 100 (e^{r_t/100} - 1).$$

Розподіл часових рядів фінансових активів (в нашому випадку рентабельності) теж має свої особливості. Найбільш загальною моделлю логарифмів рентабельностей $\{r_{it}; i=1, \dots, N; t=1, \dots, T\}$ виступає функція спільного розподілу:

$$F_r(r_{11}, \dots, r_{N1}; r_{12}, \dots, r_{N2}; \dots; r_{1T}, \dots, r_{NT}; \mathbf{Y}; \boldsymbol{\theta}), \quad (1.1)$$

де \mathbf{Y} - вектор змінних, які описують середовище рентабельностей активів, а $\boldsymbol{\theta}$ - вектор параметрів функції розподілу $F_r(\cdot)$.

Розподіл ймовірностей $F_r(\cdot)$ визначає стохастичну поведінку рентабельності $\{r_{it}\}$ та \mathbf{Y} . У фінансовій літературі вектор \mathbf{Y} трактується як даний, тому емпіричний аналіз рентабельності активів полягає у оцінці невідомого параметру $\boldsymbol{\theta}$ та статистичної поведінки $\{r_{it}\}$ при заданих історичних даних рентабельності.

Деякі наукові джерела з фінансів здійснюють фокус на спільному розподілі N рентабельностей в один момент часу t , тобто розподілу $\{r_{1t}, \dots, r_{Nt}\}$. Також, можна зустріти акцент на динамічній структурі рентабельності індивідуальних активів, тобто розподілу $\{r_{i1}, \dots, r_{iT}\}$ для заданого активу i . Однак, найбільше досліджується спільний розподіл $\{r_{it}\}_{t=1}^T$ для активу i . При цьому варто здійснити перетворення спільного розподілу (1.1) як:

$$\begin{aligned} F(r_{i1}, \dots, r_{iT}; \boldsymbol{\theta}) &= F(r_{i1})F(r_{i2} | r_{i1}) \dots F(r_{iT} | r_{i,t-1}, \dots, r_{i1}) = \\ &= F(r_{i1}) \prod_{t=2}^T F(r_{it} | r_{i,t-1}, \dots, r_{i1}), \end{aligned} \quad (1.2)$$

де з метою спрощення параметр $\boldsymbol{\theta}$ опускається.

Такий поділ підкреслює часову залежність логарифмів рентабельності r_{it} . В такому випадку головною проблемою виступає визначення умовного розподілу $F(r_{it} | r_{i,t-1}, \cdot)$, зокрема, як він змінюється в часі. У фінансах різні характеристики розподілу призводять до різних гіпотез. Наприклад, одна з версій гіпотези випадкової зміни цін активів полягає в тому, що умовний

розподіл $F(r_{it} | r_{i,t-1}, \dots, r_{i1})$ відповідає безумовному розподілу $F(r_{it})$. В такому випадку рентабельності є тимчасово незалежними, а отже, не передбачуваними.

Зауважимо, що рентабельність активів, також, прийнято розглядати як неперервну випадкову величину (особливо у випадку аналізу фондових індексів з низькою періодичністю, як у нашому випадку) та використовувати при цьому щільність розподілу. В такому випадку, з (1.2) маємо:

$$F(r_{i1}, \dots, r_{iT}; \theta) = F(r_{i1}; \theta) \prod_{t=2}^T F(r_{it} | r_{i,t-1}, \dots, r_{i1}; \theta).$$

Із даного співвідношення випливає, що умовний розподіл більш доречний, ніж відособлений при вивченні рентабельності фінансових активів. Однак, безумовні розподіли все ще можуть викликати певний інтерес, зокрема, в такому розподілі легше оцінити граничні розподіли, ніж умовні розподіли, використовуючи історичні часові ряди. Крім того, у деяких випадках доходність активів має слабку емпіричну автокореляцію, а, отже, їх граничні розподіли близькі до їх умовних розподілів.

Загалом у науковій літературі зустрічається декілька статистичних розподілів рентабельності фінансових активів, включаючи нормальний розподіл, логнормальний розподіл, стійкий розподіл та скінченна суміш нормальних розподілів.

В деяких дослідженнях [9, 10] використовується скінченна суміш нормальних розподілів. Прикладом такого розподілу може бути:

$$r_t \sim (1 - X)N(\mu, \sigma_1^2) + XN(\mu, \sigma_2^2),$$

де X - випадкова величина Бернуллі ($P(X = 1) = \alpha$ та $P(X = 0) = 1 - \alpha$ з $0 < \alpha < 1$).

У фінансовій літературі існує також поняття багатовимірної рентабельності. Позначимо через $\mathbf{r}_t = (r_{1t}, \dots, r_{Nt})'$ лог-рентабельність N активів в момент часу t , а через $\{r_t\}_{t=1}^T$ відповідний спільний розподіл, який може

бути заданий аналогічно до (1.2). Аналіз рентабельності в такому випадку зосереджений на специфікації функції умовного розподілу $F(\mathbf{r}_t | \mathbf{r}_{t-1}, \dots, \mathbf{r}_1, \theta)$. Вектор математичного сподівання та коваріаційна матриця випадкового вектора $\mathbf{X} = (X_1, \dots, X_p)$ визначається як:

$$E(\mathbf{X}) = \boldsymbol{\mu}_x = [E(X_1), \dots, E(X_p)]',$$

$$Cov(\mathbf{X}) = \Sigma_x = E[(\mathbf{X} - \boldsymbol{\mu})(\mathbf{X} - \boldsymbol{\mu}_x)'],$$

Завершуючи теоретичне визначення рентабельності фінансових активів зазначимо, що співвідношення (1.2) може бути використано для отримання функції правдоподібності лог-рентабельностей $\{r_1, \dots, r_T\}$ відповідних активів. Якщо умовний розподіл $f(r_t | r_{t-1}, \dots, r_1, \theta)$ є нормальним з математичним сподіванням μ_t та дисперсією σ_t^2 , тоді θ складається з параметрів в μ_t та σ_t^2 , а функція правдоподібності має вигляд:

$$f(r_1, \dots, r_T; \theta) = f(r_1; \theta) \prod_{t=2}^T \frac{1}{\sqrt{2\pi\sigma_t}} \exp\left[-\frac{(r_t - \mu_t)^2}{2\sigma_t^2}\right],$$

де $f(r_1; \theta)$ - щільність розподілу першого спостереження r_1 . Значення θ , що максимізує цю функцію правдоподібності, є оцінкою максимальної правдоподібності θ . Оскільки лог-функція є монотонною, то оцінка максимальної правдоподібності може бути отримана шляхом максимізації логарифмічної функції правдоподібності, яке легше оцінювати на практиці:

$$\ln f(r_1, \dots, r_T; \theta) = \ln f(r_1; \theta) - \frac{1}{2} \sum_{t=2}^T \left[\ln(2\pi) + \ln(\sigma_t^2) + \frac{(r_t - \mu_t)^2}{\sigma_t^2} \right].$$

Логарифмічна функція правдоподібності може бути отримана аналогічним шляхом, якщо умовний розподіл $f(r_t | r_{t-1}, \dots, r_1; \theta)$ не є нормальним.

Таким чином, прийняття важливих фінансових рішень та здійснення відповідних прогнозів базується на використанні широкого класу статистичних моделей. Аналіз часових рядів передбачає використання

численних методів їх аналізу з метою визначення значимих статистик та інших характеристик даних. Однак, фінансові часові ряди володіють високою дисперсією, є складними, динамічними, нелінійними, непараметричними, а також мають хаотичну природу. Тому, на противагу вищеописаній статистичній методиці, при моделюванні та прогнозуванні фінансових часових рядів широко використовуються методи машинного навчання, які детально розглянемо нижче.

1.2. Системи штучного інтелекту у фундаментальному і технічному аналізі та прогнозуванні фінансових ринків: критичний огляд наукових джерел

Відомо, що одна із задач прогнозування фінансових часових рядів полягає у вивченні ряду даних:

$$\{(\mathbf{x}_1, y_1), \dots, (\mathbf{x}_n, y_n)\}, \quad (1.3)$$

де $\mathbf{x}_i \in X$ - набір фактичних спостережень, $y_i \in \mathbb{R}$ - прогнозне значення.

З метою побудови та аналізу (1.3) на сьогодні розроблено велику кількість моделей прогнозування фінансових ринків, які ґрунтуються на використанні статистичних методів: регресійні моделі, вищезгадані авторегресійні моделі, методи експоненційного згладжування, методи максимальної правдоподібності тощо. При цьому всі підходи передбачають моделювання фінансових даних у вигляді часових рядів.

Однак, в останні роки виокремилось два головних напрямки підходів до прогнозування фінансових фондових індексів та розробки відповідних систем підтримки прийняття рішень [7, 11, 34], які умовно можна поділити на статистичні моделі та методи машинного навчання. Традиційні статистичні методи передбачають, що досліджувані часові ряди генеруються

лінійним процесом [12] та, відповідно, з їх допомогою моделюється процес поведінки часових рядів з метою прогнозування їх майбутніх значень (1.3).

Розвиток другого напрямку пов'язаний з тим, що фінансові часові ряди за своєю суттю володіють високою дисперсією, є складними, динамічними, нелінійними, непараметричними, а також мають хаотичну природу [13]. На противагу вищеописаній статистичній методиці, методи машинного навчання мають численні приклади відносно успішного їх використання у моделюванні та прогнозуванні фінансових часових рядів [14], оскільки більшість методів машинного навчання здатні враховувати нелінійні зв'язки між відповідними факторами без попередніх знань про вхідні дані [15].

Оскільки, тільки деякі джерела [14-15] вказують на відносно успішне використання методів штучного інтелекту у прогнозуванні часових рядів, то нами був проведений детальний аналіз відповідної наукової літератури в наукометричних базах даних Scopus, Science Direct, Google Scholar, IEEEExplore та Springer. В розглянутій літературі присутня мала кількість оглядових статей, які присвячені аналізу використання методів штучного інтелекту у проблематиці фінансового прогнозування. Більшість цих статей досліджують штучні нейронні мережі (*англ. artificial neural networks, ANN*) або генетичні алгоритми (*англ. genetic algorithm, GA*). Зокрема, в праці [16] запропоновано емпіричний метод для побудови автоматичної торговельної системи із використанням ANN; авторами описано ключові кроки використання штучної нейронної мережі, а саме: вибір шару входу та виходу, визначення нейронної архітектури та спосіб обробки вихідних сигналів, згенерованих нейронною мережею, з метою створення правил входу і виходу з торгівлі. В праці [17] розглянуто застосування нейронних мереж на певних типах фінансових ринків. Автори детально дослідили деякі первинні дослідження, які застосовують метод ANN до прогнозування обмінних курсів, фондових ринків та банківської (фінансової) кризи. Однак, у даному короткому дослідженні відсутній детальний опис методів ANN, які при цьому використовувались. У дослідженні [20] автори використовують метод

ANN у прогнозуванні фінансових часових рядів; в даній роботі автори розглянули літературу в двох аспектах: механізми попередньої обробки даних, що застосовуються до вхідних даних, та безпосередня нейронної мережі, що використовується в первинних дослідженнях. С. Соні [110], також, досліджував використання методів ANN у прогнозуванні фондових ринків, в якій наводиться короткий опис історії цієї методики, основні поняття фондового ринку, а також перелік деяких первинних досліджень, які використовують підхід ANN для вирішення проблеми прогнозування фондового ринку.

Однак, незважаючи на те, що вищеперераховані праці детально пов'язані з об'єктом нашого дослідження, вони мають обмежену сферу застосування оскільки зосереджені лише на одному сімействі алгоритмів (ANN) штучного інтелекту. Одна з останніх публікацій [15] присвячена аналізу фундаментальних досліджень, які використовують вищезгаданий метод ANN в різних сферах обліку та аудиту, кредитування, фінансового аналізу, інфляції, маркетингу та ін. Інші публікації [23, 26] теж ґрунтуються на аналізі тільки певних методів штучного інтелекту, таких як генетичні алгоритми, генетичне програмування, еволюційні алгоритми.

Фундаментальний аналіз передбачає, що ціна акцій компанії формується на основі політичних та економічних факторів, які є внутрішніми або зовнішніми для неї. В праці [24] було розроблено декілька кількісних інструментів та індикаторів, щоб допомогти у вивченні основ компанії, таких як політика управління, маркетингова стратегія, інноваційні продукти, фінансові коефіцієнти. З метою отримання більш якісного аналізу у вивченні фундаментальних властивостей компанії використовувались деякі інші фактори, такі як ринкові тенденції, законодавство і навіть фінансові новини, веб-сайти і соціальні мережі.

Прихильниками технічного аналізу розроблено велику кількість методів та моделей, однак всі вони ґрунтуються на аналізі часових рядів та виявленні трендів, на основі яких можна спрогнозувати поведінку ціни в майбутньому.

Говорячи про технічний аналіз, варто виділити ґрунтовне дослідження [27], в якому проводиться детальний аналіз наукової літератури, що присвячена даній тематиці. Важливою перевагою використання технічного аналізу є те, що він зводить проблему прогнозування фондових ринків до проблеми розпізнавання паттернів, в яких вхідні дані формуються з історичних цін та технічних індексів, а виходи оцінюються з цих же минулих даних, тобто здійснюється аналіз часових рядів.

Не дивлячись на те, що на сьогодні існує певна кількість розроблених моделей прогнозування фондових індексів із використанням алгоритмів штучного інтелекту, у науковій літературі відсутня усталена методології, яка визначає головні елементи та етапи алгоритму прогнозування фінансових часових рядів [28]. Якщо узагальнювати, то перший етап підготовки даних полягає у їх первинній обробці для процесу навчання. Основними напрямками діяльності на цьому етапі є збір даних, їх попередня обробка, нормалізація та структурування, в результаті чого отримується набір навчальних та тестових даних. Другий етап визначення алгоритму, полягає у визначенні методів штучного інтелекту, які будуть використовуватися для моделювання та прогнозування даних. Етап навчання полягає у виборі алгоритму навчання, налаштуванні параметрів навчання та виконання процедури навчання з метою створення моделі прогнозування часових рядів. Останній етап оцінки полягає у визначенні метрик оцінки та розрахунку точності результатів, отриманих під час виконання навчання на тестовому наборі даних [29, 93].

Однак, до методології фінансового прогнозування повинно бути додано два додаткових етапи: 1) стратегія торгів та 2) оцінка прибутковості. Стратегія торгів полягає у веденні переговорів на реальних ринках, базуючись на прогнозних значеннях відповідних фінансових змінних. Згідно з [31] стратегія торгів передбачає: входження і виходження з торгів, визначення механізмів контролю ризиків, а також управління фінансовими активами. Інший доданий етап полягає у оцінці прибутковості на реальних

ринках, що є надзвичайно важливим у прийнятті управлінських рішень, а також автоматизації процесів здійснення торгів. На рис. 1.1 зображено модифікований алгоритм.

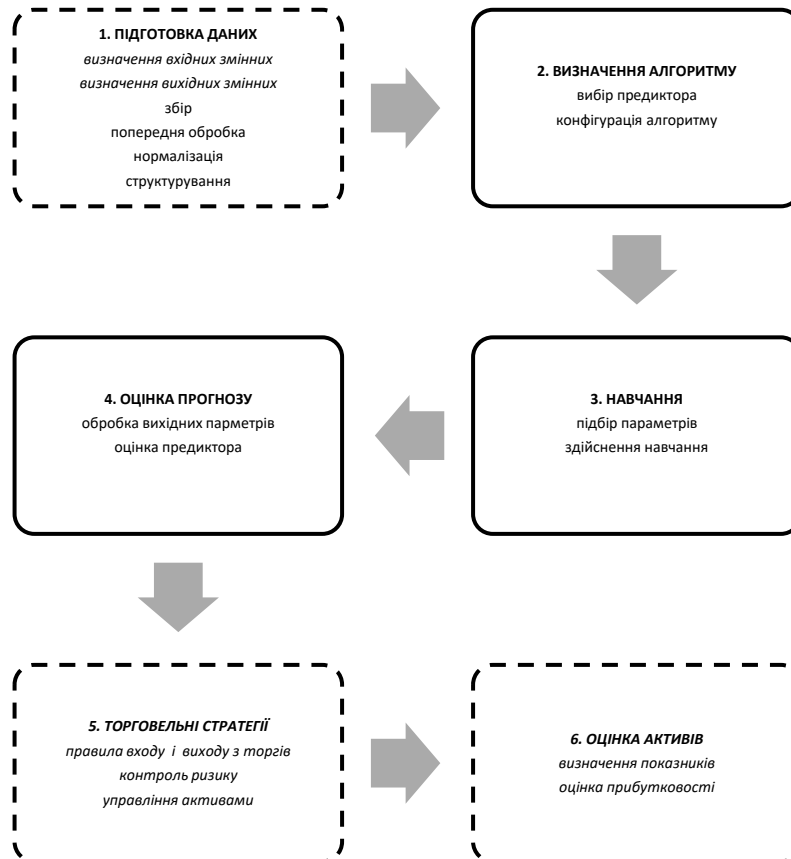


Рис. 1.1. Етапи прогнозування фінансових часових рядів на основі алгоритмів штучного інтелекту

Відповідно до алгоритму, зображеного на рис. 1.1, першочерговим є визначення відповідних змінних, які будуть використовуватись у прогнозуванні фінансових ринків. Запропоновані підходи в науковій літературі різняться щодо того які типи та кількість змінних повинна використовуватись у моделюванні фінансових ринків.

Зазначимо, що відповідно до [35] важливим на етапі підготовки даних є вибір ознак, тобто процес вибору підмножини відповідних змінних (ознак, атрибутів) з метою їх використання при побудові моделі. Алгоритм відбору ознак передбачає:

- спрощення моделей з метою їх легшої інтерпретації;
- зменшення тривалості навчання;
- зниження рівня перенавчання [39].

Головний зміст застосування алгоритму вибору ознак полягає у тому, що для даних, які містять багато ознак (які є або надлишковими, або недоречними) можна усунути ці ознаки без зниження якості інформації [39]. Поняття надлишкових та недоречних ознак різними, оскільки, наприклад доречна ознака може бути надлишковою за наявності іншої доречної ознаки, з якою вона корелює [101]. Літературі зустрічаються різні способи вибору ознак: метод опорних векторів [14], метод головних компонент, генетичний алгоритм, дерева рішень [47] та ін.

Варто відмітити, що на етапі попередньої обробки даних, потрібно здійснювати виявлення аномалій, тобто знаходження та ідентифікації елементів, даних, спостережень, які не відповідають заданому, так званому патерну (іншим елементам набору даних). Для виявлення аномалій в літературі використовуються різноманітні методи: аналіз незалежних компонент, метод k -найближчих сусідів, нечітка логіка, асоціативні правила та ін. [48, 95].

Відповідна структуризація даних на етапі їх попередньої обробки реалізується через класи методів класифікації та кластерного аналізу. Для цього науковці використовують такі методи: виявлення паралельної динаміки [49], методи нечіткої логіки [50], локальне лінійне вкладення LLE [51], самоорганізаційна карта Кохонена [52, 152].

Як зазначалося вище, прогнозування фінансових часових рядів можна вважати однією з основних задач у науковій літературі, метою яких є вивчення часових рядів [1, 5], а завдяки своєму практичному значенню, аналіз фінансових ринків в останні десятиріччя почав широко вивчатися науковцями з галузі математики, комп'ютерних наук та інженерії [4]. Фінансові часові ряди за своєю суттю володіють високою дисперсією, є складними, динамічними, нелінійними, непараметричними, а також мають

хаотичну природу, тому для їх прогнозування широко використовуються методи інтелектуального аналізу великих даних та м'яких обчислень (нечіткі множини, нечітка логіка, нечіткі регулятори, нечіткі нейронні мережі, генетичні алгоритми і еволюційне моделювання тощо). В літературі описані численні алгоритми штучного інтелекту, які використовуються у прогнозуванні фінансових часових рядів, тому їх варто розглянути детальніше.

1.3. Існуючі методи та моделі штучного інтелекту у прогнозуванні фінансових часових рядів

Комплексний аналіз наукової літератури прогнозування фінансових часових рядів із використанням алгоритмів штучного інтелекту дозволив виділити такі класи моделей:

- 1) штучні нейронні мережі (ANN);
- 2) метод опорних векторів (SVM);
- 3) гібридні системи, тобто поєднання декількох підходів;
- 4) оптимізаційні методи у поєднанні із машинним навчанням;
- 5) ансамблеві методи (*англ. ensemble methods*);
- 6) інтелектуальний аналіз тексту;
- 7) специфічні методи, орієнтовані на конкретну задачу.

Як зазначалось в попередньому підрозділі в останні роки найбільш популярним підходом машинного навчання до прогнозування фінансових часових рядів є штучні нейронні мережі, оскільки з допомогою цих моделей зручно здійснювати обробку даних, які характеризуються нелінійністю, розривами, змінною волатильністю (гетероскедастичністю) тощо. Такі методи керуються самими даними та є самоадаптивними, що дозволяє включати нелінійні поведінки часових рядів без будь-яких статистичних припущень щодо даних [95]. В літературі присутні декілька видів цього класу моделей, метою яких є прогнозування фінансових часових рядів.

Більшість таких робіт успішно використовували навчання багатошарового процептрона Румельхарта методом (MLP) зворотного поширення помилки.

Наприклад, Чен та Ду [23] застосували такий підхід для вивчення певних фінансових коефіцієнтів, з метою класифікації компаній, які мають фінансові труднощі. Так, автори наводять $l - m - n$ (через l позначено вхідні нейрони, m – приховані нейрони, n – вихідні нейрони) архітектуру в методі зворотного поширення помилки.

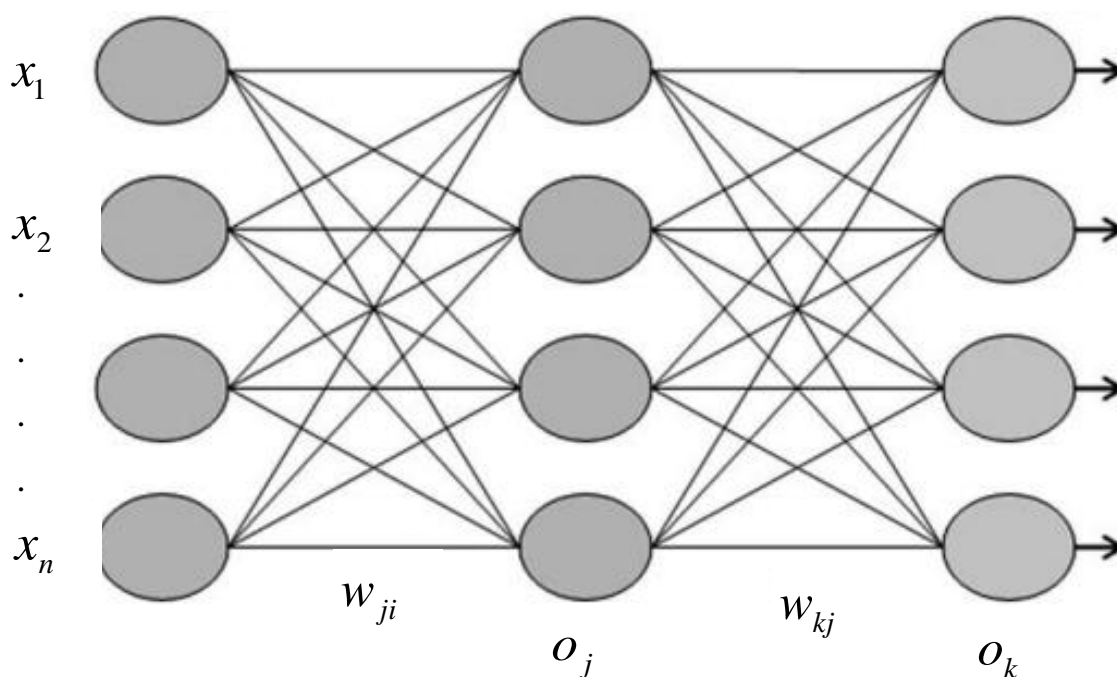


Рис. 1.2. Архітектура мережі зворотного поширення помилки

Прихований шар визначає співвідношення відображення між вхідними (x_1, \dots, x_n) та вихідними шарами, тоді як співвідношення між нейронами зберігаються як ваги вузлів. Вхідні сигнали змінюються за допомогою ваги зв'язку w_{ij} , який відображає зв'язок i -го вузла першого шару з j -м вузлом другого шару. Сума модифікованих сигналів модифікується сигмоїдною передавальною функцією (f). Аналогічно, вихідні сигнали прихованого шару модифікуються за допомогою ваги зв'язку w_{kj} k -го вузла вихідного шару з j -м вузлом прихованого шару. Сума модифікованих сигналів потім змінюється

сигмоїдною передавальною функцією (f), а і виходи збираються на вихідному шарі.

Позначимо через $I_p = (I_{p1}, I_{p2}, \dots, I_{pl})$, $p = 1, 2, \dots, N$ p -й патерн з поміж N вхідних патернів. Тоді:

- вихід з нейрона вхідного шару

$$O_{pi} = I_{pi}, \quad i = 1, 2, \dots, l;$$

- вихід з нейрона прихованого шару:

$$O_{pj} = f(NE_{pj}) = f\left(\sum_{i=0}^l w_{ij} o_{pi}\right), \quad j = 1, 2, \dots, m;$$

- вихід з нейрона вихідного шару:

$$O_{pk} = f(NE_{pk}) = f\left(\sum_{j=0}^m w_{kj} o_{pj}\right), \quad k = 1, 2, \dots, n,$$

де (f) - сигмоїдна передавальна функція, тобто

$$f(x) = 1 / (1 + e^{-x}).$$

Автори [59] підкреслюють важливість застосування мережі зворотного поширення помилки у вивченні фінансових ринків. Вони порівняли результати, отримані даним методом з іншими та отримали, порівняно, більшу точність фінансових прогнозів.

Мартінез та ін. [2] впровадили MLP з метою вивчення взаємозв'язку між деякими технічними показниками і прогнозування щоденних максимальних і мінімальних цін на акції. Прогнозовані ціни використовуються в щоденній системі торгів, яка веде переговори на реальному фондовому ринку.

Автори [108] взяли за основу метод MLP з метою прогнозування індексу індійської фондової біржі на момент закриття. MLP, що використовувалася в цій роботі, була класичною тришаровою мережею, навченою алгоритмом зворотного поширення помилки. Автори досліджували кілька комбінацій параметрів мережі, а саме: кількість введених нейронів, прихованих нейронів і швидкості навчання, щоб знайти кращу конфігурацію MLP з точки зору точності.

В науковій літературі зустрічаються і численні модифікації методу ANN. Зокрема, в [6] здійснюється прогноз обмінних курсів на основі штучної мережі з функціональними зв'язками (англ. *functional link artificial neural network FLANN*) та штучної мережі з каскадними функціональними зв'язками (англ. *ascaded functional link artificial neural network, CFLANN*). Такі нейронні мережі є більш надійними та мають меншу обчислювальну складність у порівнянні з MLP, яка пройшла навчання методом зворотного поширення помилки.

Група авторів [60] теж досліджувала нейронні мережі на ринку FOREX, запропонувавши алгоритм поліноміальної нейронної мережі з регуляризацією Тихонова (англ. *dynamic ridge polynomial neural network, DRPNN*).

В роботі [62] досліджується 3 типи нейронних мереж для прогнозування деяких фондових індексів Болгарії: 1) навчання багат шарового процептрона Румельхарта методом зворотного поширення помилки, 2) нейронна мережа Гаусової радіальної базисної функції, 3) нейронна мережа загальної регресії (англ. *general regression neural network, GRNN*).

Автор [42] запропонував використовувати Баєсові мережі для прогнозування фондових індексів, оскільки Баєсова мережа є типом штучної нейронної мережі, в якій ваги зв'язків розглядаються як випадкові величини, а функції щільності записуються відповідно до правила Байєса. Коригування ваг під час процесу навчання складається з визначення функції щільності ймовірності для кожної ваги. У цій роботі для прогнозування динаміки цінних паперів використовується тришарова нейронна мережа.

Не дивлячись на те, що алгоритми штучних нейронних широко використовуються для прогнозування фінансових часових рядів, ці підходи все ж мають певні обмеження у моделях машинного навчання фінансових часових рядів, які характеризуються високою волатильністю, розмірністю, та нестационарністю. На противагу, цьому методу метод опорних векторів (англ. *support vector machines, SVM*) широко використовується як альтернатива в

задачах розпізнавання. Зокрема, автор [65] підкреслює, що з допомогою методу опорних векторів можна отримати глобальний оптимум, тоді як алгоритми штучних нейронних мереж продукують тільки локальний оптимум. В зазначеній публікації детально описано алгоритм використання методу опорних векторів у прогнозуванні зміни часових рядів. Для подільшого розгляду введемо такі позначення:

m – кількість базових часових одиниць (зазвичай днів), за які досліджується фінансовий часовий ряд;

n – кількість факторів, які впливають на напрямок зміни часового ряду;

A – матриця фінансових часових рядів розмірності $m \times n$;

D – одинична діагональна матриця напрямку зміни часових рядів;

ω - вектор рішень, який використовується для визначення гіперплощини з метою прогнозування зміни;

e – одиничний вектор, елементи якого дорівнюють 1;

x – вхідний вектор наявних даних, який використовується для прогнозування зміни фінансових часових рядів;

γ - дійсне число, яке використовується як параметр прогнозування.

Визначимо лінійно розділюючу гіперплощину:

$$P = \{x \mid x \in R^n, x^T \omega = \gamma\}. \quad (1.4)$$

Даний простір може бути використано для поділу R^n на дві частини (ліву і праву, верхню і нижню тощо). Якщо ми визначимо функцію рішення як:

$$f(x) = \begin{cases} +1, & \text{якщо } x^T \omega \geq \gamma, \\ -1, & \text{якщо } x^T \omega < \gamma, \end{cases}$$

то вона може бути використана для визначення розміщення вхідної точки x на основі розділюючої гіперплощини (1.4). Наприклад, якщо $f(x) = 1$ позначає, що точка розміщена в правій площині, то відповідно $f(x) = -1$ - в лівій.

Для заданого набору даних A та відомих розміщень, потрібно знайти відповідну гіперплощину з метою їх класифікації. Класичний метод опорних векторів якраз це і передбачає, що виражається у відповідній оптимізаційній задачі:

$$\begin{cases} \min_{(\omega, \gamma) \in R^{n+1}} \frac{1}{2} \omega^T \omega, \\ D(A\omega - e\gamma) \geq e. \end{cases} \quad (1.5)$$

Іншими словами, мета класичного методу опорних векторів – знайти ω та γ із оптимізаційної задачі (1.5).

Завдяки своїй нелінійній апроксимації, SVM застосовувалися як в кластеризації (англ. *support vector clustering, SVC*), так і в задачах відновлення регресії (англ. *support vector regression, SVR*).

Зокрема, в праці [68] вивчається проблема побудови (1.3) на основі методу SVR. В стандартному методі SVR приймається простір ℓ_p -norm ($p=1$ або 2) з метою контролю функціональної складності, а також вибирається \mathcal{E} -чутлива функція втрат для знаходження емпіричного ризику.

Нехай, дано тренувальний набір даних

$$\{(\mathbf{x}_1, y_1), \dots, (\mathbf{x}_N, y_N)\},$$

де $\mathbf{x}_i \in X$, $y_i \in \mathbb{R}$, N – кількість тренувальних точок даних, а X позначає простір вхідних даних \mathbb{R}^n .

Мета регресії полягає у такої пошуку функції, яка зможе не тільки наблизити добре ці дані, але також може передбачити точно значення y для майбутніх значень \mathbf{x} .

В даному випадку апроксимуюча функція набуває лінійної форми:

$$f(\mathbf{x}) = \mathbf{w}^T \mathbf{x} + b,$$

де $\mathbf{w} \in \mathbb{R}^n$ та $b \in \mathbb{R}$. Слід додати, що дана лінійна регресія може бути розширена до нелінійного вигляду застосовуючи ядро Мерсера [24].

Виходячи із вищесказаного, розглянемо задачу визначення \mathbf{w} та b з тренувального набору даних шляхом мінімізації ризику регресії $R_{reg}(f)$, визначеного як:

$$R_{reg}(f) = \Omega[f] + C \sum_{i=1}^N \Gamma(f(\mathbf{x}_i) - y_i),$$

де $\Omega[f]$ - очікуваний ризик, $\Gamma(\cdot)$ - функція втрат, C - предвизначене значення компромісу.

В узагальненому методі SVR, $\Omega[f]$ приймає форму $\|\mathbf{w}\|_1$ при ℓ_1 -SVR та $\frac{1}{2} \mathbf{w}^T \mathbf{w}$ при ℓ_2 -SVR.

Емпірична функція втрат приймає форму функції втрат з ε -чутливістю, визначена як:

$$\Gamma(f(\mathbf{x}) - y) = \begin{cases} 0, & \text{якщо } |y - f(\mathbf{x})| < \varepsilon, \\ |y - f(\mathbf{x})| - \varepsilon & \text{в іншому випадку.} \end{cases}$$

В даній функції, якщо точки даних знаходяться в межах $\pm\varepsilon$, то вони не належать до емпіричної похибки.

Відповідна оптимізація SVR (чи якщо бути більш точним ℓ_1 -SVR) може бути визначена як:

$$\begin{cases} \min_{\mathbf{w}, b, \xi_i, \xi_i^*} \|\mathbf{w}\|_1 + C \sum_{i=1}^N (\xi_i + \xi_i^*) \\ y_i - (\mathbf{w}^T \mathbf{x}_i + b) \leq \xi_i, \\ (\mathbf{w}^T \mathbf{x}_i + b) - y_i \leq \xi_i^*, \\ \xi_i \geq 0, \xi_i^* \geq 0, i = 1, \dots, N. \end{cases} \quad (1.6)$$

де ξ_i та ξ_i^* - відповідні додатні та від'ємні похибки в i -й точці.

Оптимізаційна задача (1.6) може бути розв'язано відомим методом лінійного програмування. Якщо структурний ризик $\Omega[f]$ приймає форму

$\frac{1}{2} \mathbf{w}^T \mathbf{w}$, (тобто $\ell_2 - SVR$), то оптимізаційна задача (1.6) зводиться до задачі квадратичного програмування.

В задачі (1.6) стандартний SVR метод фіксує границі ε глобально для всіх точок даних. Не дивлячись на те, що цей метод успішно використовується в багатьох сферах, він має значний недолік, який полягає у виявленні волатильності даних, що притаманне фінансовим ринкам.

Ми коротко розглянули окремі підходи машинного навчання при вирішенні проблем прогнозування фінансових часових рядів. Однак важливо відзначити, що ці індивідуальні підходи мають деякі недоліки, такі як наявність локальних оптимумів, перенавчання, труднощі у виборі багатьох параметрів, які безпосередньо впливають на точність прогнозування [69]. При виборі єдиного підходу до вирішення проблеми навчання інші алгоритми або архітектури одного і того ж алгоритму відкидаються з рішення. Для вирішення цієї проблеми кілька дослідників запропонували використовувати гібридні механізми, які поєднують індивідуальні рішення різних підходів. Комбінація індивідуальних підходів може дозволити знизити неточності регулювання параметрів і стохастичу в навчанні.

У літературі з прогнозування фінансових ринків було реалізовано декілька підходів для об'єднання індивідуальних алгоритмів для підвищення точності прогнозування. Ліанг [70] запропонував каскадний метод, що складається з параметричного та непараметричного підходів. Параметричний підхід спочатку виконується для моделювання тенденції руху цін. Далі використовується непараметричний підхід для вивчення залишків, які є результатом застосування параметричного підходу. Параметричними методами, дослідженими в цій роботі були метод біноміального дерева, метод скінченних різниць і метод Монте-Карло. Розглянутими непараметричними методами були лінійні нейронні мережі, MLP та SVR.

Подібне дослідження запропонували науковці в роботі [71], де запропоновано прогнозування цін на електроенергію на день вперед,

використовуючи гібридний механізм, що складається з трьох підходів. Першим методом є авторегресійна модель ковзної середньої з екзогенною змінною, яка застосовується при моделюванні лінійної залежності між історичними ціновими прибутками та іншими пояснювальними змінними. Потім використовується модель GARCH для ідентифікації гетероскедастичних характеристик залишків моделі ARMAX. Нарешті, для прогнозування залишків з GARCH була застосована адаптивна нейронна мережа (*англ. adaptive wavelet neural network, AWNN*).

Використання декількох алгоритмів машинного навчання призвело до появи окремого класу моделей, який отримав назву ансамблевих методів (*англ. Ensemble methods*). Такі методи використовують декілька методів машинного навчання з метою отримання кращої ефективності прогнозу, ніж можна було б отримати з допомогою кожного алгоритму окремо. У відмінність від статистичного ансамблю в статистичній механіці, зазвичай, ансамблеві методи складаються з конкретного скінченного ряду альтернативних моделей, але, як правило, дозволяють більш гнучко структурувати оцінку на певних підмножинах даних. Прикладом успішного використання ансамблевих методів у прогнозуванні фондових індексів є праця [72], в якій автори здійснюють порівняльний аналіз ефективності оцінки ансамблевих методів, а саме випадкового лісу і adaboost та моделей ANN, SVR, SVM і k-NN, для прогнозування зміни цін на акції. Отримані результати показали, що ансамблеві методи перевершують якість індивідуальних методів. Випадковий ліс забезпечив найкращу загальну продуктивність.

Інші дослідники вивчали використання алгоритмів оптимізації для покращення точності прогнозування деяких алгоритмів штучного інтелекту. Наприклад, в дослідженні [93] використовується генетичний алгоритм з метою пошуку кращої топології мережі MLP для покращення точності прогнозування. Генетичний алгоритм також можна використати для оптимізації параметрів методу опорних векторів та вибору входів моделі.

Відмітимо, що у фінансовій науковій літературі визначається, що математичного аналізу часових рядів недостатньо для здійснення прогнозів на індексів, оскільки існує велика кількість суб'єктивних факторів, таких як, політична ситуація, настрої населення, соціальні напруження, тощо. Наприклад, Ванг [7] наголосив, що важливим навіть є тональність у новинах, що може бути використана для прогнозування індексів. Тут важливо відмітити публікацію [73], в якій автори наводять новий підхід до прогнозування фінансових часових рядів базуючись на інтелектуальному аналізі тексту (що також поєднує статистичні методи та алгоритми штучного інтелекту). Нехай, існує деякий часовий ряд

$$Y = \{y_1, \dots, y_t, \dots, y_N\},$$

де y_t - значення в момент часу t ; N - кількість спостережень. Крім того, також існує набір ринкової інформації (думок, настроїв)

$$T = \{T_1, \dots, T_t, \dots, T_N\},$$

де T_t - текстові дані наявні в момент часу t .

В іншому дослідженні [92] автори запропонували новий підхід до прогнозування фондових індексів базуючись на історичних даних про щоденні ціни закриття акцій, здійснюючи поєднання аналізу часових рядів та методу k -найближчих сусідів (k -NN). Авторами було вибрано чотири індикатори:

- 1) Проста ковзна середня:

$$SMA(t) = \frac{1}{n} \sum_{i=t-n}^t x(i),$$

де $x(i)$ - часовий ряд цін акцій на момент закриття;

- 2) Індекс відносної сили RSI , який визначає силу тренду та ймовірність його зміни:

$$RSI = 100 - \frac{100}{1 + RS},$$

де $RS = \frac{\bar{U}}{\bar{D}}$, \bar{U} - згладжене ковзне середнє значення позитивних цїнових змін,
 \bar{D} - згладжене ковзне середнє значення негативних цїнових змін;

3) Стохастичний осцилятор (або стохастик), який будується з двох лїній $\%K$ (швидкий стохастик) та $\%D$ (повільний стохастик), які розраховуються як:

$$\%K_t = \frac{C_t - Ln}{Hn - Ln} \cdot 100,$$

де C_t - цїна закриття поточного перїоду; Ln - найнижча цїна за останнї n перїодів; Hn - найвища цїна за останнї n перїодів. $\%D$ розраховується як ковзна середня вїдносно $\%K$ з малим перїодом усереднення.

4) Лїнії Боллінджера, які були побудованї з метою порівняння волатильностї та рївня цїн за вїдповїдний перїод часу.

Як зазначалось вище, для прогнозування змїни 15 фондкових індексїв Бразилїї було використано методу k -найближчих сусїдїв (k -NN), який як зазначають автори, найпростїший з алгоритмїв машинного навчання. На основї часового ряду історичний даних цїн акцїй автори видїлили для кожного дня t вхїднї змїннї k -NN:

$$I_1 = RSI(t),$$

$$I_2 = (x(t) - bollinger_{upper}) / bollinger_{upper},$$

$$I_3 = (x(t) - bollinger_{lower}) / bollinger_{lower},$$

$$I_4 = \%K(t),$$

$$I_5 = \%D(t),$$

$$I_6 = \%K(t) - \%K(t-1),$$

$$I_7 = \%D(t) - \%D(t-1),$$

$$I_8 = (x(t) - x(t-1)) / x(t-1),$$

$$I_9 = (x(t) - x_t(t)) / (x_h(t) - x_t(t)),$$

$$I_{10} = (PMA_s(t) - PMA_s(t-1)) / PMA_s(t-1),$$

$$I_{11} = (PMA_l(t) - PMA_l(t-1)) / PMA_l(t-1),$$

$$I_{12} = (PMA_s(t) - PMA_l(t-1)) / PMA_l(t-1),$$

$$I_{13} = (x(t) - PMA_l(t)) / PMA_l(t),$$

$$I_{14} = x(t) - \min(x(t), \dots, x(t-5)) / \min(x(t), \dots, x(t-5)),$$

$$I_{15} = x(t) - \max(x(t), \dots, x(t-5)) / \max(x(t), \dots, x(t-5)),$$

$$I_{16} = (v(t) - v(t-1)) / v(t-1),$$

$$I_{17} = (VMA_s(t) - VMA_s(t-1)) / VMA_s(t-1),$$

$$I_{18} = (VMA_l(t) - VMA_l(t-1)) / VMA_l(t-1),$$

$$I_{19} = (VMA_s(t) - VMA_l(t-1)) / VMA_l(t-1),$$

$$I_{20} = (v(t) - VMA_l(t)) / VMA_l(t),$$

$$I_{21} = v(t) - \min(v(t), \dots, v(t-5)) / \min(v(t), \dots, v(t-5)),$$

$$I_{22} = v(t) - \max(v(t), \dots, v(t-5)) / \max(v(t), \dots, v(t-5)),$$

де $x(t)$, $x_h(t)$ та $x_l(t)$ - відповідно середні ціни закриття, найвищої та найнижчої цін в день t , $v(t)$ - кількість торгів в день t , $bollinger_{upper}$ та $bollinger_{lower}$ - відповідно верхня та нижня криві Боллінджера, PMA_s та PMA_l - відповідно коротко- та довгострокові прості ковзні середні цін, VMA_s та VMA_l - відповідно коротко та довгострокові прості ковзні середні кількостей.

Результати відповідних експериментальних досліджень показали важливість їх використання на практиці, а також можливість використання й інших методів машинного навчання для прогнозування фондових індексів, зокрема методу опорних векторів, машини екстремального навчання (*англ. Extreme Learning Machines, ELM*) та багат шарового прецептрона Румельхарта.

Таким чином, ми на сьогодні методи та моделі машинного навчання широко використовуються для вирішення завдань на фінансових ринках. Систематизуємо інформацію, отриману в даному розділі. Перш за все, об'єктом дослідження у досліджуваній літературі виступають численні Змінні фундаментального, технічного та текстового аналізів (рис. 1.3)

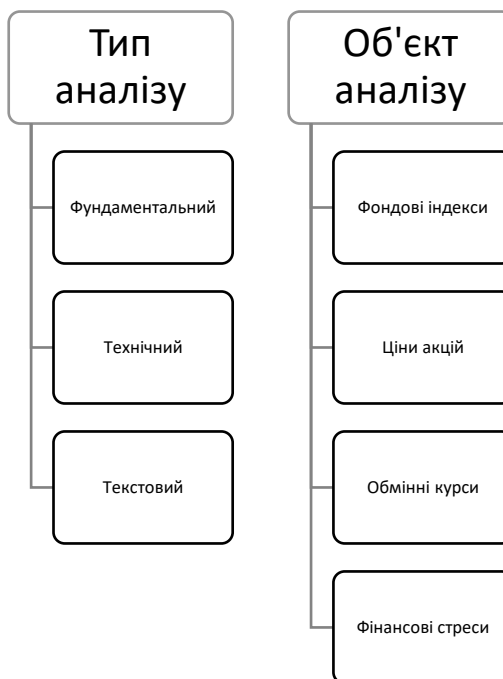


Рис. 1.3. Типи змінних, які використовуються у дослідженні фінансових ринків методами штучного інтелекту

Велика кількість літературних джерел присвячена попередній обробці даних з допомогою алгоритмів штучного інтелекту:



Рис. 1.4. Методи та моделі штучного інтелекту, які використовуються для первинної обробки фінансової інформації

Головною метою більшості проаналізованих методів та моделей, запропонованих у літературі, є прогнозування майбутньої поведінки ринку з максимальною точністю. Прогнозні значення можуть бути використані для підтримки прийняття рішень про купівлю або продаж фінансових активів з метою максимізації прибутку інвесторів. Проте важливо відзначити, що, на сьогодні не існує єдиної методології, яка б реалізувала побудову успішної інтелектуальної торговельної системи.

Таким чином, в даному розділі нами проведено детальний систематизований аналіз методів та моделей прогнозування фінансових часових рядів із застосуванням алгоритмів машинного навчання. Нами був проведений детальний аналіз відповідної наукової літератури в наукометричних базах даних Scopus, Science Direct, Google Scholar, IEEEExplore та Springer. Виявлено, що з метою забезпечення підтримки прийняття рішень інвесторами в різних сегментах фінансових ринків широко використовувались методи інтелектуального аналізу великих даних та м'яких обчислень (нечіткі множини, нечітка логіка, нечіткі регулятори, нечіткі нейронні мережі, генетичні алгоритми і еволюційне моделювання тощо).

Висновки до розділу 1

1. Виявлено, що прогнозування фінансових часових вважається однією з основних задач у науковій літературі, метою яких є вивчення часових рядів, а завдяки своєму практичному значенню, аналіз фінансових ринків в останні десятиріччя почав широко вивчатися науковцями з галузі математики, комп'ютерних наук та інженерії, адже успіх інвестора залежить від якості інформації, яку він використовує для підтримки прийняття.

2. Математичні моделі оцінки фінансових часових рядів на практиці мають багато видів, а також моделюють різноманітні стохастичні процеси.

Виявлено, що дослідження взаємозалежності між декількома часовими рядами фінансових даних здійснюється з допомогою багатовимірних VAR моделей часових рядів. Виділено, що нелінійна природа фінансових рядів часових даних є наслідком наявності багатьох класів нелінійних моделей, в яких випадкові величини мають різну дисперсію (гетероскедастичність), тому в такому випадку опис та моделювання відповідних часових рядів здійснюються з допомогою авторегресивних моделей умовної гетероскедастичності ARCH, а також її узагальнень (GARCH, TARARCH, EGARCH, FIGARCH, CGARCH тощо).

3. Теоретичний аналіз використання статичних методів аналізу фінансових часових рядів продемонстрував, що фінансові часові ряди володіють високою дисперсією, є складними, динамічними, нелінійними, непараметричними, а також мають хаотичну природу. Тому, зроблено висновок, що на противагу описаній статистичній методиці при моделюванні та прогнозуванні фінансових часових рядів, широко використовуються методи машинного навчання. В межах даного розділу здійснено детальний аналіз наукової літератури в наукометричних базах даних Scopus, Science Direct, Google Scholar, IEEEExplore та Springer, який дозволив виділити такі класи моделей: штучні нейронні мережі ANN; метод опорних векторів SVM; гібридні системи, тобто поєднання декількох підходів; оптимізаційні методи у поєднанні із машинним навчанням; ансамблеві методи (англ. ensemble methods); інтелектуальний аналіз тексту; специфічні методи. При цьому зроблено головний висновок, що найбільшого поширення набули інтегровані моделі авторегресії-ковзної середньої (ARIMA) та штучних нейронних мереж (ANN).

4. Розроблено формалізований алгоритм прогнозування фінансових часових рядів на основі методів штучного інтелекту, який передбачає такі етапи: підготовку даних, визначення алгоритму, навчання, оцінку прогнозу, розробку торговельної стратегії та оцінку активів

РОЗДІЛ 2. ГІБРИДНІ МЕТОДИ ТА МОДЕЛІ: ПОРІВНЯЛЬНИЙ АНАЛІЗ

2.1. Алгоритм прогнозування часових рядів фондових індексів на основі гібридної моделі ARMA моделі та штучної нейронної мережі

В розділі 1 даного дисертаційного дослідження підкреслено, що на сьогодні розроблено велику кількість моделей прогнозування часових рядів фінансових даних, які ґрунтуються на використанні статистичних моделей та методів машинного навчання. Зокрема, серед статистичних методів найбільшого поширення набув широкий клас інтегрованих моделей авторегресії-ковзної середньої (ARIMA) [3], а серед методів машинного навчання – штучні нейронні мережі (ANN).

Моделі ARIMA, які також відомі як моделі Бокса-Дженкінса [4], виявилися особливо ефективними при здійсненні короткотермінових прогнозів та, часто, перевершують більш складні структурні моделі [3]. Крім того, зазначені моделі розглядаються як найбільш ефективні методики прогнозування в соціально-економічних науках та широко використовується в аналізі та прогнозуванні саме фінансових часових рядів. Використання моделі ARIMA для прогнозування часових рядів має істотне значення в умовах невизначеності, оскільки не передбачає знання будь-якої базової моделі або відносин, як у деяких інших методах: вона спирається на минулі значення рядів даних, а також помилок прогнозування [2, 4].

Штучні нейронні мережі (ANN), як частина методів машинного навчання, є найбільш точними та широко використовуються для прогнозування фондових індексів, курсів валют, мікро- макроекономічних показників тощо [1, 8, 9, 10, 92]. Таке розповсюдження цієї методики зумовлено наявністю відмінних рис для штучних нейронних мереж, які роблять їх привабливими як для науковців-дослідників, так і для практиків. Як зазначається в праці [1], методи ANN керуються великими даними, а

також є самоадаптивними з кількома вхідними припущеннями. Крім того, методи ANN виступають універсальним апроксиматором, оскільки мережа може ефективно апроксимувати неперервну функцію до бажаного рівня точності. Важливо, що штучні нейронні мережі ефективні у розв'язанні нелінійних задач, у тому числі емпіричних [1]: це головна перевага методології на відміну від багатьох традиційних методів прогнозування часових рядів, таких як ARIMA, які припускають, що ряди генеруються з лінійних процесів, і як наслідок, можуть бути неприйнятними для більшості практичних задач, які є нелінійними [8, 92]. На сьогодні актуальним є розв'язання багатьох нелінійних задач, зокрема прогнозування часових рядів фондових ринків, які є нелінійними з невизначеною поведінкою та змінюються з часом [9, 10].

Враховуючи вищесказане, розглянемо детальніше ефективність моделей ANN та ARIMA (а також їх поєднання) у прогнозуванні фондових індексів з метою отримання більш точних прогнозів.

Відомо, що інтегрована модель авторегресії-ковзної середньої є лінійною та поєднує окремі види однойменних моделей. Так, модель авторегресії (AR) передбачає, що значення x_t часового ряду може бути задане як функція його p останніх значень $x_{t-1}, x_{t-2}, \dots, x_{t-p}$. Тоді модель авторегресії $AR(p)$ порядку p має вигляд:

$$x_t = \phi_1 x_{t-1} + \phi_2 x_{t-2} + \dots + \phi_p x_{t-p} + \omega_t, \quad (2.1)$$

де x_t - стаціонарний випадковий процес, $\phi_1, \phi_2, \dots, \phi_p$ - константи ($\phi_p \neq 0$), а ω_t - гауссівський білий шум із середнім значенням, що дорівнює 0 та дисперсією σ_ω^2 .

З іншого боку, модель ковзної середньої (MA) передбачає, що значення x_t часового ряду може бути задане, як функція його q останніх значень білого шуму $\omega_{t-1}, \omega_{t-2}, \dots, \omega_{t-q}$. Відповідно, модель ковзної середньої порядку q , $MA(q)$ матиме вигляд:

$$x_t = \omega_t + \theta_1 \omega_{t-1} + \theta_2 \omega_{t-2} + \dots + \theta_q \omega_{t-q}, \quad (2.2)$$

де q – кількість лагів ковзних середніх, а $\theta_1, \theta_2, \dots, \theta_q$ ($\theta_q \neq 0$) - параметри.

Моделі (2.1) та (2.2), відповідно, можуть бути об'єднані в, так звану, модель авторегресії-ковзної середньої (ARMA):

$$\phi(B)x_t = \theta(B)\omega_t, \quad (2.3)$$

$$\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p, \quad (2.4)$$

$$\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q. \quad (2.5)$$

Модель $ARMA(p, q)$ визначена в (2.3), де $\phi(B)$ та $\theta(B)$ - оператори AR та AM (B – оператор зворотного зсуву), відповідно визначені в (2.4) та (2.5).

Досліджувана модель ARIMA є розширенням моделі ARMA для нестационарних часових рядів. Зокрема, несезонний процес $ARIMA(p, d, q)$ задається як:

$$\phi(B)(1-B)^d x_t = c + \theta(B)\omega_t.$$

Відповідно, сезонний процес $ARIMA(p, d, q)(P, D, Q)_m$ задається як:

$$\Phi(B^m)\phi(B)(1-B^m)^D(1-B)^d x_t = c + \Theta(B^m)\theta(B)\omega_t,$$

де $\Phi(z)$ та $\Theta(z)$ - поліноми порядку P та Q , відповідно, m – компонента сезонності (наприклад, кількість місяців, кварталів у році). Якщо $c \neq 0$, то маємо поліном порядку $d + D$.

Варто наголосити, що на сьогодні існує недолік використання ARIMA-моделей у прогнозуванні фінансових часових рядів, який полягає у тому, що процес вибору відповідного порядку є суб'єктивним та складним на практиці. В науковій літературі існують підходи до вирішення зазначеної проблеми. Так, в дослідженні [34] розроблено метод визначення порядку ARIMA для стаціонарних рядів. Метод передбачає побудову авторегресійної моделі виходячи із наявних даних з подальшою оцінкою функції правдоподібності потенційних моделей через набір стандартних регресій. Автори дослідження визначили асимптотичні властивості методу на занадто

загальний умовах. Гомес [11] розширив попереднє дослідження оцінкою мультиплікативної сезонної моделі ARIMA. В своїй наступній праці [7] автор запропонував підхід автоматичної ідентифікації моделі на основі програмного забезпечення TRAMO and SEATS: для заданих рядів даних оцінка моделей здійснювалась шляхом знаходження мінімального значення баєсового інформаційного критерію (*BIC*) Ліу [12] запропонував алгоритм ідентифікації сезонної ARIMA моделі із використанням методу фільтрації та евристичних правил; алгоритм передбачає використання програмного забезпечення SCA-Expert. Зустрічається, також, використання програмного забезпечення Forecast Pro, Autobox, Time Series Expert (TSE-AX) для оцінки параметрів моделі ARIMA.

Тому, при використанні методу ARIMA повинен використовуватись алгоритм автоматичного вибору найбільш адекватної моделі, який має значно спростити (автоматизувати) численну кількість типових розрахунків. Так званий, автоматичний алгоритми прогнозування повинен забезпечувати вибір відповідної адекватної моделі, оцінку її параметрів та безпосереднє здійснення прогнозу часового ряду. Даний алгоритм повинен враховувати велику кількість модифікацій інтегрованих моделей авторегресії-ковзної середньої і застосовуватись до великих масивів рядів даних без втручання користувача. Головна задача в автоматичному прогнозуванні ARIMA полягає у виборі водовідвідного порядку моделі, тобто значень p, q, P, Q, D, d . Якщо значення d та D відомі, то ми можемо вибрати порядки p, q, P та Q через інформаційний критерій Акаїке (*AIC*):

$$AIC = -2\log(L) + 2(p + q + P + Q + k),$$

де $k = 1$, якщо $c \neq 0$ та 0 у протилежному випадку, L – функція максимальної правдоподібності моделі для першої різниці ряду $(1 - B^m)^D (1 - B)^d x_t$.

Штучні нейронні мережі ANN відносяться до нелінійних методів моделювання, які використовуються у широкому класі задач. Відомо, що ця архітектура цих мереж має два або більше шарів, зокрема, наприклад,

тришарова модель має шар входу, прихований шар виходу та шар виходу. У прогнозуванні часових рядів розглядається багатошаровий перцептрон, навчений методом зворотного поширення помилки. x_t розглядається як нелінійна функція від x_{t-1}, \dots, x_{t-N} :

$$x_t = g(x_{t-1}, x_{t-2}, \dots, x_{t-N}) + v_t, \quad (2.6)$$

де g – нелінійна функція, v_t - шум або похибка.

Передавальна функція прихованого шару, як правило, є сигмоїдною функцією:

$$S(x) = \frac{1}{1 + e^{-x}}.$$

З метою оцінки параметрів моделі наявний ряд даних задається як вхідний шар нейронної мережі. Навчання нейронної мережі може розглядатися як зведення до мінімуму помилки, утвореної значеннями ваг зв'язків. Мінімізація помилки за допомогою багатошарового перцептрону реалізується через алгоритм зворотного поширення для тренування нейронних мереж. Ефективність прогнозування при використанні штучних нейронних мереж вища, ніж у ARIMA, оскільки перші моделюють нелінійні часові ряди більш точно, що впливає із (2.6).

Часто на практиці вхідні дані можуть мати як лінійну, так і нелінійну складову, тому у прогнозуванні фінансових часових рядів набуло поширення поєднання використання обох методів, які тримали назву гібридні моделі ARIMA-ANN, серед яких найбільшого використання набули 2 методики: Жанга [13] та Хашей-Біжарі [14]. Метод Хашей-Біжарі, також, розглядає часові ряди, як суму двох компонент. На першому етапі здійснюється оцінка ARIMA, після чого послідовність помилок розглядається як нелінійна компонента та моделюється з допомогою штучної нейронної мережі. Прогнози, отримані з цих двох методів, поєднуються з метою отримання остаточного прогнозу. Автори показали кращу ефективність прогнозу у порівнянні із моделлю Жанга. Жанг показав кращу ефективність гібридної моделі у поєднанні із індивідуальним використанням двох методів.

Дослідимо ефективність зазначеного методу. Так, загальному випадку в гібридних моделях будь-який часовий ряд задається як сума лінійної та нелінійної компоненти:

$$x = L + N,$$

де L – лінійна складова, а N – нелінійна.

В нашому випадку, лінійна компонента L оцінюватиметься з допомогою ARIMA, а нелінійна N - з допомогою багат шарового перцептрону (MLP) використовуючи залишки з оцінки ARIMA. Головна ідея такого гібридного поєднання полягає в тому, що якщо часові ряди будуть оцінюватись лінійною моделлю (в нашому випадку ARIMA), тоді її залишки міститимуть тільки нелінійну складову. Таким чином, комбіновані моделі дають змогу використовувати унікальні властивості та переваги ARIMA та багат шарового перцептрону у розв'язанні задач різного класу. У дослідженні [58] автором дисертаційного дослідження було розроблено модель, яка поєднує використання ARIMA моделі для часового ряду та нейронної мережі для опису динаміки залишків. Інший підхід, заснований на використанні карт Кохонена, використаний в роботі [152].

Надалі, позначимо через e_t - залишки моделі ARIMA в час t , тоді

$$e_t = x_t - \hat{L}_t, \quad (2.7)$$

де \hat{L}_t - відповідно оцінка на основі ARIMA моделі, яка містить відповідну нелінійну складову в залишках; відповідно, на другому етапі, моделюючи залишки з допомогою багат шарового перцептрону, можна оцінити нелінійні характеристики. З кількістю нодів n , MLP-модель залишків матиме вигляд:

$$e_t = f(e_{t-1}, e_{t-2}, \dots, e_{t-n}) + \varepsilon_t \Rightarrow \hat{N}'_t = \hat{e}_t = f(e_{t-1}, e_{t-2}, \dots, e_{t-n}), \quad (2.8)$$

де f – нелінійна функція, що визначена багат шаровим перцептроном, \hat{N}'_t - прогнозне значення в момент часу t , ε - випадкова похибка.

Відповідну оцінку точності прогнозів здійснюватимемо з допомогою таких індикаторів, як середньоквадратичної похибки MSE та $RMSE$,

середньої абсолютної похибки MAE та середньої абсолютної відсоткової похибки $MAPE$:

$$\begin{aligned}
 MAE &= \frac{1}{N} \sum_{i=1}^N |e_i|, \\
 MSE &= \frac{1}{N} \sum_{i=1}^N (e_i)^2, \\
 MAPE &= \frac{1}{N} \sum_{i=1}^N \frac{|e_i|}{|x_i|}, \\
 RMSE &= \sqrt{\frac{1}{N} \sum_{i=1}^N (e_i)^2}.
 \end{aligned}
 \tag{2.9}$$

Для практичної реалізації алгоритму здійснимо поділ ряду даних x_t на навчальний ряд ($x_{t,навч}$) та тестовий ($x_{t,тест}$, останній рік) (рис. 2.1).

Крім того, для всіх розрахунків тут і нижче здійснюється попередня обробка даних шляхом їх нормалізації:

$$x_t = \frac{x_{t-1} - x_{\min}}{x_{\max} - x_{t-1}}.$$

Окремі емпіричні розрахунки (ARIMA) за описаним алгоритмом здійснювались на мові Python, Matlab, розрахунки, що пов'язані з машинним навчанням – у відкритих бібліотеках Python.



Рис. 2.1. Часові ряд SP500 та його поділ на навчальний та тестовий

Попередній візуальний аналіз часового ряду (рис. 2.1) демонструє відсутність очевидного глобального чи сезонного тренду. Суттєвий спад, який спостерігається, пов'язаний із пандемією COVID-19, яка суттєво

вплинула на світові фінансові ринки. Додатково здійснимо візуальний аналіз автокореляційної (ACF) та часткової автокореляційної функцій (PACF) (рис. 2.2).

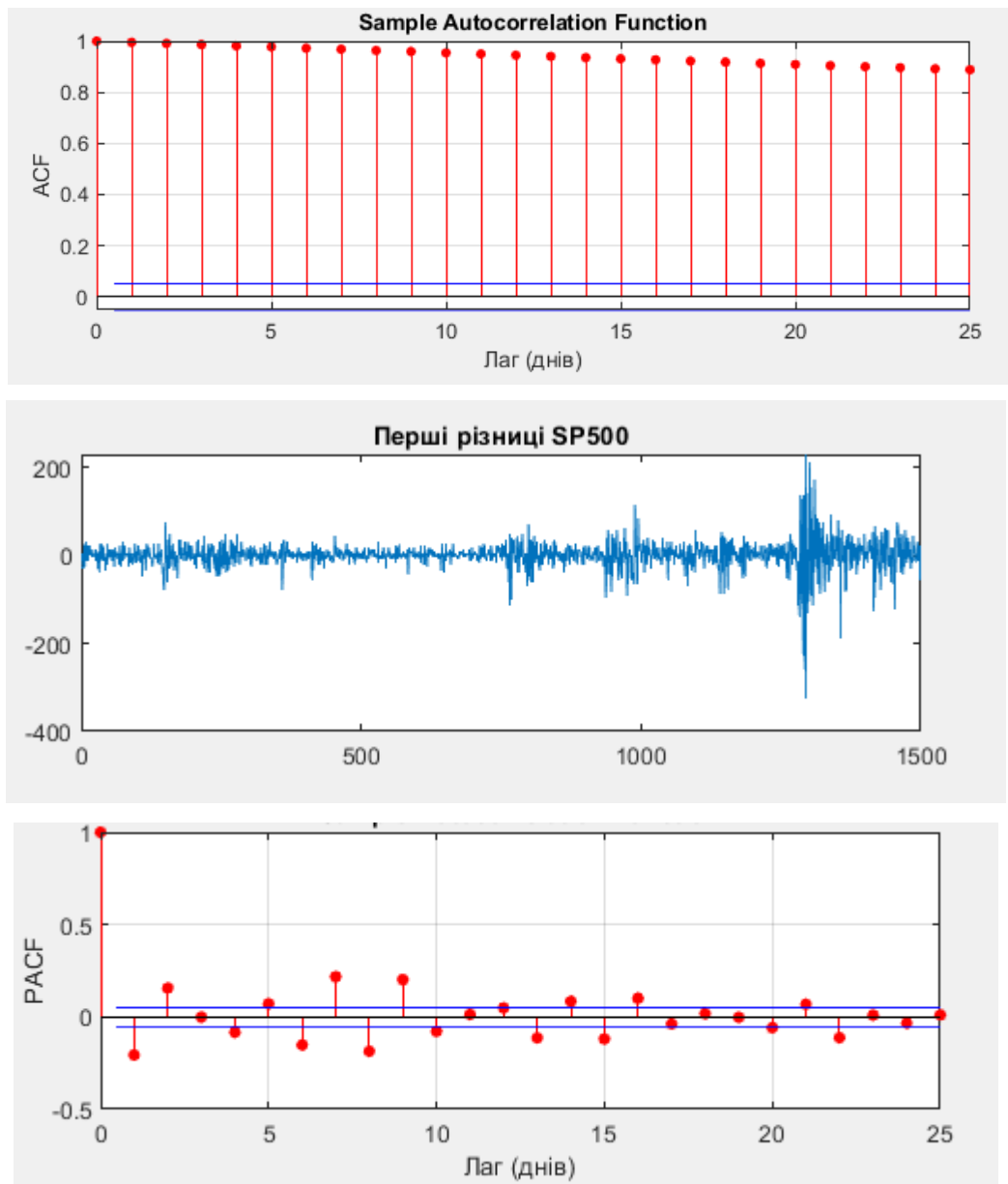


Рис. 2.2. Автокореляційна функція (ACF) та частинна автокореляційна функція (PACF)

Автокореляційна функція (рис. 2.2) містить значення, які суттєво відрізняються від нуля; часткова автокореляційна функція вказує на те, що автокореляція стрімко спадає на першому лазі, тобто ряд нестационарний. Відповідне взяття перших різниць та аналіз їх корелограми (рис. 2.2) вказує, що ряд є стаціонарним.

Відповідне взяття перших різниць і отримання стаціонарного ряду говорить про те, що подальша оцінка ARIMA повинна здійснюватися із параметрами: $AR(0)$ та $MA(0)$. Не дивлячись на те, що попередній аналіз виявив, що найкращою є модель ARIMA порядку $(0,1,0)$, використаємо алгоритм автоматичного вибору відповідної моделі про що наголошувалось вище. Один з таких відомих алгоритмів розв'язку цієї задачі запропоновано в Дурбіном та Купманом [15], який реалізований у функції *arima()* програмного пакету *R* і який часто використовується як вітчизняними, так і іноземними науковцями. Даний алгоритм передбачає, що часові ряди мають середнє значення, що дорівнює нулю та високу дисперсію. Однак, вибір d та D шляхом мінімізації інформаційного критерію Акаїке з використанням даного алгоритму призводить до надлишкової кількості взяття перших різниць ряду даних. Оскільки, об'єктом даного дослідження є прогнозування фінансових часових рядів, то на нашу думку кількість взяття різниць ряду даних має бути мінімальною, оскільки їх надлишковість призводить до зниження точності прогнозу та збільшує довірчі інтервали. Тому використаємо алгоритм Хіндмана-Кандакара [16], який базується на тесті існування одиничних коренів. В зазначеному алгоритмі для несезонних даних розглядається модель $ARIMA(p,d,q)$, в якій значення d вибирається на основі успішного тесту Квятковського-Філіпса-Шмідта-Шина (KPSS): якщо результат тесту є значимим, то проводиться тест перших різниць ряду на наявність одиничного кореня; процедура повторюється доти, поки не отримається перший незначимий результат. Якщо знайдено модель з найменшим значенням інформаційного критерію Акаїке, то вона стає новою поточною моделлю та процедура повторюється. Алгоритм завершується коли неможливо знайти модель, яка наближена до поточної моделі із найменшим рівнем інформаційного критерію Акаїке.

Перейдемо до емпіричного аналізу запропонованого методу. Результати побудови ARIMA-моделі наведено в Табл. 2.1:

Табл. 2.1. Результати алгоритму автоматичного вибору порядку

Порядок моделі	ARIMA-моделі	
	AIC	BIC
	Теоретична ARIMA	
ARIMA(0,1,0)	10346,461	-
	Автоматична ARIMA	
ARIMA(3,1,1)	9472,866	6503,657

Для подальших розрахунків була вибрана не теоретична ARIMA-модель, а модель за параметрами ARIMA(3,1,1) оскільки вона має найменше значення інформаційного критерію Акаїке.

Для побудови гібридної моделі згідно (2.7)-(2.8) вихідний ряд даних x_t представлено сумою лінійної компоненти e_t (яка буда оцінена за відповідною моделлю ARIMA(3,1,1) та нелінійної складової \hat{L}_t , яка оцінена з допомогою штучної нейронної мережі. Нелінійна модель представлена у вигляді $N^{x,y,z}$, де x – кількість вхідних шарів, y – розмір прихованого шару, z – кількість вихідних шарів.

Відповідні прогностні значення фондових індексів, отриманих за методологією ARIMA та гібридної моделі наведені на рис. 2.3:

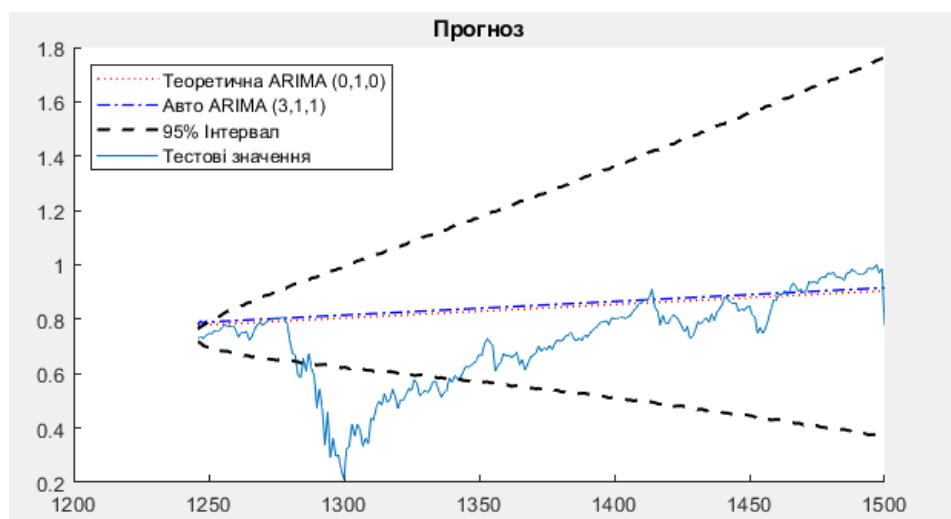




Рис. 2.3. Тестові та прогнозовані значення оцінені з допомогою ARIMA (зверху) та гібридної моделі (знизу)

На основі отриманих даних згідно алгоритму, зображеного на рис. 2.1, було здійснено повторний прогноз фондових індексів. Для порівняння ефективності здійснених прогнозів за моделлю ARIMA та побудованої гібридної моделі зобразимо графіки середньої квадратичної похибки, розрахованої за (2.9). Оцінка прогнозу наведено в наступній таблиці:

Табл. 2.2. Оцінка прогнозу моделлю ARIMA та гібридною моделлю

	ARIMA (0,1,0)	ARIMA (3,1,1)	ARIMA-MLP
MAE	0,136	0,143	0,1324
MSE	666,053	717,329	1993,534
MAPE	26,417	27,599	11,485
RMSE	25,808	26,783	44,649

Відповідні значення табл. 2.2. вказують на більшу точність прогнозів фондового індексів S&P 500, отриманих за допомогою алгоритму гібридної

інтегрованої моделі авторегресії-ковзної середньої та штучної нейронної мережі, для якого значення відповідних оцінок нижчі.

Таким чином, часові ряди фондових індексів містять як лінійну, так і нелінійну складову, а отже, окремо лінійна модель ARIMA та нелінійна модель ANN не можуть дати точну оцінку таких часових рядів. Тому при прогнозуванні часових рядів фондових індексів потрібно використовувати, так звані, гібридні моделі, які поєднують переваги обидвох методик.

2.2. Алгоритм прогнозування часових рядів фондових індексів на основі гібридної моделі авторегресії-ковзної середньої та рекурентної нейронної мережі

Підкреслимо, що в попередньому дослідженні використання методів машинного навчання показало широке поширення має класу ANN. Однак, однією із найбільших проблем навчання (перенавчання) ANN є надмірне тренування та чутливість до кількості прихованих нейронів. Тому в дослідженні [17] вказується на перевагу рекурентних нейронних мереж у порівнянні з попереднім класом. Зокрема автор зазначає, що обчислення ANN є повільнішими, оскільки вони передбачають нелінійні оцінки для кожного нейрона в архітектурі. Це робить рекурентні нейронні мережі особливо привабливим для використання їх у контексті обмежень в режимі реального часу, а це особливо важливо для фінансових часових рядів. Виходячи з цього, по аналогії можна було б запропонувати гібридну модель ARIMA-RNN. Однак, на нашу думку метод, описаний в попередньому підрозділі метод має суттєвий недолік, оскільки ряд даних не розкладається безпосередньо на лінійну та нелінійну складову, а здійснюється оцінка лінійної моделі ARIMA, а набір помилок розглядається як нелінійна складова. Таким чином, обидві гібридні методи передбачають, що модель ARIMA лінійна. Однак, з іншого боку, ARIMA-моделі мають багато цікавих

статистичних властивостей. По-перше, як зазначалось вище, ω_t в (2.1) за своєю природою є гауссівським білим шумом; гауссівський випадковий процес $\{X_t\}_{t \in T}$ представлений випадковим вектором $(X_{t_1}, \dots, X_{t_m})^T$ за своєю природою має багатовимірний нормальний розподіл. По-друге, гауссівський процес завжди є стаціонарним (стаціонарним у вузькому сенсі). Таким чином, якщо припустити, що часовий ряд ARMA є строго стаціонарним, то відповідно, він є і гауссівським за деяких додаткових умов на залишки ω_t . Зазвичай, після приведення часового ряду до стаціонарного, оцінка коефіцієнтів моделі ARIMA здійснюється з допомогою узагальненого методу максимальної правдоподібності, де заданий часовий ряд є гауссівським. Таким чином, можна зробити висновок, якщо часовий ряд є строго стаціонарним, то ARIMA модель є більш адекватною для гауссівських рядів.

В загальному випадку, з метою перевірки гіпотези про нормальність розподілу використовується тест Жарка-Бера, який, зокрема, передбачає перевірку значення ексцесу:

$$k = \frac{E\{(x - E\{x\})^4\}}{\left(E\{(x - E\{x\})^2\}\right)^2}, \quad (2.10)$$

де x – випадкова величина, для якої розраховується коефіцієнт ексцесу, E – математичне сподівання.

Як що в (2.10) значення ексцесу дорівнює 3, то ряд є гауссівським; такий ряд даних вважатимемо рядом з низькою волатильністю. Якщо ж значення ексцесу для заданого ряду даних буде відмінне 3, то, відповідно, вважатимемо його рядом з високою волатильністю. Зауважимо, що часовий ряд з високою волатильністю або лептокуртичний, або платікуртичний характер, а це означає, що відповідний розподіл не гауссівським. Таким чином, ми прийшли до висновку, що ARIMA-моделі є більш адекватними для часових рядів даних, для яких значення ексцесу наближене до 3.

Розглянемо метод оцінки прогнозу часового ряду, запропонованого в роботах [150, 157, 158].

Так, ряд даних x_t представимо сумою компонент з низькою волатильністю l_t та високою волатильністю h_t :

$$x_t = l_t + h_t. \quad (2.11)$$

Після перевірки часового ряду l_t на стаціонарність, він моделюється з допомогою ARIMA, як лінійна функція минулих значень $l_{t-1}, l_{t-2}, \dots, l_{t-p}$ та випадкових помилок $n_t, n_{t-1}, \dots, n_{t-q}$:

$$\hat{l}_t = f(l_{t-1}, l_{t-2}, \dots, l_{t-p}, n_t, n_{t-1}, \dots, n_{t-q}). \quad (2.12)$$

Аналогічно, h_t виражається як нелінійна функція від $h_{t-1}, h_{t-2}, \dots, h_{t-N}$ та моделюється з допомогою штучної нейронної мережі:

$$\hat{h}_t = g(h_{t-1}, h_{t-2}, \dots, h_{t-N}) + \varepsilon_t. \quad (2.13)$$

В (2.11) - (2.13) f – лінійна функція, g – нелінійна функція, ε_t - помилка.

Прогнозні значення \hat{y}_t часового ряду отримуються як сума:

$$\hat{x}_t = \hat{l}_t + \hat{h}_t. \quad (2.14)$$

Розглянемо алгоритм детальніше (рис. 2.1):

2) з допомогою простої ковзної середньої здійснюється поділ часового ряду на 2 компоненти: з низькою волатильністю та високою волатильністю.

$$x_{tr} = \frac{1}{m} \sum_{i=t-m+1}^t x_i,$$

Довжина згладжування m змінюється до тих пір, поки не буде досягнуто відповідного значення ексцесу. Перша компонента (згладжений тренд x_{tr}), яка розглядається як компонентна з низькою волатильністю. Друга компонента, з високою волатильністю, - це залишки, які обчислюються як $x_r = x_t - x_{tr}$;

3) компонента з низькою волатильністю x_{it} і значенням ексцесу $k = 3$ оцінюється з допомогою моделі ARIMA, а прогнозні значення отримуються з (2.8);

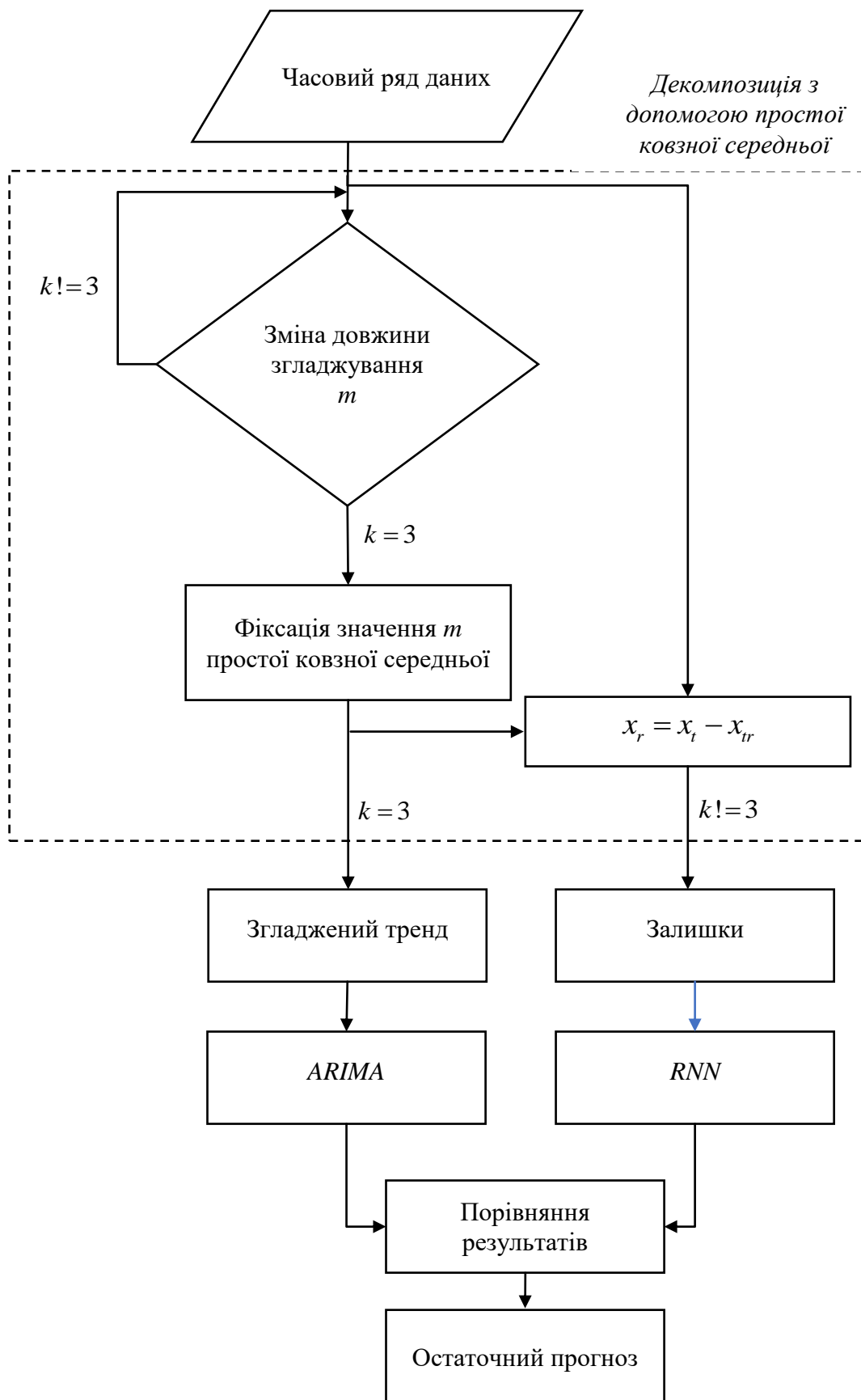


Рис. 2.4. Алгоритм побудови гібридної моделі ARIMA-RNN

4) компонента з низькою волатильністю x_r і значенням ексцесу $k!=3$ оцінюється з допомогою штучної нейронної мережі, а прогнозні значення отримуються з (2.13);

5) здійснюється остаточний прогноз на основі (2.14).

Отже, покроковий алгоритм виглядає так - рис. 2.4.

Нелінійна функція g в (2.13) в нашому випадку є моделлю рекурентної нейронної мережі, в якій з'єднання між її вузлами формують граф, що орієнтований у часі. Це, в свою чергу формує внутрішній стан мережі, що дає можливість оцінювати динамічну поведінку в часі:

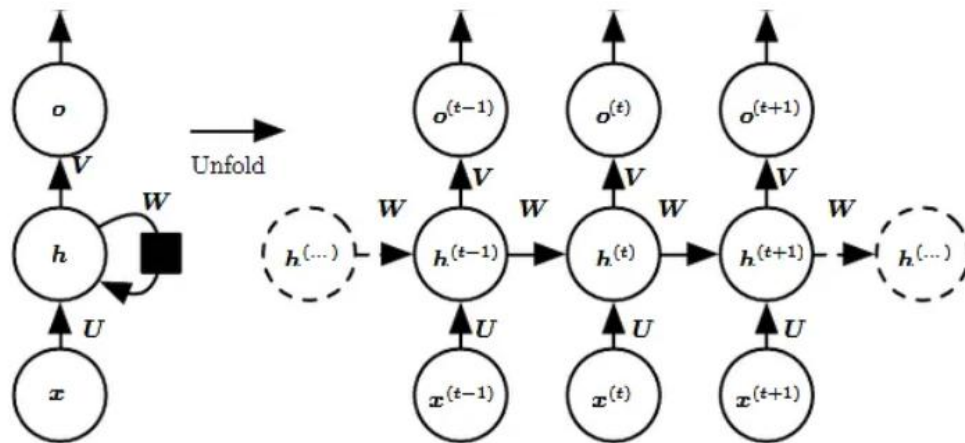


Рис. 2.5. RNN-структура

Вхідний шар представлений x , а N – кількість нейронів; відповідно прихований шар – h та кількість нейронів – G , вихідний шар – o з кількістю нейронів M . Позначимо через w_{gv} вагу v -го нейрона у вхідному шарі до g -го нейрона в прихованому шарі; $w_{g'g}$ – вага g -го нейрона в прихованому шарі до g' -го нейрона в прихованому шарі, а w_{oh} – вага g -го нейрона до o -го нейрона у вихідному шарі; a – сумарні значення, тобто, a_h^t відображає зважене вхідне значення в прихованому шарі в момент часу t ; b – оцінено функцією активації, тобто b_h^t відображає вихідне значення, оцінене функцією активації у прихованому шарі в момент часу t ; L – функція втрат; x_i^t – вхідне значення i -го нейрона у вхідному шарі в момент часу t .

Відповідно оцінка (тренування) прямим поширенням відбувається наступним чином:

$$a_h^t = \sum_{i=1}^N w_{gv} x_i^t + \sum_{g'=1}^G w_{g'g} b_{g'}^{t-1},$$

$$b_t^g = \theta_g(a_g^t),$$

$$a_o^t = \sum_{g=1}^G w_{og} b_g^t,$$

$$b_o^t = \theta_o(a_o^t).$$

Процес зворотного оновлення параметрів здійснюється на основі методу зворотного поширення в часі:

$$\frac{\partial L}{\partial w_{gk}} = \sum_{t=1}^T \frac{\partial L}{\partial a_o^t} \frac{\partial a_o^t}{\partial w_{gk}},$$

$$\delta_o^t \square \frac{\partial L}{\partial a_o^t}.$$

Уточнення δ_o^t ваги w_{og} розраховується як:

$$\delta_o^t = \frac{\partial L}{\partial w_{gk}} = \sum_{t=1}^T \frac{\partial L}{\partial a_o^t} \frac{\partial a_o^t}{\partial w_{og}} = \sum_t \delta_j^t b_i^t,$$

А уточнення δ_g^t ваги в прихованому шарі як:

$$\delta_g^t = \theta'(a_g^t) \left(\sum_{o=1}^O \delta_o^t w_{go} + \sum_{g'=1}^G \delta_{g'}^{t+1} w_{g'g} \right).$$

Запропонований алгоритм може дати більшу точність порівняно з іншими гібридними моделями. Частина з низькою волатильністю має Гауссівську природу, а тому вона може бути більш точно оцінена з використанням лінійної моделі ARIMA. Аналогічно, часовий ряд з високою волатильністю можна розглядати як нелінійний. Таким чином, коли ряд даних розкладається на компоненти з низькою та високою волатильністю (лінійну та нелінійну), то сумарна помилка моделі буде низькою.

Згідно (2.11)-(2.14) вихідний ряд даних x_t представимо сумою компонент з низькою волатильністю l_t та високою волатильністю h_t .

Компонента вихідного часового ряду з низькою волатильністю l_t (лінійна частина) була оцінена за відповідною моделлю ARIMA(3,1,1) (2.12). Компонента з високою волатильністю h_t (нелінійна частина) оцінена з допомогою рекурентної нейронної мережі (2.13).

Аналогічно, для експериментальних досліджень розробленої моделі використовувалась відкрита бібліотека Python Keras. Ціна закриття індексу SP500 на 5й день прогнозується на основі значень за попередні 4 дні. Кількість нейронів у вхідному, прихованому і вихідному шарах рекурентної нейронної мережі – 4, 4 та 1 відповідно.

Відповідні прогностичні значення фондових індексів, отриманих за методологією запропонованої гібридної моделі наведені на рис. 2.6:



Рис. 2.6. Тестові та прогностичні значення оцінені з допомогою гібридної моделі

Оцінка прогнозу наведена у наступній таблиці:

Табл. 2.3. Оцінка прогнозу за гібридною моделлю ARIMA та рекурентної нейронної мережі

	ARIMA-RNN
MAE	39,525
MSE	3600,330

MAPE	1,321
RMSE	60,003

У даному дослідженні розроблена удосконалена гібридна модель прогнозування на основі поєднання ARIMA та RNN, яка базується на статистичних властивостях ARIMA. Розроблений алгоритм передбачає використання фільтра ковзної середньої для розкладання часових рядів на дві компоненти: з низькою та високою волатильністю. Наступний крок передбачає оцінку двох компонент, за моделями ARIMA та RNN, відповідно. Прогноз, отриманий з гібридної моделі – це сума прогнозів, що тримані з двох індивідуальних моделей.

Висновки до розділу 2

Основні результати даного розділу були зосереджені на дослідженні гібридних моделей, в яких основний процес оцінювався на основі моделі часового ряду, а залишки моделі оцінювалися на основі нейронних мереж. Використання такого підходу дало змогу розширення використання класичних моделей часових рядів на випадок відсутності гомоскедастичності у залишках та відсутності незалежності залишків часового ряду. А саме:

1. Доведено, що часові ряди фондових індексів містять як лінійну, так і не лінійну складову, а отже, окремо лінійна модель ARIMA та нелінійна модель ANN не можуть дати точну оцінку таких часових рядів. Тому при прогнозуванні часових рядів фондових індексів потрібно використовувати, так звані, гібридні моделі, які поєднують переваги обидвох методик.

2. Розроблено гібридну модель моделі авторегресії-ковзної середньої та штучної нейронної мережі прогнозування часових рядів фондових індексів. Емпіричні розрахунки здійснювались на мові Python, Matlab, розрахунки, що пов'язані з машинним навчанням – у відкритій бібліотеці Python Keras. Побудована модель показала більшу точність

прогнозів, отриманих за допомогою алгоритму гібридної інтегрованої моделі авторегресії-ковзної середньої та штучної нейронної мережі, для якого значення відповідних оцінок похибок нижчі.

3. Оскільки найпоширенішою проблемою навчання штучної нейронної мережі є надмірне тренування та чутливість до кількості прихованих нейронів, розроблено гібридну модель авторегресії-ковзної середньої та рекурентної нейронної мережі. Доведено, якщо часовий ряд є строго стаціонарним, то ARIMA модель є більш адекватною для гауссівських рядів. В загальному випадку, з метою перевірки гіпотези про нормальність розподілу використовується тест Жарка-Бера, який, зокрема, передбачає перевірку значення ексцесу. Якщо значення ексцесу дорівнює 3, то ряд є гауссівським; такий ряд даних вважатимемо рядом з низькою волатильністю. Якщо ж значення ексцесу для заданого ряду даних буде відмінне 3, то, відповідно, вважатимемо його рядом з високою волатильністю. Таким чином, ми прийшли до висновку, що ARIMA-моделі є більш адекватними для часових рядів даних, для яких значення ексцесу наближене до 3. Відповідно у розробленій гібридній моделі ряд даних x_t представлено сумою компонент з низькою волатильністю l_t та високою волатильністю h_t : компонента з низькою волатильністю x_{tr} і значенням ексцесу $k=3$ оцінюється з допомогою моделі ARIMA, а компонента з високою волатильністю x_r і значенням ексцесу $k \neq 3$ оцінюється з допомогою рекурентної нейронної мережі. Запропонований алгоритм продемонстрував більшу точність порівняно з іншими гібридними моделями, оскільки ряд даних розкладається на компоненти з низькою та високою волатильністю (лінійну та нелінійну), а отже сумарна похибка моделі є низькою.

РОЗДІЛ 3. ВИКОРИСТАННЯ МЕТОДІВ МАШИННОГО НАВЧАННЯ ДЛЯ ДОСЛІДЖЕННЯ ХАОТИЧНИХ ФІНАНСОВИХ ЧАСОВИХ РЯДІВ

3.1. Огляд теореми про вкладення та штучний інтелект

Існують чотири фундаментальні характеристики хаотичних систем: аперіодичність характеризує системи, в яких кожен стан системи не повторюється; обмеженість означає, що сусідні стани залишаються в скінченному діапазоні і не наближаються до нескінченності; детермінованість означає, що існує керуюче правило без випадкових умов для прогнозування майбутнього стану системи; а також чутливість до початкових умов, що означає, що невелика різниця в початкових умовах призводить до розбіжності двох близьких один до одного станів у міру розвитку системи. Квазіперіодичні та гіперхаотичні системи відносяться до хаотичних систем [22, 64, 79, 114, 118, 136]. Нелінійні системи зворотного зв'язку, в яких поточний стан системи відводиться назад для створення майбутнього стану, мають потенціал стати хаотичними [67, 87, 116, 139].

Мета цього розділу – представити необхідні означення та інструменти, які використовуються в роботі для розробки оригінального методу прогнозування хаотичних фінансових часових рядів.

Характеристики хаотичних систем

Вкрай важливо визначити характеристики хаотичних систем, щоб точно вибрати правильні інструменти для аналізу та дослідження динамічних систем. Wilding [132] представив комплексний огляд характеристик хаотичних систем.

Розглянемо короткий опис хаотичних детермінованих систем та їх основних характеристик. Система є детермінованою, коли закон керування визначає майбутній стан системи на основі попередніх станів без залучення випадкових умов [64, 139]. Комплексні системи демонструють обмежувальну

поведінку. Коли точки даних, які виділяються з динамічної системи, залишаються в скінченному діапазоні і не наближаються до нескінченності, система обмежена. Іншою вимогою до обмеженої системи є стаціонарність. Бути стаціонарною означає завжди демонструвати однакову поведінку, або коли середнє та стандартне відхилення залишаються незмінними.

Аперіодична поведінка є ключовим елементом хаотичних систем. Аперіодична характеристика визначається як властивість системи показувати нерегулярні коливання, які не збільшуються, не зменшуються і не залишаються нерухомими. Іншими словами, система не є ні випадковою, ні періодичною [22, 64, 132].

Система, яка працює у стабільному режимі, може перейти і інші режими, такі як періодичний або хаотичний, якщо змінити параметри системи. Крім того, складні системи часто перемикаються між різними режимами, коли система розвивається з плином часу [54].

Чутливість до початкових умов є найвідомішою характеристикою хаотичних систем [54]. Навіть незначна відмінність початкових умов з часом посилюється, доки майбутні стани двох точок, спочатку близьких одна до одної, не стануть абсолютно відмінними. Ця дуже важлива властивість хаотичних систем унеможливує довгострокове передбачення складних систем. Показник Ляпунова є відповідним інструментом для кількісної оцінки збіжності чи розбіжності системи та визначення перспектив передбачуваності [21].

Хаотичні системи породжують закономірності, незважаючи на поведінку, схожу на випадкову. Сформовані шаблони не повторюються, але вони мають подібні властивості. Сніжинки є прикладом таких типів візерунків. Детермінована система породжує форму сніжинок; однак невелика зміна впливає на остаточну форму сніжинок. У літературі про хаос атрактор визначається як шаблони, що генеруються хаотичними системами [21, 54].

Термін "атрактор" визначається Crutchfield та ін. [53] як "набір точок у фазовому просторі, які відвідує розв'язок рівняння еволюції задовго після того, як (початкові) перехідні процеси зникли". Аналогічний підхід до визначення атрактора наявний і в інших роботах, наприклад [33]. Таким чином, атрактори геометрично формують довгострокову проекцію системи в режимі фазового простору. Для стабільної системи атрактор – це атрактор з фіксованою точкою. Атрактор для періодичних систем є циклічним. Квазіперіодична система має атрактор, який являє собою комбінацію двох або більше періодів. Хаотичні системи мають дивний атрактор [18, 33], який описується операцією розтягування та згортання. Процес розтягування та згортання створює візерунки, відомі як фрактали.

Принципи редукціонізму не підходять для аналізу хаотичних систем. Згідно принципів редукціонізму вважається, що складну проблему можна розбити на прості складові [18, 33]. Пізніше, аналізуючи прості частини складної проблеми, загальну проблему можна розв'язати. Goldratt [90] продемонстрував, що в хаотичних системах ця методологія не спрацьовує, і незначна зміна у будь-якій з простих складових може призвести до різкої зміни загальної реакції системи.

Хаотичні системи, проаналізовані за допомогою комп'ютера, можуть давати різні результати. Ідентичний код може давати абсолютно різні результати при опрацюванні двома різними комп'ютерами або з різним програмним забезпеченням. Багато дослідників [18, 33, 90] попереджають про ризик покладатися лише на результат комп'ютерних пакетів. Вони підкреслюють, що необхідно розуміти результати комп'ютерних пакетів та причини, що лежать в основі отриманих результатів.

Розроблений автором метод використання методів хаотичних систем в поднанні з теорією часових рядів, представлений в роботі [77].

3.2. Основні показники хаотичності системи

Теорія хаосу – це нова область математики, яка також відома як складність або нелінійна динаміка [18, 33, 90]. Хоча хаотичні системи детерміновані та регулюються певними законами, через нелінійні ефекти та чутливість до початкових умов вони передбачувані в межах обмеженого горизонту.

Багато природних систем демонструють нелінійну або хаотичну поведінку. Використовуючи теорію хаосу, ці системи були описані математичними рівняннями. Для хаотичної системи фазовий простір визначається як векторний простір R^n , при цьому кожна точка у фазовому просторі описується n -вимірним вектором $s(t)$, необхідним для опису динаміки системи, тобто $s(t)$ визначається рівнянням:

$$s(t) = [s_1(t), s_2(t), s_3(t), \dots, s_n(t)]$$

де t – індекс часового ряду, а n – розмірність векторного простору R^n . За допомогою нелінійної функції $Q: R^n \rightarrow R^n$, яка описує систему, майбутню величину системи в момент часу $t+\tau$ можна визначити за рівнянням

$$s(t) \rightarrow Q\{s(t)\} = s(t + \tau)$$

Невелика зміна стану системи $s(t)$ істотно вплине на тенденцію системи, і після кількох ітерацій система стане непередбачуваною. Така поведінка динамічних систем відома як чутливість до початкових умов або ефект метелика [88, 115]. Прогресування не випадкової системи створює траєкторію, яка називається атрактором. Теорема вкладення Такенса [46, 100, 119] стверджує, що оскільки величина $s(t)$ та його компонентів, $s_1(t), s_2(t), s_3(t), \dots$ у хаотичній системі невідомі, то якщо можна спостерігати одиничну величину або змінну $x(t)$ з цієї динамічної системи, тоді атрактор можна розгорнути з цього набору спостережуваних зразків. Це означає, що якщо окрему величину $x(t)$ спостерігати з хаотичної системи, відновлена динаміка системи $Y(t) = [x(t), x(t+T), x(t+2T) \dots]$ з T , визначеним як затримка часу, геометрично подібна до вихідного атрактора.

Отже, якщо існує динамічна система $s(t) \rightarrow s(t + 1)$, то послідовний порядок реконструйованих точок фазового простору $Y(t) \rightarrow Y(t + 1)$ слідує за невідомою динамікою $s(t) \rightarrow s(t + 1)$. Тому поведінка власне системи відображається у спостережуваних часових рядах, згенерованих системою [46, 100].

3.2.1 Хаотичні часові ряди

Часовий ряд – це послідовність скалярних значень у часі. Дослідники описують часовий ряд як “послідовність регулярно відібраних величин із спостережуваної системи”. Канц і Шнайбер [68] визначають часовий ряд як набір скалярних значень, які вимірюють стан певної системи в часі. Часовий ряд – це історичний запис системи, при цьому вимірювання здійснюються через рівні проміжки часу з узгодженістю в методі вимірювання та системі. Це корисне джерело інформації для аналізу та дослідження характеристик і поведінки системи. Аперіодичність, обмеженість, детермінованість і чутливість до початкових умов відрізняють хаотичні часові ряди від інших типів часових рядів [46, 100]. Аналіз часових рядів включає в себе методи, які дозволяють маніпулювати, характеризувати та виконувати кількісний та якісний аналіз для розуміння основних характеристик системи [59].

3.2.2 Визначення хаосу в системах

В аналізі часових рядів важливим кроком є визначення характеристик даних. Для того, щоб відрізнити хаотичні дані від нехаотичних даних були використані наступні методи. Метод показника Ляпунова, за яким шукають основні характеристики хаотичних систем, чутливість до початкових умов, найбільше переважає в літературі. Перетворення Фур'є може бути використано для ідентифікації хаосу в даному часовому ряду. Перетворення Фур'є обчислює наявні частоти у часовому ряду. Перетворення Фур'є

часового ряду, $x(t), t = 0, 1, 2, \dots, N - 1$, де N – кількість точок, яке можна обчислити наступним чином

$$X(k) = \frac{1}{N} \sum_{t=0}^{N-1} x(t) e^{-2\pi i \left(\frac{kt}{N}\right)} \quad k = 0, 1, 2, \dots, (N - 1)$$

де $i = \sqrt{-1}$ – уявнй одиниця. Спектр потужностей є квадратом $X(k)$ і може бути обчислений наступним чином

$$P(k) = \frac{1}{N^2} \left| \sum_{t=0}^{N-1} x(t) e^{-2\pi i \left(\frac{kt}{N}\right)} \right|^2 \quad k = 0, 1, 2, \dots, (N - 1)$$

Побудова діапазону потужностей часового ряду може допомогти визначити природу часового ряду. Спектр для хаотичних часових рядів буде широкосмуговим, що означає, що різниця між найвищим значенням потужності та найнижчим значенням на діаграмі спектра потужності є значною, з широким піком, що означає, що на графіку спектра потужності доступний широкий діапазон частот. Випадкові дані повинні мати постійно визначений спектр потужності. Для періодичних даних спектр потужності зростає на частотах, що характеризують систему, і залишається близьким до нуля для інших систем.

Показник Ляпунова

Важливою характеристикою хаотичних систем, що визначається як ефект метелика, є висока чутливість системи до початкових умов [19, 21]. Показник Ляпунова – найбільш практичний метод виявлення хаотичної поведінки в системі. Показник Ляпунова кількісно визначає збіжність та розбіжність сусідніх траєкторій. Якщо існує функція, яка відображає $x(t)$ на $x(t + 1)$, $x(t) \rightarrow R(x(t)) = x(t + 1)$, для двох найближчих початкових точок у $x(0)$ і $x(0) + \Delta x(0)$, після однієї ітерації поділ точок можна обчислити наступним чином:

$$\Delta x(1) = R[x(0) + \Delta x(0) - R(x(0))] \approx \Delta x(0) R'\{x(0)\}.$$

Показник Ляпунова при $x(0)$ можна визначити таким чином, що $e^\lambda = \left| \frac{\Delta x(1)}{\Delta x(0)} \right|$, або його можна переставити, як показано в співвідношенні (3.1):

$$\lambda = \ln \left| \frac{\Delta x(1)}{\Delta x(0)} \right| \approx \ln |R'[x(0)]| \quad (3.1)$$

Величина $\frac{\Delta x(1)}{\Delta x(0)}$ є показником Ляпунова або виміром розтягування при $x = x(0)$. Якщо $\frac{\Delta x(1)}{\Delta x(0)}$ – від'ємне, це означає, що дві сусідні точки обмінюються порядком (більший стає меншим, і навпаки) після ітерації. Для отримання глобального показника Ляпунова середнє значення рівності (3.1) на великій кількості ітерацій обчислюється, як обчислюється наступним чином

$$\lambda = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{t=0}^{N-1} \ln |R'[x(t)]|.$$

Нульове значення інтерпретується як циклічна поведінка, негативний показник означає нехаотичну поведінку, а позитивний показник Ляпунова доводить існування хаосу в системі [19, 21]. Числовий приклад міститься у Додатку В.

Показник Герста

Гарольд Едвін Герст відомий тим, що ввів показник Герста як міру передбачуваності часового ряду [32, 59, 103, 124]. Показник Герста визначається за допомогою аналізу $\frac{R}{S}$. Для заданого часового ряду з N точок та вибраним p , де $10 \leq p < \frac{N}{2}$, часовий ряд можна розділити на блоки $\frac{N}{p}$. Для кожного блоку обчислюється максимальний діапазон, середнє значення та стандартне відхилення. Потім для кожного блоку отримується значення (діапазон)/(стандартне відхилення), rs , і усереднюється по всіх блоках. Наступна рівність показує зв'язок між середнім значенням rs і показником Герста:

$$rs = \left(\frac{p}{2} \right)^H,$$

де H – показник Герста. Показники Герста можуть змінюватися між 0 і 1. Показник Герста 0,5 вказує на справжнє випадкове блукання. Значення від 0 до 0,5 вказує на нестійку поведінку, тобто дані не випадкові, але нинішня

тенденція навряд чи збережеться. Показник Герста між 0,5 і 1 доводить, що дані є більш стійкими і поточний напрямок, імовірно, продовжиться. Показник Герста 0 означає, що часовий ряд змінює напрямок з кожною вибіркою. Постійний часовий ряд з ненульовим градієнтом призведе до значення Герста 1.

Фрактальна розмірність

Іншим методом виявлення існування хаосу в системі є фрактальна розмірність. Неціла фрактальна розмірність вказує, що система хаотична. Кореляційна або фрактальна розмірність є однією з найпоширеніших фрактальних розмірностей, використання якої відображене в літературі [37]. Якщо сфера з радіусом r відцентрована на певній точці D -вимірного простору, то середнє значення точок у сфері $C(r)$, за винятком центральної точки, можна обчислити наступним чином

$$C(r) = \frac{1}{N(N-1)} \sum_{j=0}^{(N-DT-1)} \sum_{k=0, k \neq j}^{(N-DT-1)} \Theta(r - |Y(j) - Y(k)|)$$

де $\Theta(x)$ – індикатор функції, визначений як:

$$\Theta(x) = \begin{cases} 0 & \text{if } x < 0 \\ 1 & \text{if } x \geq 0 \end{cases}.$$

Графічно $C(r)$ проти r повинна давати приблизно пряму лінію, нахил якої є кореляційною розмірністю d_c , як сформульовано у наступному співвідношенні

$$d_c = \lim_{R \rightarrow 0} \frac{d \log C(r)}{d \log r}.$$

Кореляційна розмірність d_c , представлена цілою величиною, вказує, що аттрактор – це простий геометричний об'єкт. Неціле значення для d_c вказує на те, що аттрактор є дивним, а система – хаотичною [18, 33, 37, 119].

3.2.3. Теорема вкладення

За відсутності керівного рівняння для хаотичної системи точки фазового простору генеруються з вихідних часових рядів методом вкладення [46, 100, 119]. Теорема вкладення передбачає, що якщо для хаотичного часового ряду

вибрана відповідна розмірність, дані можна реконструювати, а розгорнуті дані можуть виявити приховану інформацію системи.

Хаотичний часовий ряд може бути вкладений у багатовимірний простір шляхом побудови графіку точки даних $Y(t), t = 1, 2, \dots, N - (D - 1)T$, де $Y(t) = [x(t), x(t + T), \dots, x(t + (D - 1)T)]$, N – довжина вихідного часового ряду, D – розмірність вкладення часового ряду, а T – затримка часу. Метод вкладення спостережуваних даних у D – вимірний простір відомий як реконструкція фазового простору. Для належного застосування теореми Такенса необхідний відповідний вибір розмірності вкладення D і затримки часу T .

3.2.4. Визначення затримки часу

Щоб вибрати належну затримку часу T , слід забезпечити низьку кореляцію між найближчими елементами в точках фазового простору, яка не буде занадто довгою. Перший мінімум середньої функції взаємної інформації може бути обраний як затримка часу.

Середня взаємна інформація – це теоретичний метод з'єднання двох наборів вимірювань між собою та встановлення критерію їх взаємної залежності на основі поняття інформаційного зв'язку між ними. Середня взаємна інформація використовується для точного визначення уявлення про те, що вимірювання $x(t)$ у момент часу t теоретично пов'язані з вимірами $x(t + T)$ у момент часу $t + T$. Отже, середня взаємна інформація між $x(t)$ та $x(t + T)$ може бути визначена наступним чином:

$$I(T) = \sum_{x(t), x(t+T)} \Pr [x(t), x(t + T)] \log_2 \left[\frac{\Pr [x(t), x(t + T)]}{\Pr [x(t)] \Pr [x(t + T)]} \right]$$

де $\Pr(x(t), x(t + T))$ – сумісна щільність $x(t)$ і $x(t + T)$, $\Pr(x(t))$ та $\Pr(x(t + T))$ – це щільності $x(t)$ і $x(t + T)$ відповідно.

$I(T)$ визначає середню кількість інформації, що поділяється двома значеннями у часовому ряду. Коли значення T збільшується, $x(t)$ і $x(t + T)$ стають незалежними, а $I(T)$ прямує до нуля.

3.2.5. Визначення розмірності вкладення

Щоб вибрати відповідну розмірність вкладення D , важливо зрозуміти вплив розмірності вкладення на структуру атрактора. Зі збільшенням розмірності вкладення атрактор розгортається. Коли атрактор розгортається повністю, та сама точка на атракторі не перетинається.

Метод хибних найближчих сусідів передбачає, що при перетині траєкторій атрактора дві сусідні точки фазового простору стоять надзвичайно далеко одна від одної в послідовному порядку вкладених часових рядів. Автори робіт [38, 56] детально описали процедуру оцінки розмірності вкладення (розмірності атрактора). В роботах [38, 56] ефективну оцінку розмірності вкладення пропонуться визначати наступним чином:

$$E(D) = \frac{1}{N-DT} \sum_{t=0}^{N-DT-1} \frac{\|Y_{D+1}(t) - Y_{D+1}^{NN}(t)\|}{\|Y_D(t) - Y_D^{NN}(t)\|}$$

де N – довжина вихідного часового ряду, індекс D відноситься до розмірності вкладення, а індекс NN означає найближчого сусіда до іншого вектора, що обчислюється наступним чином

$$\|Y_D(t) - Y_D^{NN}(t)\| = \max_{0 \leq j \leq D-1} |x(t + jT) - x^{NN}(t + T)|$$

Оскільки D стає досить великим, $E1(D) = \frac{E(D+1)}{E(D)}$ – відношення $E(D + 1)$ до $E(D)$, наближається до 1.

Відповідна розмірність вкладення задається значенням D , де $E1(D)$ перестає змінюватися. Практично це означає вибір значення D , де $E1(D)$ починає вирівнюватись, близько до одиниць.

Узагальнюючи все вище сказане, ми можемо визначити загальний алгоритм оцінки прогнозу часового ряду з корекцією на наявність хаотичної складової, як показано на рис. 3.1.

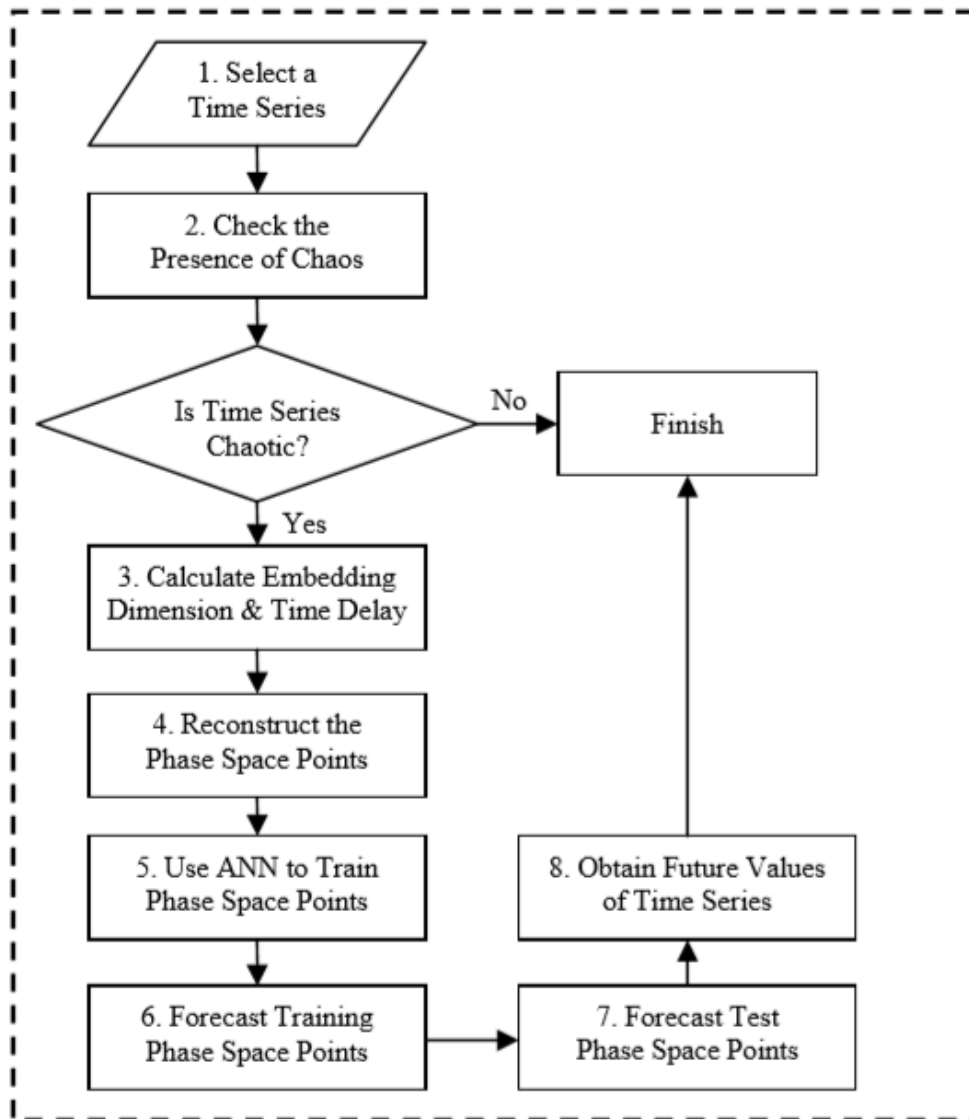


Рис. 3.1. Схема дослідження часових рядів з корекцією детермінованого хаосом

3.3. Генетичні алгоритми

Генетичні алгоритми – це еволюційний метод, який імітує процес природного відбору, запропонований Дарвіном у 1859 р. Генетичні алгоритми дотримуються стратегії виживання, яка використовується природою для модифікації видів. У 1975 році Голландія вперше запровадила метод ГА для цілей оптимізації.

У методі ГА параметри оптимізації представлені у вигляді рядків двійкових цифр для імітації біологічних хромосом. Кросовер та мутація

використовуються для створення нового потомства. Формат хромосом є першим кроком у розробці структури пошуку ГА. Параметри реструктуризуються у двійкові рядки для формування рядків цифр [40, 42, 80]. У літературі, присвяченій ГА, простір, призначений для кожного параметра, називається геном. Кожен ген містить кілька двійкових цифр. Хромосома утворюється при поєднанні кількох генів, як зображено на рис. 3.2.



Рис. 3.2. Рядок, що представляє ген та хромосому в генетичних алгоритмах.

Визначаються мінімальні та максимальні значення параметрів. Значення генів потрібно утримувати між цими двома значеннями. Наступним кроком є вибір чисельності популяції, яка є кількістю хромосом для кожного покоління. Перед запуском процесу ГА потрібно визначити локації кросоверу та ймовірність мутації.

У методі ГА перша популяція хромосом генерується випадково. Хромосоми розбиваються на гени. Значення генів, які представляють значення відповідних параметрів, приміряються до загальної функції. Вимірюється придатність кожної хромосоми. За допомогою рівняння (3.2) визначається ймовірність відбору хромосом для спарювання:

$$p_t = \frac{f_t}{\sum_{i=1}^N f_i} \quad (3.2)$$

де f_t позначає придатність хромосоми, а N – чисельність популяції.

На основі рівняння (3.2) вибираються спарювані хромосоми і відповідно до заданої ймовірності визначаються місця кросоверу. Коли відбудеться кросовер, мутація буде застосовуватися до кожного біта хромосом на основі заданої ймовірності. Отримані хромосоми передаються наступному

поколінню, і популяція наступного покоління формується, як показано на рис. 3.3. З усіх поколінь найкраща хромосома, яка оптимізує загальну функцію, відбирається як найкраща хромосома, і відповідно до цього обираються значення параметрів.

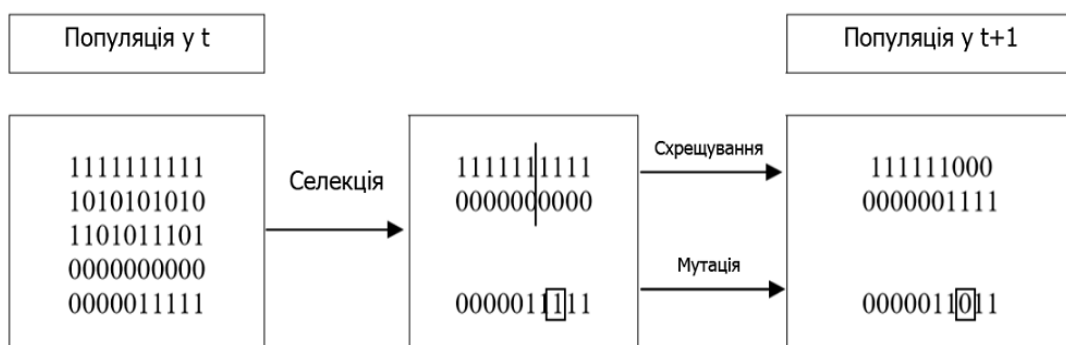


Рис. 3.3. Схематична ілюстрація генетичних алгоритмів

3.4. Використання машинного навчання для прогнозування фінансових хаотичних часових рядів

Розглянемо більш детально аналіз часових рядів на основі оцінки хаотичної складової в часовому ряді. Зауважимо, що оцінка хаотичності системи буде здійснюватися на основі показників, описаних вище. В результаті оцінки хаотичності на основі показника Герста вдається спростити задачу побудови прогнозу часового ряду з врахуванням теореми вкладення та відповідних оцінок на динаміку вихідного процесу.

Реалізація розрахунків здійснювалась на мові Python з допомогою відкритої бібліотеки нейронної мережі Keras, яка в нашому прикладі налаштована, як надбудова над фреймворком TensorFlow. Дана бібліотека призначена для оперативної роботи з мережами глибинного навчання та, при цьому спроектована так, щоб бути компактною, модульною та такою, що легко розширюється.

Для здійснення оцінок прогнозу була використана архітектура рекурентних нейронної мереж довга короткочасна па'мять (LSTM).

Модель LSTM в Keras передбачає, що дані поділяються на вхідні (X) та вихідні (Y) компоненти. У задачах прогнозування часових рядів спостереження на в момент часу ($t-1$) використовуються як вхідні дані, а спостереження на в момент часу (t) як вихідні дані.

Відповідно, на а наступному кроці було об'єднано вхідні та вихідні компоненти з метою створення дата фрейму, який можна використати при навчанні з учителем. Для створення та використання вхідних даних був згенерований масив NumPy вихідних та лагових даних.

На наступному кроці здійснювалась перевірка на стаціонарність та здійснювалась їх диференціація вбудованими засобами бібліотеки Pandas. Відомо, що стаціонарні ряди легше моделювати, а їх оцінки мають більш точні прогнози.

Як і інші нейронні мережі, LSTM передбачає, що дані знаходяться в межах масштабу функції активації, яка використовується мережею. В LSTM функцією активації за замовчуванням є гіперболічний тангенс (\tanh), виходом якої є значення від -1 до 1, що є найбільш «зручним» діапазоном для даних часових рядів. Щоб зробити експеримент адекватним, мінімальні та максимальні значення коефіцієнтів масштабування розраховані на навчальних даних та застосовані для масштабування тестів даних та прогнозів. Це зроблено для того, щоб вивільнити з експерименту знання, які походять з тестового набору даних. Перетворення набору даних у діапазон $[-1, 1]$ здійснено з використанням класу MinMaxScaler. Як і інші класи перетворення в scikit-learn, для цього потрібні дані, надані у матричному форматі.

LSTM є різновидом рекурентних нейронних мереж. Перевага цього типу мережі полягає в тому, що вона може вчитись і запам'ятовувати протягом довгих послідовностей і не покладається на попередньо задане вікно з лагом, як вхід. У Keras це реалізується через параметр `stateful` при визначенні рівня LSTM.

За замовчуванням шар LSTM у Keras підтримує стан між даними в межах одного батчу. Батч - це фіксована кількість рядків з навчального ляду даних, яка визначає, скільки шаблонів потрібно обробити перед оновленням ваг мережі. Стан в шарі LSTM очищується між батчами за замовчуванням, тому LSTM потрібно задати з параметром `stateful`.

Шар LSTM задається з такими вхідними даними:

- зразки (`samples`) - незалежні спостереження, як правило вихідні ряди даних;
- часові кроки (`time steps`) - окремі часові кроки даної змінної для даного спостереження.
- характеристики (`features`): окремі оцінки на момент спостереження.

Враховуючи, що ряд даних нашого експерименту заданий як набір навчальних (X) та вихідних даних (Y), його змінено у формат зразки-часові кроки-характеристики наступним чином:

```
X, y = train[:, 0:-1], train[:, -1]
```

```
X = X.reshape(X.shape[0], 1, X.shape[1])
```

Форму вхідних даних потрібно вказати на рівні LSTM, використовуючи аргумент `batch_input_shape`, як кортеж, який визначає очікувану кількість спостережень для зчитування кожного батчу, кроків часу та характеристик.

Розмір батчу, як правило, набагато менший, ніж загальна кількість зразків. Він, поряд із кількістю епох визначає, наскільки швидко мережа засвоює дані (як часто оновлюються ваги).

Параметром імпорту при визначенні рівня LSTM є кількість нейронів, яку також називають кількістю одиниць пам'яті або блоків. Це досить проста проблема, а число від 1 до 5 має як правило є достатнім.

Мережі потрібен один нейрон у вихідному шарі з лінійною активацією, щоб здійснити прогноз на момент часу $t + 1$.

Після того, як мережа задана, вона повинна бути скопільована в ефективне символічне представлення за допомогою серверної математичної

бібліотеки; в межах даного експерименту було була використана бібліотека TensorFlow.

Компілюючи мережу, потрібно вказати функцію втрат та алгоритм оптимізації: для була використана функція `mean_squared_error` як функція втрат, оскільки для подальшої оцінки прогнозу була використана середньоквадратична похибка (RMSE); в якості алгоритму оптимізації було використано ADAM.

З використанням API `Sequential Keras` нами було створена та скомпільована нейронна мережа.

За замовчуванням зразки протягом епохи перемішуються перед тим, як потрапити в мережу, що є небажано для LSTM, оскільки мета експерименту полягає у тому, щоб мережа створювала стан, коли вона вивчає послідовність спостережень. Для цього була відключена перетасовка зразків (параметр `shuffle`).

Були задані такі параметри:

- батчів 1
- епох 100
- нейронів 4.

Після оцінки моделлю LSTM навчальних даних, її можна використовувати для прогнозування. Для складання прогнозу використана функція `predict()`. Для цього потрібно як аргумент задати масив 3D NumPy: в даному випадку це буде масив з одним значенням спостереження на попередньому кроці часу. Функція `predict()` повертає масив передбачень для кожного введеного рядка введення. Оскільки в даному експерименті існує один вхід, то на виході генерується 2D-масив NumPy з одним значенням.

Підсумовуючи вищесказане алгоритм прогнозування в даному експерименті полягає в наступному:

1. завантаження ряду даних із файлу CSV;
2. прероцесінг набору даних з метою його використання в моделі LSTM, зокрема:

- перетворення даних для навчання з учителем;
 - оцінка стаціонарності;
 - перетворення даних таким чином, щоб вони мали масштаб -1 до 1.
3. підгонка (fitting) моделі мережі LSTM на навчальних даних;
 4. оцінка статичної моделі LSTM на основі тестових;
 5. оцінка результатів прогнозів (RMSE).

Файл Лог 1

- 1500 (навчання) + 500 (прогноз)
- Ряд даних переформатований у ряд з учителем (supervised learning)
- Метод LSTM (Довга короткочасна пам'ять) з класу методів рекурентних нейронних мереж (RNN)

RMSE (середньоквадратичне відхилення) = 0.044

- Батчів 1
- Епох 100
- Нейронів 4
- Реалізовано в Python. Бібліотеки: Pandas, Keras, Tensorflow

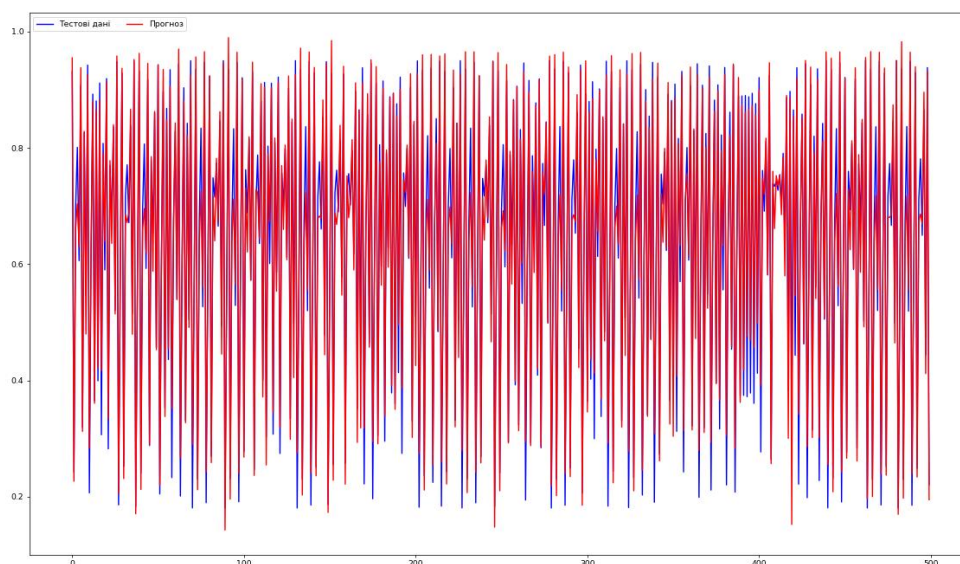


Рис. 3.4. Оцінка прогнозу часового ряду із врахуванням прогнозу,

RMSE = 0.44

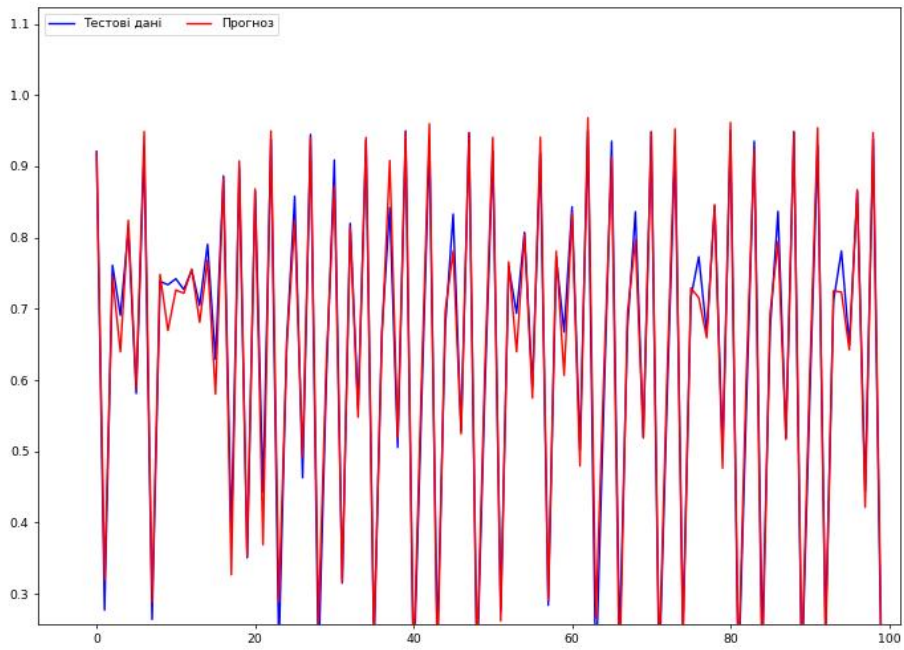


Рис.3.5. Рис. 3.4. Оцінка прогнозу часового ряду із врахуванням прогнозу,
 $RMSE = 0.03$

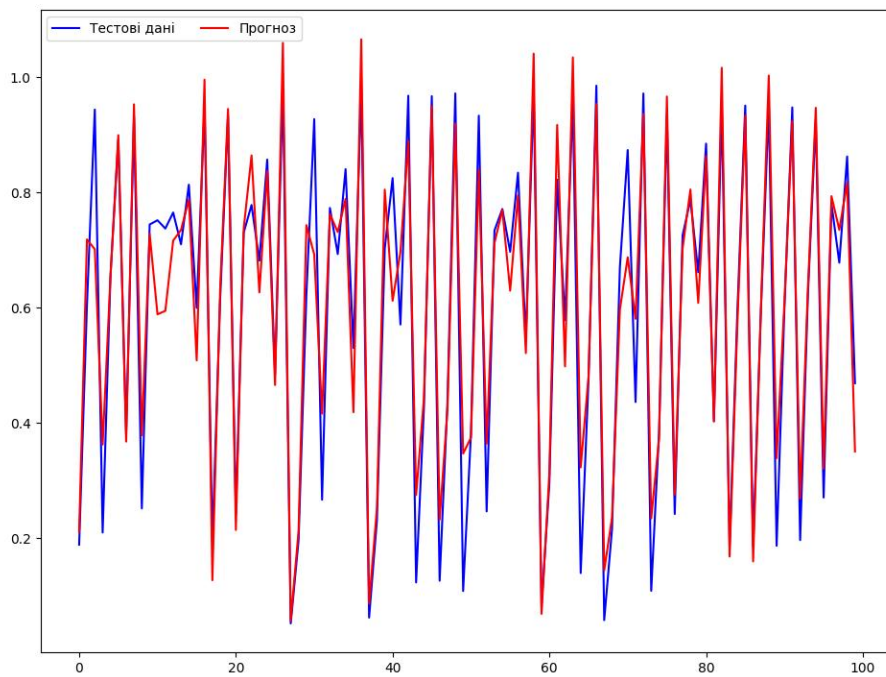


Рис. 3.6. Оцінка прогнозу часового ряду із врахуванням прогнозу,
 $RMSE = 0.085$

3.5. Числовий аналіз за допомогою нейронної мережі прямого поширення

Для того, щоб оцінити ефективність запропонованого методу, розроблена методика застосовується до трьох хаотичних часових рядів: Маккея-Глесса, логістичного та Ено. У цьому пункті для оцінки невідомої хаотичної функції використовується нейронна мережа прямого поширення (НМПП).

Рівняння Маккея-Глесса

Рівняння Маккея-Глесса спочатку було запропоновано як модель регулювання клітин крові. Часові ряди Маккея-Глесса використовувалися в літературі як еталонна модель через їх хаотичні характеристики. Диференціальне рівняння, який визначає часовий ряд, визначається наступним чином

$$\frac{dx}{dt} = \frac{ax(t-\tau)}{1+x^c(t-\tau)} - bx(t) \quad (3.3)$$

У рівнянні Маккея-Глесса параметр затримки τ визначає характер хаотичної поведінки рівняння. (3.3), тобто $\tau > 16,8$ за умови наявності хаосу. Надалі будемо розглядати наступні значення констант $a = 0,2, b = 0,1$ та $c = 10$, а хаотичний часовий ряд генерується $\tau = 17$ і початковим значенням $x(0) = 1,2$.

Хаотичний набір даних з 1000 точками даних генерується рівнянням (3.3) і реконструюється у точки фазового простору. Перші 500 точок даних використовуються для навчання, а решта 500 точок даних зберігаються для перевірки ефективності запропонованого методу.

Розмірність вкладення D і затримка часу T обчислюються за допомогою TSTOOL, додаткового набору інструментів для MATLAB. Часові ряди відтворюються з $D = 3$ та $T = 7$. Дані у часовій області та фазовому просторі представлені на рис. 3.7. Точки вкладеного фазового простору $Y(t)$ генеруються на основі перших 500 точок даних. Нейронна мережа прямого

поширення навчається з використанням реконструйованих точок фазового простору. Кількість вузлів прихованого шару вибирається методом спроб і помилок, що є найпоширенішим способом у літературі. Ваги та зміщення нейронних мереж зберігаються для прогнозування невідомих точок фазового простору. Зрештою прогнозовані фазові точки розкладаються для прогнозування майбутніх значень часового ряду, як показано на рис. 3.8.

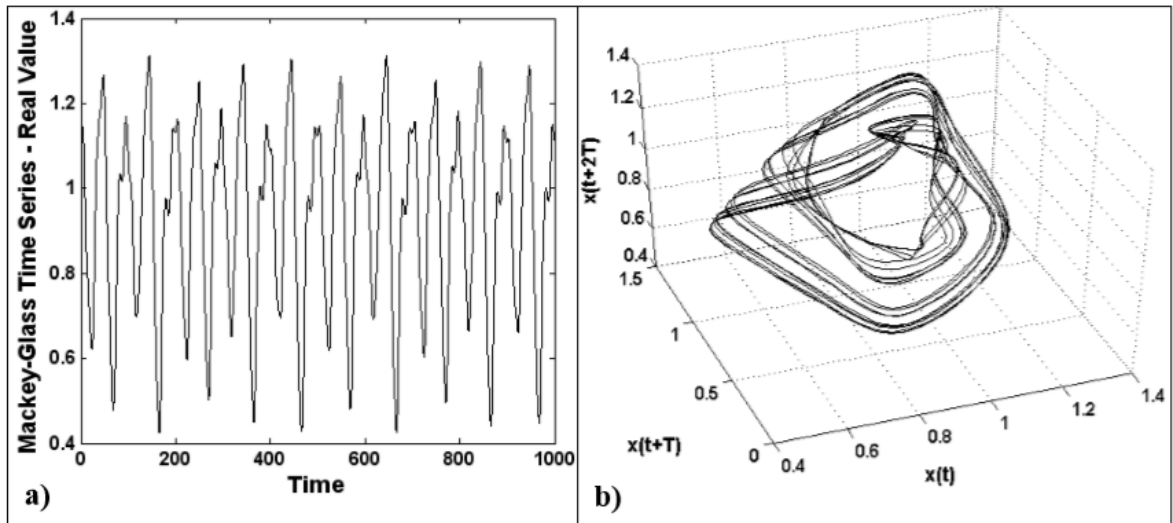


Рис. 3.7 Хаотичний часовий ряд Маккея-Глесса а) у часовій області, б) у фазово-просторовій області.

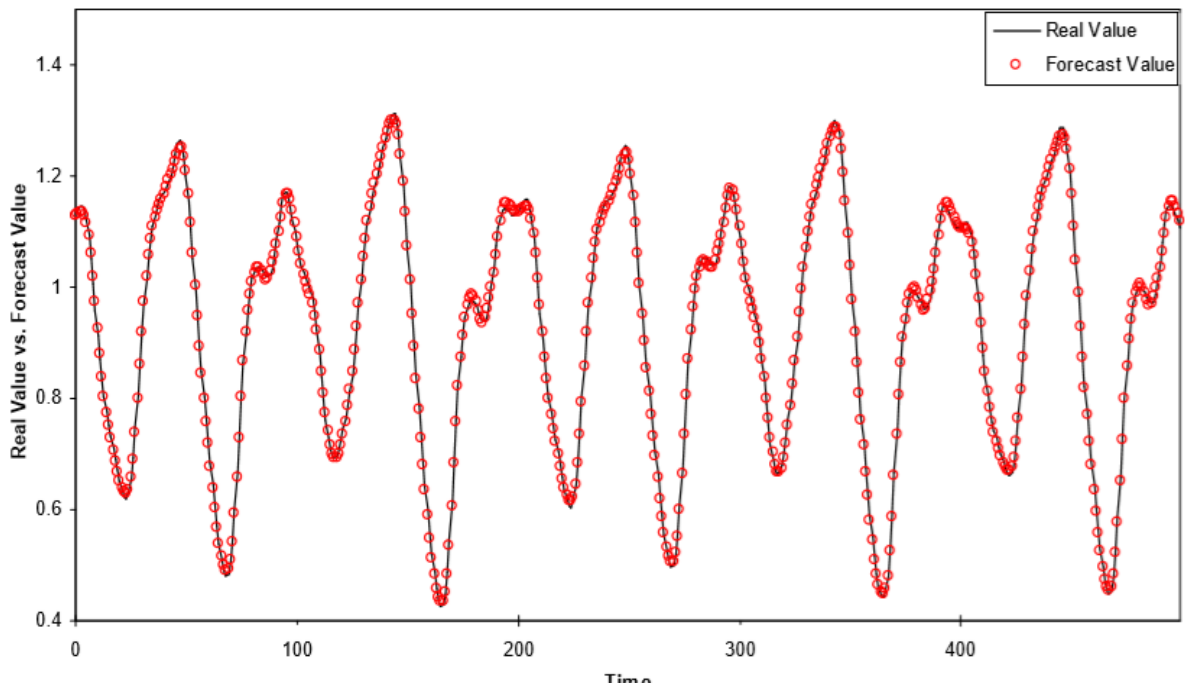


Рис. 3.8 Порівняння часових рядів Маккі-Глесса та значень прогнозу за допомогою НМПП

На основі вибраних дослідницьких праць для порівняння розраховуються індекси помилок. Для того, щоб оцінити ефективність прогнозування та порівняти її з результатами, наведеними в літературі, середньоквадратична похибка (RMSE) та нормалізована середньоквадратична похибка (NRMSE) обчислюються, де NRMSE обчислюється наступним чином

$$NRMSE = \left(\frac{\sum_{i=1}^N (y_i - \widehat{y}_i)^2}{\sum_{i=1}^N (y_i - \bar{y})^2} \right)$$

де y_i , \widehat{y}_i та \bar{y} – дані спостережень, прогнозовані дані та середні дані спостережень відповідно, а N – довжина даних спостережень. Значення RMSE та NRMSE відображені в таблиці 3.1.

Таблиця 3.1 Часовий ряд Маккея-Глесса – RMSE та NRMSE для 500 контрольних зразків

	RMSE	NRMSE
Запропонований метод – НМПП	5.98E-05	1.20E-03

Логістичне рівняння

Логістичне рівняння є найвідомішою, найстарішою і найпростішою хаотичною математичною системою, хоча воно має лише один параметр управління та одну змінну, він демонструє хаотичну поведінку. Логістичне рівняння було використано у дослідженнях комах та зростання популяції риб, проте фізик і математичний біолог Мей [104] виявив важливість логістичного рівняння як хаотичної системи. Логістичне рівняння записується має наступний вигляд

$$x(t + 1) = rx(t)[1 - x(t)]$$

з початковою умовою $x(0) = 0,36$ та значенням параметру $r = 4,0$. За допомогою даного рівняння формується хаотичний часовий ряд із 612 точками даних. Дані у часовій області показані на рис. 4.4. Розмірність

вкладення $D = 3$ і затримка часу $T = 1$ вибираються на основі Zhang та ін. [87], і реконструюється часовий ряд.

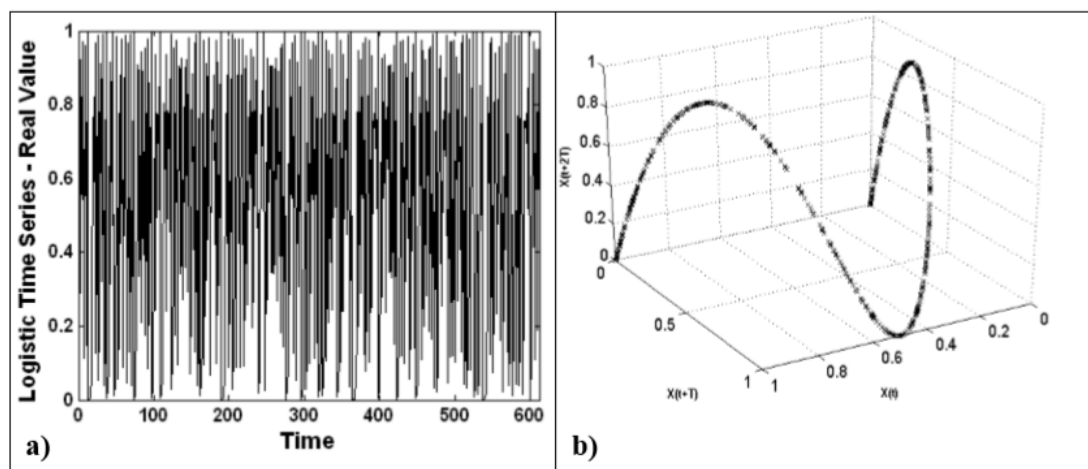


Рис. 3.9. Логістичний хаотичний часовий ряд а) у часовій області, б) у фазово-просторовій області

Точки вкладеного фазового простору $Y(t)$ отримуються на основі перших 512 точок даних. Нейронна мережа прямого поширення навчається з використанням точок фазового простору. Решта 100 точок даних зберігаються для перевірки ефективності запропонованого методу прогнозування. Ваги та зміщення зберігаються для прогнозування невідомих точок фазового простору, а отже, і майбутніх значень часового ряду, як показано на рис. 3.10.

Для оцінки ефективності запропонованого методу, застосованого до логістичного часового ряду обчислюються RMSE та NRMSE, показані в таблиці 3.2.

Таблиця 4.2 Логістичні часові ряди – RMSE та NRMSE для 100 контрольних точок даних

	RMSE	NRMSE
Запропонований метод – НМПП	4.00E-07	3.00E-06

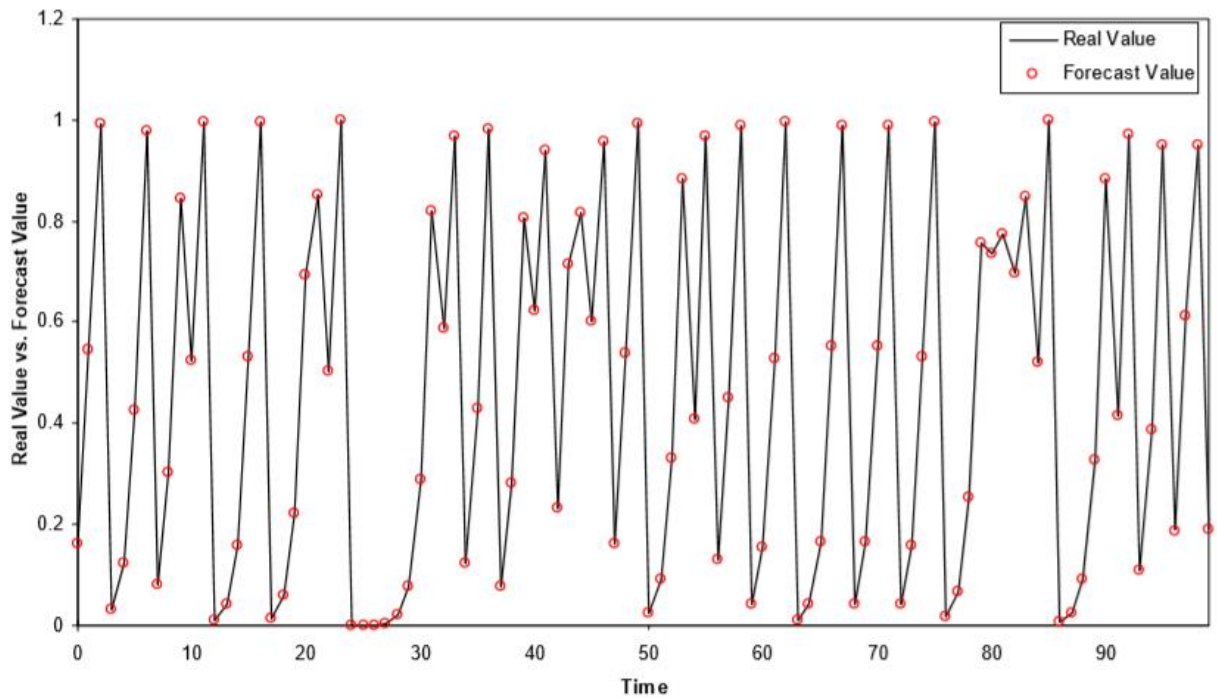


Рис. 3.10. Порівняння даних логістичних часових рядів та значень прогнозу з використанням НМПП.

Рівняння Ено

Рівняння Ено [55] було введено Ено та має простий формат, проте в ньому представлено багато аспектів динамічної поведінки складніших хаотичних систем. Рівняння Ено описується рівнянням, що породжує хаотичні часові ряди

$$x_{t+1} = Bx_{t-1} + 1 - Ax_t^2 \quad (3.4)$$

де $A = 1,4, B = 0,3, x_0 = 0,3$ та $x_1 = 0,3$.

Хаотичний часовий ряд із вибіркою 1000 генерується рівнянням (3.4). Перші 500 точок даних використовуються для навчання, а решта 500 точок зберігаються для тестування. Розмірність вкладення D і затримка часу T обчислюються за допомогою TSTOOL. Часові ряди відтворюються з $D = 2, T = 1$. Дані у часовій області та області фазового простору показані на рис. 3.11.

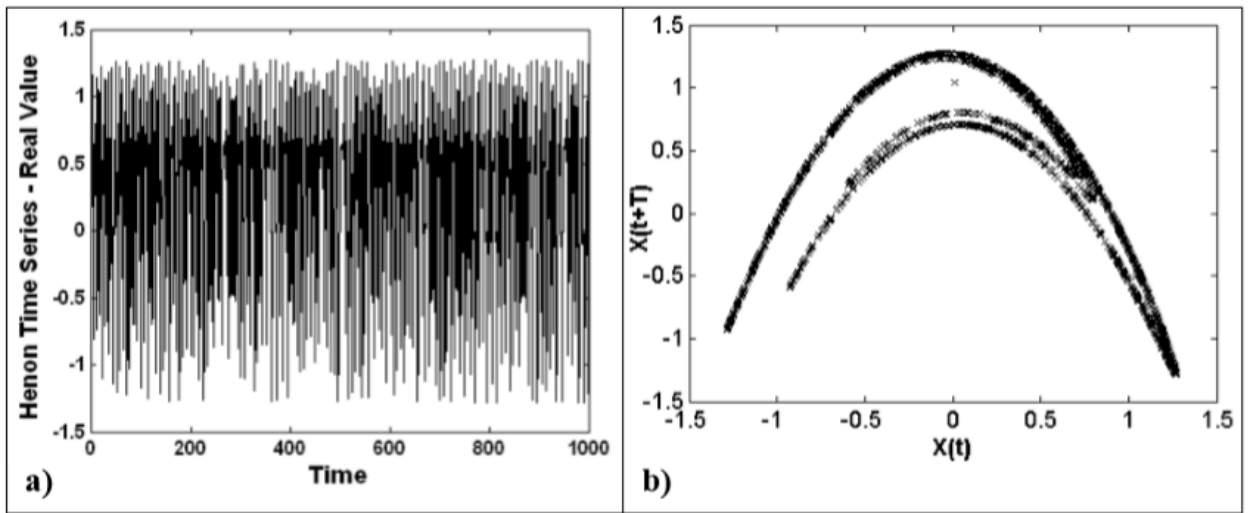


Рис. 3.11. Хаотичні часові ряди Ено а) у часовій області, б) в області фазового простору.

Точки вбудованого фазового простору $Y(t)$ генеруються на основі перших 500 точок даних. Решта 500 точок даних зберігаються для тестування. Вибирається та навчається архітектура нейронної мережі прямого поширення. Кількість вузлів прихованого шару вибирається методом спроб і помилок. Ваги та зміщення нейронних мереж зберігаються для прогнозування невідомих точок фазового простору. Зрештою, прогнозовані фазові точки розкладаються для прогнозування майбутніх значень часового ряду, як показано на рис. 3.12.

Отримані RMSE та NRMSE порівнюються з дослідженнями, про які повідомляється у літературі, та відображаються у Таблиці 3.3. Результати показують, що запропонований метод може ефективно прогнозувати часові ряди Ено.

Таблиця 4.3 Часові ряди Ено – RMSE та NRMSE для 500 контрольних точок даних

	RMSE	NRMSE
Запропонований метод – НМПП	8.51E-04	1.6

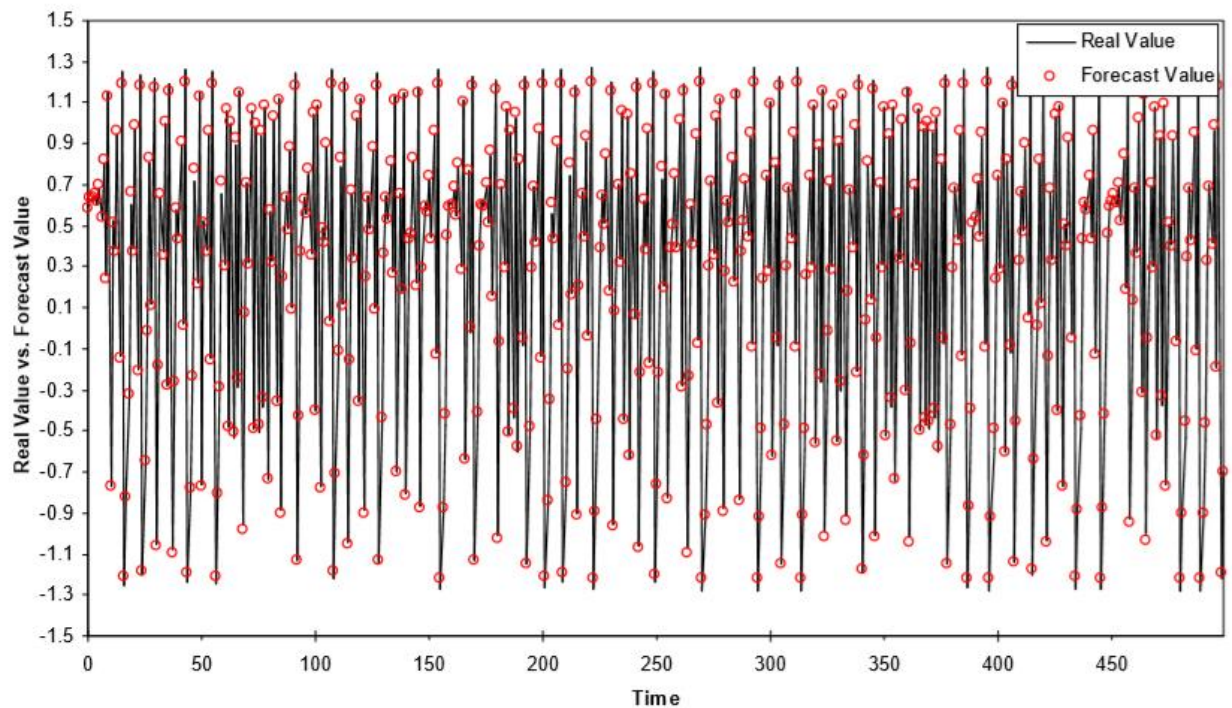


Рис. 3.12. Порівняння даних часових рядів Ено та значень прогнозу за допомогою НМПП

Числовий аналіз із використанням нейронної мережі Елмана

Наведені вище обчислення, які виконуються з нейронною мережею прямого поширення, повторюються з рекурентною нейронною мережею Елмана. Умови залишаються незмінними, щоб мати можливість порівняти результати. Єдина відмінність між рекурентною мережею Елмана та нейронною мережею прямого поширення полягає в механізмі зворотного зв'язку, який передає інформацію в прихованому шарі на вхідний рівень. На рисунках 3.13 – 3.15 показані результати запропонованого методу прогнозування за допомогою рекурентної мережі Елмана у разі застосування до часових рядів Маккея-Глесса (М-Г), логістичного рівняння та моделі Ено.

У таблиці 3.4 представлено результативність запропонованого методу прогнозування з використанням рекурентної мережі Елмана для трьох хаотичних часових рядів: Маккея-Глесса, логістичного та Ено.

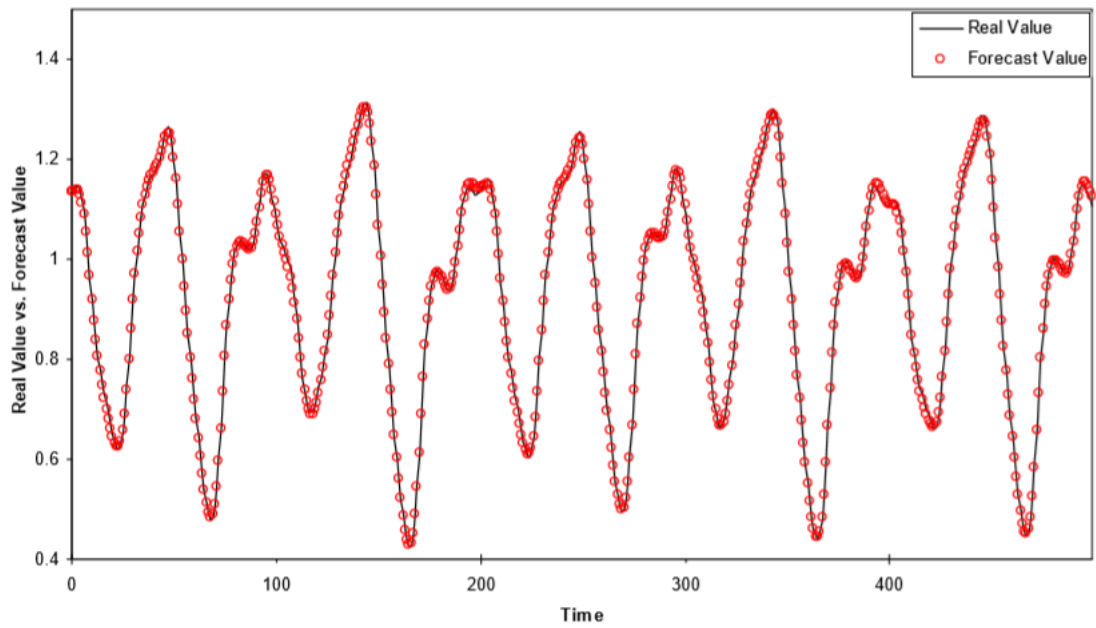


Рис. 3.13. Порівняння даних часового ряду М-Г та прогнозованих значень за допомогою ШНМ Елмана

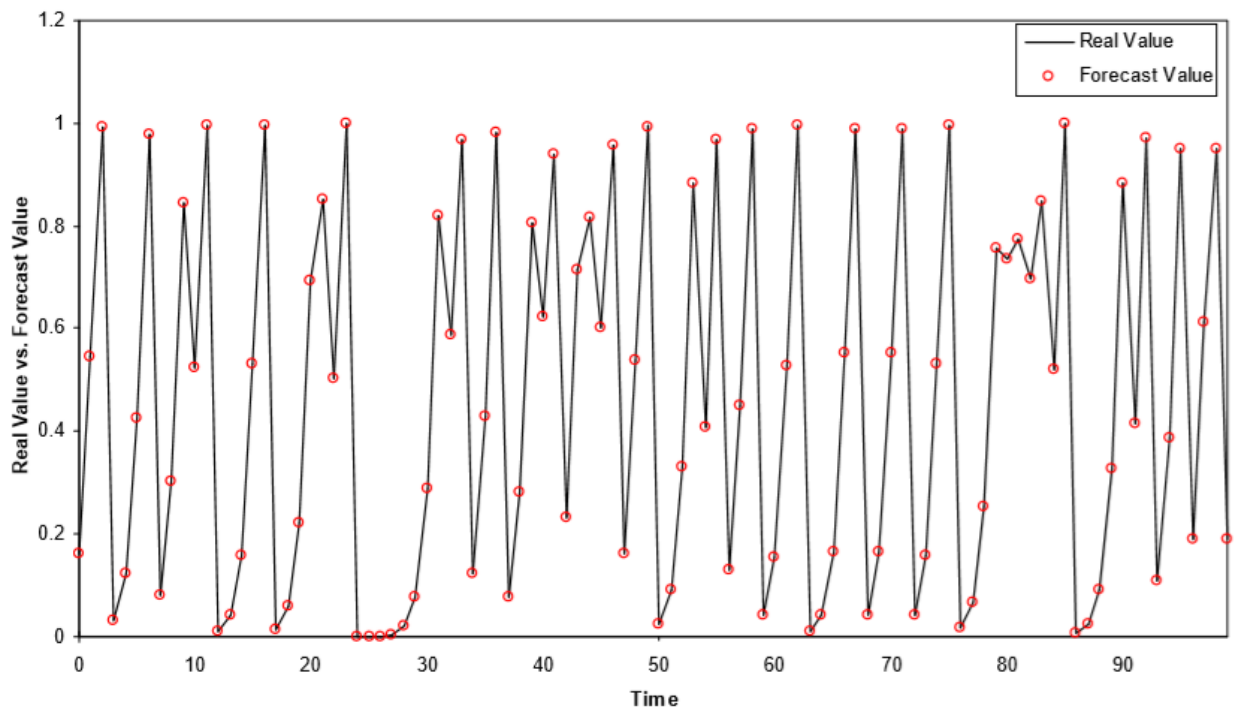


Рис. 3.14. Порівняння даних логістичного часового ряду та прогнозованих значень за допомогою РНМ Елмана

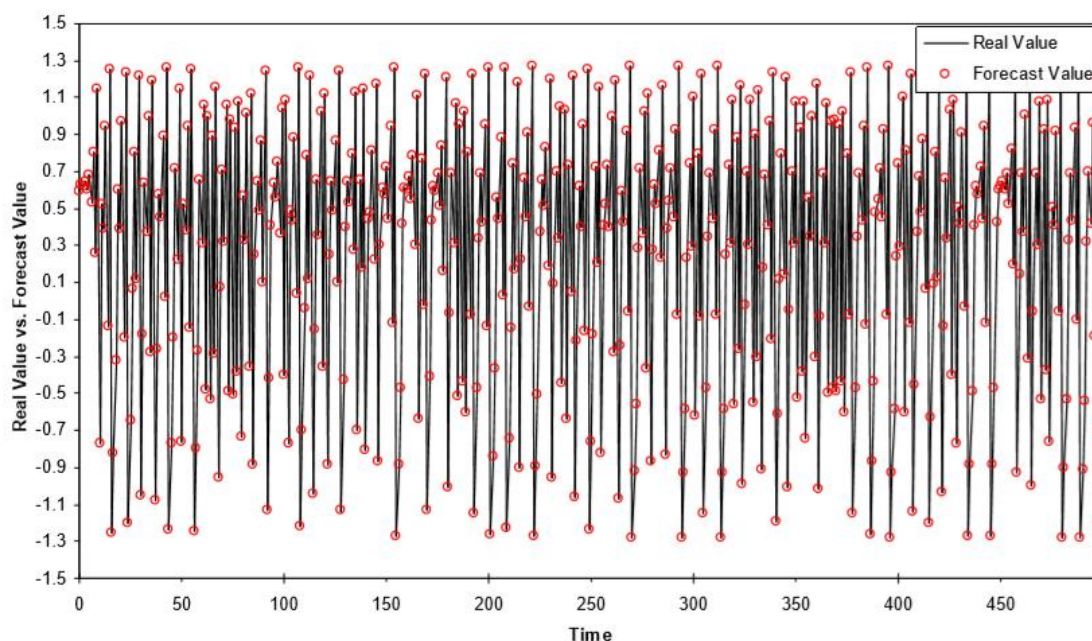


Рис. 3.15. Порівняння даних часового ряду Ено та прогнозованих значень за допомогою РНМ Елмана

Таблиця 3.4 Ефективність прогнозування з використанням РНМ Елмана

Часовий ряд	RMSE	NRMSE
Маккея-Глесса	4.12E-05	8.05E-04
Логістичне рівняння	2.44E-11	1.89E-10
Модель Ено	5.41E-09	1.01E-08

Покращення результатів за допомогою генетичних алгоритмів

У більшості алгоритмів навчання початкові ваги та зміщення нейронної мережі вибираються випадковим чином. Це спричинить варіації у вихідних даних нейронної мережі. Щоб зменшити варіації виходу нейронної мережі та покращити продуктивність прогнозування, для пошуку відповідних початкових ваг та зміщень прийнятними є генетичні алгоритми (ГА). Метою ГА є мінімізація середньої квадратичної помилки (RMSE) між прогнозованими значеннями та вихідними значеннями. Змінні ГА – це ваги

та зміщення нейронної мережі. Коли застосовується ГА, вибираються найкращі ваги та зміщення, які покращують RMSE.

У таблиці 3.5 наведено результативність запропонованого методу прогнозування, покращеного за допомогою генетичних алгоритмів як для нейронних мереж прямого поширення, так і для рекурентних нейронних мереж Елмана у разі застосування до трьох хаотичних часових рядів: Маккея-Глесса, логістичного та Ено.

Таблиця 3.5. Ефективність прогнозування, покращена за допомогою генетичних алгоритмів

Часовий ряд	RMSE	NRMSE
Маккея-Глесса (пряме поширення помилки)	5.24E-05	1.0E-03
Маккея-Глесса (Елмана)	3.67E-05	7.16E-04
Логістичне рівняння (пряме поширення помилки)	2.54E-08	1.90E-07
Логістичне рівняння (Елмана)	1.62E-11	1.21E-10
Модель Ено (пряме поширення помилки)	6.53E-04	1.20E-03
Модель Ено (Елмана)	4.76E-09	9.01E-09

Результати, отримані в цьому розділі на основі запропонованого методу, порівнюються з результатами, представленими в літературі. Для часового ряду Маккея-Глесса результативність запропонованого в цьому розділі методу порівнюється з результатами простої нейронної мережі прямого поширення та методами, запропонованими Chen та ін. [87]. У методі, запропонованому в Chen та ін. [87] використовується методика вкладення координат затримки в систему нечіткого виходу на основі адаптивної мережі з багатьма входами і виходами (MANFIS) для прогнозування хаотичних часових рядів. Серед іншого автори використовували рекурентну нейронну мережу Елмана для прогнозування хаотичних часових рядів. Щоб поліпшити свої результати, авторами було запропоновано багатовимірний метод, який

поєднує вихід кількох навчених нейронних мереж з різними розмірностями вкладення.

У таблиці 3.6 порівнюється ефективність запропонованого методу з результатами, представленими в літературі. На основі цього порівняння запропонований метод з використанням рекурентної нейронної мережі Елмана, вдосконалений генетичними алгоритмами, демонструє кращу продуктивність.

Таблиця 3.6 Часові ряди Маккея-Глесса – порівняння ефективності прогнозування

Метод прогнозування	RMSE
Нейро-нечіткі	1.26E-03
РНМ – D = 2	1.02E-03
РНМ – D = 3	4.15E-04
РНМ – Багатовимірні	3.81E-04
Прості прямого поширення	9.75E-04
Запропонований метод – Елмана	4.12E-05
Запропонований метод – Елмана та ГА	3.67E-05

Для логістичних часових рядів результативність запропонованого методу порівнюється з результатами простої нейронної мережі з прямим поширенням та результатами, представленими у дослідницьких роботах Chen та ін. [87]. У даній роботі розроблено метод прогнозування, заснований на фундаментальних характеристиках хаотичних систем, які є чутливістю до початкових умов. Показник Ляпунова забезпечує кількісне вимірювання динаміки хаотичної системи в різних розмірностях вкладення. У цьому методі використано модель показника Ляпунова для прогнозування невідомих точок фазового простору і, отже, невідомих значень у часових рядах.

У [33] запропоновано метод прогнозування на основі еволюціонуючих рекурентних нейронних мереж (ERNN), яким оцінюються належні параметри вкладення для реконструкції хаотичного часового ряду та покращується структура РНМ за допомогою еволюційних алгоритмів.

У таблиці 3.7 порівнюється ефективність запропонованого у цьому дослідженні методу прогнозування у разі застосування до логістичного часового ряду з результатами, представленими в літературі. Поєднання методу РНМ Елмана та ГА показує найкращий результат, представлений у таблиці 3.7.

Таблиця 3.7 Логістичний часовий ряд – порівняння ефективності прогнозування

Метод прогнозування	RMSE
Метод Ляпунова – $D = 2$	1.44E-02
Метод Ляпунова – $D = 3$	3.88E-02
Метод Ляпунова – Багатовимірний	3.18E-04
ЕРНМ	6.45E-10
Прості прямого поширення	3.98E-04
Запропонований метод – Елмана	2.44E-11
Запропонований метод – Елмана та ГА	1.62E-11

Для часових рядів Ено результативність запропонованого у цьому дослідженні методу прогнозування порівнюється з результатами простої нейронної мережі прямого поширення і результатами, отриманими з регресії опорного вектора (РОВ), багат шарового перцептрона із відгалуженою лінією затримки (TDL-MLP), розподілені локальні моделі (DLM), розподілені місцеві експерти на основі векторного квантування з використанням теоретико-інформаційної концепції (DLE-VQIT), ансамблеві самогенеруючі нейронні мережі (ESGNN), ГА поєднані з ансамблевими самогенеруючими нейронними мережами (GAESGNN) та генетичні вейвлет-нейронні мережі (GWNN). Метод рекурентних нейронних мереж Елмана, покращений за допомогою генетичних алгоритмів, показує найкращі результати порівняно з іншими методами, як показано в Таблиці 3.8.

Таблиця 4.8 Часовий ряд Ено – порівняння ефективності прогнозування

Метод прогнозування	RMSE
SVR	2.10E-02
TDL-MLP	7.59E-02
DLM	1.01E-01

DLE-VQIT	1.92E-02
ESGNN	2.23E-03
GAESGNN	2.01E-03
GWNN	8.42E-05
Прості прямого поширення	3.2E-03
Запропонований метод – Елмана	5.41E-09
Запропонований метод – Елмана та ГА	4.76E-09

Висновки до розділу 3

Першим кроком для аналізу системи є виявлення природи системи. Системи поділяються на три основні категорії: стабільні, періодичні та хаотичні. Хаотичні системи можна розпізнати по тому, що вони детерміновані, обмежені, аперіодичні та чутливі до початкових умов, змінюваних областей, що мають хаотичні закономірності та хаотичний атрактор, роблячи недійсним підхід редуccionізму та недооцінюючи точність комп'ютера. Теорія хаосу може математично описати складні системи. Теорема вкладення Такенса стверджує, що в хаотичних системах точки фазового простору слідуєть за невідомою динамікою вихідної системи. Послідовності скалярних значень у часі, які приймаються через рівні проміжки часу і мають хаотичні характеристики, називаються хаотичними часовими рядами.

Метою цього розділу було представити інструменти й теорії, необхідні для розробки методу прогнозування, запропонованого у цій дисертації. Для визначення існування хаосу в системі були представлені чотири методи: перетворення Фур'є, показник Ляпунова, показник Герста і фрактальна розмірність. Застосування даних показників відображено у наступних пунктах. Виходячи зі теоретичних викладок розділу зрозуміло, що для застосування теореми вкладення потрібні відповідна затримка часу та розмір вкладення.

Штучні нейронні мережі – це нелінійні моделі, які імітують нейронну систему людини та виконують функції самоорганізації, керування даними,

самонавчання, самоадаптації та пов'язаної з ними пам'яті. ШНН здатні фіксувати приховані функціональні зв'язки між заданими даними. Нейронна мережа прямого поширення є прикладом статичних нейронних мереж, а мережі Елмана та NARX є прикладами динамічних нейронних мереж. Генетичні алгоритми як еволюційний метод, що імітує процес природного відбору, можна використовувати з метою оптимізації. Генетичні алгоритми, розглянуті в п 3.4, можуть бути використані для покращення результату роботи нейронних мереж.

Метод реконструкції фазового простору використовується для реконструкції точок фазового простору з хаотичних часових рядів з використанням розмірності вкладення D та часової затримки T . Відповідно до методу реконструкції фазового простору, точки фазового простору відтворюють ключові характеристики вихідного часового ряду. Для прогнозування майбутнього значення точок фазового простору, а отже, і хаотичних часових рядів, були використані штучні нейронні мережі, які можуть фіксувати відносини між даними, коли основна функція невідома або її складно визначити. Для прогнозування на один крок вперед було обрано нейронні мережі прямого поширення та Елмана. Для підтвердження ефективності запропонованої методики, вона була застосована до трьох різних хаотичних часових рядів: Маккея-Глесса, логістичного та Ено. Запропонований метод об'єднувався з генетичними алгоритмами для покращення результатів прогнозування. Було помічено, що новий метод прогнозування, покращений за допомогою GA, демонструє визначні результати. Порівняння результатів запропонованої методики з результатами, наведеними в літературі, обґрунтовує ефективність запропонованого методу.

Дость цікавим є використання машинного навчання для прогнозування фінансових хаотичних часових рядів, що показало високі результати з дуже малими похибкам. Таким чином, в майбутніх дослідженнях є можливість не враховувати аналітичні рівняння для опису детермінованого хаосу. Це в свою

чергу забезпечить неможливість розкриття природи хаотичних рівнянь при прогнозуванні фінансових та інших часових рядів.

4.1. Неперервний аналог часових рядів та відповідна оптимізаційна задача

4.1.1. Мотивація використання часових рядів

Аналіз та прогнозування складних систем являється однією із найбільш досліджуваних задач різних прикладних областей математики [11, 25, 39, 141, 148, 148]. Розглянемо найбільш досліджувані моделі, які використовуються при дослідженні складних систем. Насамперед, слід відмітити часові ряди, які відіграють одну із важливіших ролей у моделюванні та прогнозуванні фінансових часових рядів [26, 144]. Використання часових рядів обумовлене простотою задання моделі та інтерпретацією результатів оцінки параметрів моделі [156], що включає в себе велику кількість тестів для визначення стаціонарності досліджуваного процесу, наявності одиничного кореня, перевірки гомоскедастичності залишків тощо. Немає сумніву у тому, що використання саме часових рядів і стало відправною точкою для розробки багатьох інших математичних моделей, що в деякій мірі є розширюють сферу використання останніх. Нижче будуть наведені декілька основних моделей, які являються продовженням часових рядів та є об'єктом дослідження даної роботи.

Найперше розглянемо, так звані неперервні авторегресійні моделі (continuous autoregressive models, CAR model) або неперервні авторегресійні моделі рухомого середнього (continuous parameter ARMA, CARMA model) [26, 30, 68, 73, 98, 144, 156], основна увага яких базується на припущенні про неперервність досліджуваного процесу. Найчастіше процесу такого описуються за допомогою стохастичних диференціальних рівнянь Іто, оскільки дані рівняння є природнім продовженням звичайних

диференціальних рівнянь та містять випадковості у вигляді інтегралу Іто. Наприклад, у роботі [98] основна робота базується на припущенні, що реальний процес описується за допомогою стохастичного диференціального рівняння Іто

$$dr(t) = \mu(t, r(t))dt + \sigma(t, r(t))dW(t), \quad (4.1)$$

де $W(t), t \geq 0$ – стандартний вінерівський процес відносно фільтрації $\mathcal{F} = \{F_t\}_{t \geq 0}$. У даній роботі розглянуті моделі Кокса-Росса-Інтерсолла, Холла – Уайта та деякі моделі типу процесів Орнштейна – Уленбека. У роботі [156] розглядаються CARMA моделі, які задаються за допомогою наступної моделі

$$\begin{aligned} B^p x(t) + a_1 B^{p-1} x(t) + \dots + a_p x(t) = \\ = b_0 B W(t) + b_1 B^2 W(t) + \dots + b_q B^{q+1} W(t), \end{aligned} \quad (4.2)$$

де B – лаговий оператор, що задається наступним співвідношенням

$$Bx(t) = x(t - 1).$$

Слід зауважити, що рівняння (4.2) може бути переписане у вигляді векторного рівняння (4.1), причому коефіцієнт біля dt буде лінійним функціоналом від розв'язку x на інтервалі $[t - p, t]$. За рахунок лінійності коефіцієнту біля dt , розв'язок рівняння (4.2) володіє рядом властивостей, що характеризують велику кількість реальних процесів, які зустрічаються на фінансовому ринку. Наприклад, коваріаційна функція для процесу (4.2) визначається співвідношенням

$$\gamma(h) = \text{cov}(x(t), x(t + h)) = \sum_{i=1}^p \beta_i e^{\lambda_i |h|},$$

де λ_i – власні значення допоміжної матриці рівняння (4.1), що відповідає рівнянню (4.2), $\beta_i, i = 1, \dots, p$ – деякі коефіцієнти. На відміну від роботи [156], у роботі [145] розглянуто моделі типу Орнштейно – Уленбека

$$B^p x(t) + a_1 B^{p-1} x(t) + \dots + a_p x(t) = \sigma W(t). \quad (4.3)$$

Детально розглянуто методи оцінок параметрів моделі (4.3), властивості кореляційної матриці процесі та підходи до визначення невідомого

параметру p , оцінка якого є першим кроком до оцінки ARIMA моделей чи їх модифікацій.

Основною моделлю роботи [30] є модель вигляду

$$x(t) = \int_{-\infty}^0 h(t-s; \theta) dW(s), \quad (4.4)$$

де $h(t-s; \theta)$ – деяка відома функція залежна від відомого невідомого параметру θ . Таким чином, рівняння (4.4) можна інтерпретувати як розширення рівняння (4.2) на неперервний випадок функції $h(t-s; \theta)$. Увага в даній роботі присвячується оцінці невідомого параметру θ на основі методу максимальної правдоподібності.

Деякий інший підхід, що відображає врахування екзогенних змінних як додаткового джерела невизначеності, відображений в роботах [8, 26]. У роботі [26] класична ARMA модель розвинута до ARMA моделі із врахуванням прихованих марковських процесів, тобто розглядається часовий ряд

$$x(t) - m = \sum_{i=1}^p \phi_i(\xi_t)(x(t-i) - m) + \sum_{j=0}^q \theta_j(\xi_t)\varepsilon_j \quad (4.5)$$

де ξ_t – дискретний ланцюг Маркова з матрицею переходу за 1 крок P , причому елементи матриці P задаються рівностями

$$P_{ij} = P(\xi_t = j | \xi_{t-1} = i), i, j \in S,$$

S – множина станів дискретного ланцюга Маркова ξ_t . Таким чином, розширена модель (4.5) враховує переключення, які інколи називають зовнішніми до залишків моделі ε_j . Робота [8] розширює поняття класичних ARIMA моделі на процеси Levy. Аналогічний підхід, проте із напівмарковськими процесами, запропоновано авторами роботи [147].

Ще одним напрямком розширення моделей часових рядів можна рахувати багатопарові моделі, в яких прогнозування проводиться на основі багаторівневих моделей або нейронних мереж. У роботі [155] розроблено нейронну мережу, зв'язки у якій представляють собою не класичні лінійні

моделі із функціями активації, а випадковими процесами, параметри яких залежать від оцінюваних процесів у попередні моменти часу. Таким чином, у даній роботі описано симбіоз нейронних мереж із часовими рядами. Наведено алгоритм оцінки параметрів моделі та доведено достатні умови збіжності відповідного алгоритму, який базується на байєсівському підході. Автор роботи [155] також звернув увагу на оцінку параметрів часових рядів, що разом із нейронними мережами дозволяють будувати досить точні прогнози для деяких реальних фінансових ринків. Споріднені результати із аналізом динаміки капіталу страхової компанії та її банкрутства, розглянуті в роботі [154]. Робота [9] присвячена дослідження тестів на перевірку лінійності моделей, що використовуються в нейронних мережах.

Підсумовуючи розглянуті вище моделі, можемо стверджувати, що перехід від класичних дискретних моделей до більш загальних неперервних моделей відбувається з використанням різних математичних моделей, включаючи стохастичні диференціальні рівняння та моделі нейронних мереж. Найбільшою перевагою неперервних моделей являється відсутність припущення про однакові проміжки між спостереженнями процесу. Тобто для оцінки параметрів чи гіперпараметрів неперервних моделей часи спостережень

$$x(t_1), \dots, x(t_N)$$

не повинні задовольняти класичне співвідношення

$$t_n - t_{n-1} = t_m - t_{m-1},$$

що дозволяє розглядати дані з пропусками (gaps). Крім того, сучасні методи оцінки вартості активів дозволяють відслідковувати їх динаміку практично неперервно, що знову ж є фактором, для використання неперервних моделей.

4.1.2. Основна модель та допоміжний результат

Найбільш поширеними моделями при дослідженні є моделі лінійних часових рядів, особливо $ARMA$ модель та її модифікації [17], тобто досліджуваний процес $x(t), t \geq 0$ задовольняє співвідношення

$$x(t) - m = \sum_{i=1}^p \phi_i (x(t-i) - m) + \sum_{j=0}^q \theta_j \varepsilon_j,$$

де $\phi_i, \theta_j, m \in R$; p, q – відомі параметри $ARMA$ моделі, ε_j – незалежні однаково розподілені випадкові величини із нульовим математичним сподіванням та скінченною дисперсією $\sigma^2 < \infty$. Проте, як було показано в багатьох дослідженнях, класична $ARMA(p, q)$ модель недостатньо точно визначає динаміку реальних процесів, оскільки порушуються умови гомоскедастичності дисперсії, незалежності залишків тощо. Крім того, $ARMA(p, q)$ модель є лінійною моделлю, що ускладнює вибір моделі для реальних нелінійних систем.

Розглянемо основну модель, на основі якої будемо розглядати оцінку та побудову прогнозів для реальних процесів. Основною математичною моделлю, розглянутою в даній роботі буде нелінійне стохастичне диференціально – функціональне рівняння Іто, що задається наступним співвідношенням

$$dx(t) = Lx_t dt + Gx_t dW(t) \quad (4.6)$$

за початковою умовою

$$x(t) = x_0(t), t \in [-h, 0], \quad (4.7)$$

де $x(t) \in R^d$, W – одновимірний стандартний вінерівський процес, $L, G: C([-h, 0]) \rightarrow R^d$ – функціонали, що зазвичай визначаються за допомогою інтегрального представлення,

$$Lu_t = \int_{-h}^0 l(t, s, u(t+s)) ds,$$

$$Gu_t = \int_{-h}^0 g(t, s, u(t+s)) ds.$$

У випадку лінійності функції l по третьому аргументу, отримуємо лінійний випадок. Проблема вибору моделі (4.6) базується на оцінці невідомих параметрів h , l та g . Параметр h відіграє роль максимального запізнення або передісторії, на основі якої конструюється значення процесу $x(t)$ на основі співвідношення (4.6). Для подальшого розгляду визначимо наступний детермінований процес

$$y(t) = Ex(t),$$

який задовольняє диференціально – функціональне рівняння

$$dy(t) = Ly_t dt. \quad (4.8)$$

Розглянемо деякі властивості розв'язків диференціально – функціональних рівнянь (4.6) та (4.8). Перша властивість стосується представлення розв'язку рівняння (4.6) на основі розв'язку рівняння (4.8) [76, 151].

Теорема 4.1. Нехай функція $l(t, s, x)$ є лінійною функцією за третім аргументом та не залежить від першого аргументу t . Тоді розв'язок задачі Коші (4.6), (4.7) має наступне представлення

$$x(t) = y(t) + \int_0^t X(t-s) G x_s dW(s), \quad (4.9)$$

де $y(t)$ – розв'язок задачі Коші (4.8), (4.7), $X(t)$ – розв'язок рівняння (4.8) за початковою умовою

$$X(t) = \begin{cases} 0, & t \in [-h, 0), \\ 1, & t = 0. \end{cases}$$

Доведення. Розглянемо диференціал випадкового процесу (4.9), використовуючи лінійність функціоналу L ,

$$dx(t) = d \left(y(t) + \int_0^t X(t-s) G x_s dW(s) \right) =$$

$$= dy(t) + d\left(\int_0^t X(t-s)Gx_s dW(s)\right).$$

Використовуючи властивості інтегралу Іто, диференціал від другого доданку буде рівним

$$\begin{aligned} & d\left(\int_0^t X(t-s)Gx_s dW(s)\right) = \\ &= X(t-t)Gx_t dW(t) + \int_0^t dX(t-s)Gx_s dW(s) = \\ &= Gx_t dW(t) + \int_0^t LX_{t-s}Gx_s dW(s) dt. \end{aligned}$$

Згідно нашого припущення про лінійність функціоналу L , формально можна

$$\int_0^t LX_{t-s}Gx_s dW(s) = L \int_0^t X_{t-s}Gx_s dW(s).$$

Таким чином,

$$\begin{aligned} dx(t) &= dy(t) + Gx_t dW(t) + L \int_0^t X_{t-s}Gx_s dW(s) dt = \\ &= Ly_t dt + Gx_t dW(t) + L \int_0^t X_{t-s}Gx_s dW(s) dt = \\ &= \left(Ly_t + L \int_0^t X_{t-s}Gx_s dW(s) \right) dt + Gx_t dW(t) = \\ &= L \left(y_t + \int_0^t X_{t-s}Gx_s dW(s) \right) dt + Gx_t dW(t) = \\ & \quad Lx_t dt + Gx_t dW(t), \end{aligned}$$

що і доводить твердження Теорема.

Теорема 4.1 буде нам потрібна в майбутньому для аналізу та оцінки параметрів запропонованої моделі. Надалі будемо користуватися лише стохастичними диференціально – функціональними рівняннями (4.6), коефіцієнти які будуть задовольняти умовам Теорема 4.1. Тобто, функціонал L буде мати наступне зображення

$$Lu_t = \int_{-h}^0 l(s, u(t+s)) ds = \int_{-h}^0 l(s) u(t+s) ds. \quad (4.10)$$

В подальшому нас буде цікавити оцінка моделі (4.6), а саме оцінка невідомої трійки

$$(h, l(s), g(t, s, x)). \quad (4.11)$$

на основі мінімізації наступного функціоналу якості

$$I(h, l, g) = \frac{1}{T} \int_0^T E \left(x(t) - x_{sample}(t) \right)^2 dt + h, \quad (4.12)$$

де $x_{sample}(t)$ – значення спостережуваного процесу, $[0, T]$ – час спостереження. Побудова функціоналу якості (4.12) базується на побудові квадратичного відхилення від спостережуваного значення та врахуванні величини запізнення h . Аналогічний підхід також і використовується при побудові інформаційних критеріїв Акаїке та Шварца.

4.1.3. Метод оцінки невідомих параметрів

Згідно побудови функціоналів L , G та результату Теорема 4.1, оцінку параметрів $(h, l(s), g(t, s, x))$ можна проводити на основі уточнення класів функцій $l(s)$ та $g(t, s, x)$. Одне із найбільш загальних припущення припущень щодо цих функцій є припущення щодо їх неперервності по всіх змінних, тобто

$$l \in C([-h, 0]), \quad g \in C([0, T] \times [-h, 0] \times R^d).$$

Проте дані класи є занадто широкими для практичного застосування. Тому в подальшому нами будуть вибрані більш простіші класи, які є найбільш вживаними в контексті застосування в прогнозуванні фінансових активів

$$l \in \mathbf{L}_{n,m} = \left\{ l \in C([-h, 0]): l(s) = l_0 + l_{1,1}s + \dots + l_{1,n}s^n + \right. \\ \left. + l_{2,1} \log(s+1) + \dots + l_{2,1} \log^m(s+1) \right\}$$

$$g \in \mathbf{G}_{n_t, m_t, n_s, m_s, k} =$$

$$= \left\{ \begin{array}{l} g \in C([0, T) \times [-h, 0] \times R^d): \\ g(t, s, x) = \sum_{k_0=0}^k x^{k_0} \left(\sum_{n=0}^{n_t} g_{k_0, n}^{t,1} t^n + \sum_{m=1}^{m_t} g_{k_0, m}^{t,2} \log^m(t+1) \right) * \\ \left(\sum_{n=0}^{n_s} g_{k_0, n}^{s,1} s^n + \sum_{m=1}^{m_s} g_{k_0, m}^{s,2} \log^m(s+1) \right) \end{array} \right\}.$$

Розмірності розглянутих просторів наступні

$$\dim(\mathbf{L}_{n,m}) = n + m + 1;$$

$$\dim(\mathbf{G}_{n_t, m_t, n_s, m_s, k}) = (k+1)(n_t + m_t + 1)(n_s + m_s + 1).$$

Шукаючи мінімальне значення функціоналу якості на просторах $\mathbf{L}_{n,m}$, $\mathbf{G}_{n_t, m_t, n_s, m_s, k}$, задача зводиться до параметричної задачі оптимізації на відміну від непараметричної задачі оптимізації на більш ширших просторах $C([-h, 0])$ та $C([0, T) \times [-h, 0] \times R^d)$ відповідно. Отже, звуживши простори пошуку нам вдалося перейти до параметричної задачі оптимізації, що значно полегшує наступні оцінки. Крім того, для обмежених значень $x \in R^d$, можна стверджувати що об'єднання множин $\mathbf{L}_{n,m}$ та $\mathbf{G}_{n_t, m_t, n_s, m_s, k}$ є всюдищільними в $C([-h, 0])$ та $C([0, T) \times [-h, 0] \times \{x: |x| < K\})$ відповідно для довільного $0 < K < \infty$, тобто для довільного $\varepsilon > 0$ та довільних $u_1 \in C([-h, 0])$, $u_2 \in C([0, \infty) \times [-h, 0] \times \{x: |x| < K\})$ існують значення параметрів $(n, m, n_t, m_t, n_s, m_s, k)$ та відповідних функцій $\tilde{u}_1 \in \mathbf{L}_{n,m}$ та $\tilde{u}_2 \in \mathbf{G}_{n_t, m_t, n_s, m_s, k}$ для яких виконується наступне співвідношення

$$|u_i - \tilde{u}_i| < \varepsilon, i = 1, 2.$$

Для оцінки невідомого вектору параметрів

$$par = (l_0, l_{1,1}, \dots, l_{1,n}, l_{2,1}, \dots, g_{0,0}^{t,1}, \dots, g_{m_s,k}^{s,2})$$

в даній роботі буде використано метод рою частинок [68], оскільки від виявився найбільш точним для досліджуваних даних.

4.1.4. Аналіз реальних даних

Для аналізу нової моделі використаємо три часові ряди, які відображають динаміку трьох фінансових процесів:

1. Вартість акцій компанії Apple – денні дані (1.05.2022 - 1.05.2023), 250 вимірів;
2. Обмінний курс USD-EUR – денні дані (1.05.2022 - 1.05.2023), 366 вимірів;
3. Вартість BTC(USD) – денні дані (1.05.2022 - 1.05.2023), 261 вимірів;

Динаміку розглянутих процесів проілюстровано на Рис. 4.1 – Рис. 4.3. Поряд із динамікою відповідних процесів також наведено також значення автокореляційних функцій (ACF) та частинних автокореляційних функцій (PACF). Слід відзначити, що у всіх трьох випадках процеси є нестационарні, про що свідчать результати тестів Дікі – Фуллера, Льюнга – Бокса та Квітковського – Філіпса – Шмідта – Шіна. Для першого (Apple) та третього (обмінний курс USD-BTC) процесів порядок інтегровності $d = 1$, для другого процесу (обмінний курс USD-EUR) значення порядку інтегровності рівне $d = 2$.

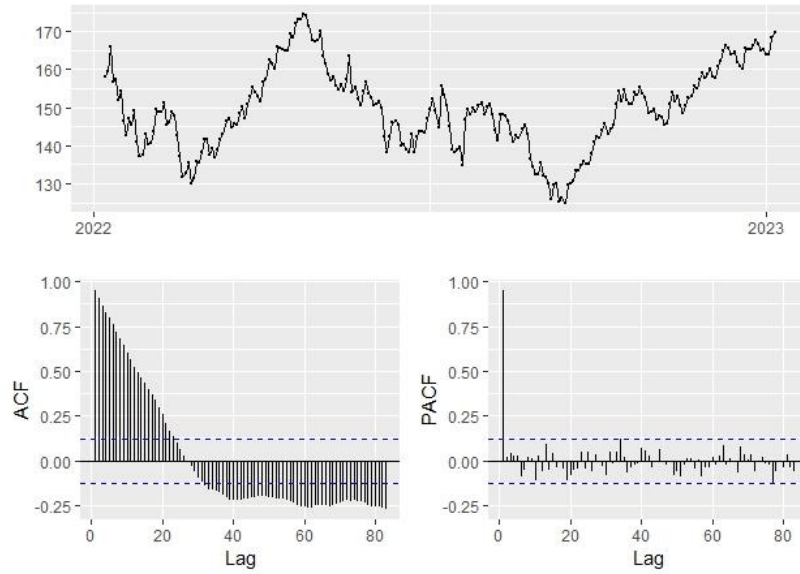


Рис. 4.1. Вартість акцій компаній Apple у період 1.05.2022 - 1.05.2023 разом із автокореляційною функцією (ACF) та частинною автокореляційною функцією (PACF)

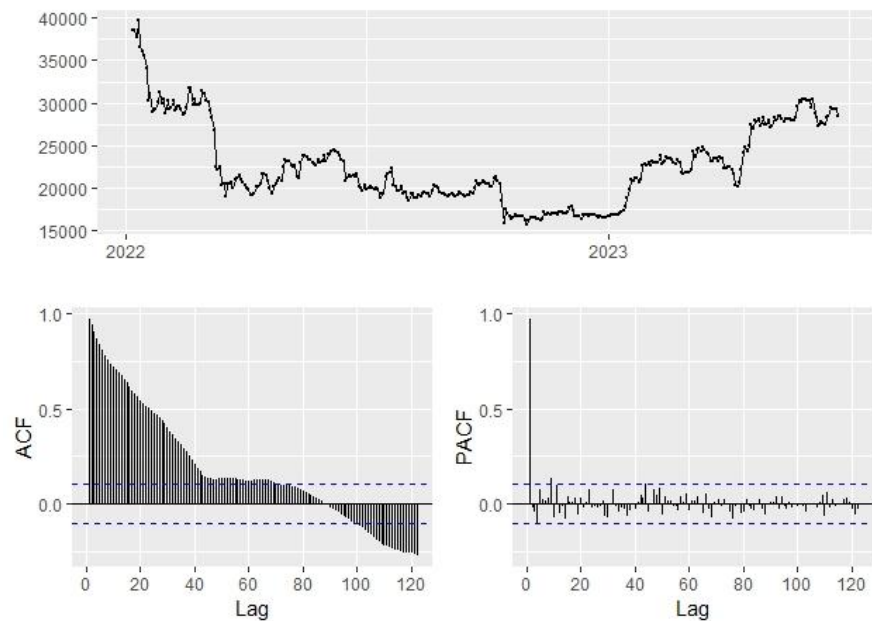


Рис. 4.2. Вартість BTC (у USD) у період 1.05.2022 - 1.05.2023 разом із автокореляційною функцією (ACF) та частинною автокореляційною функцією (PACF)

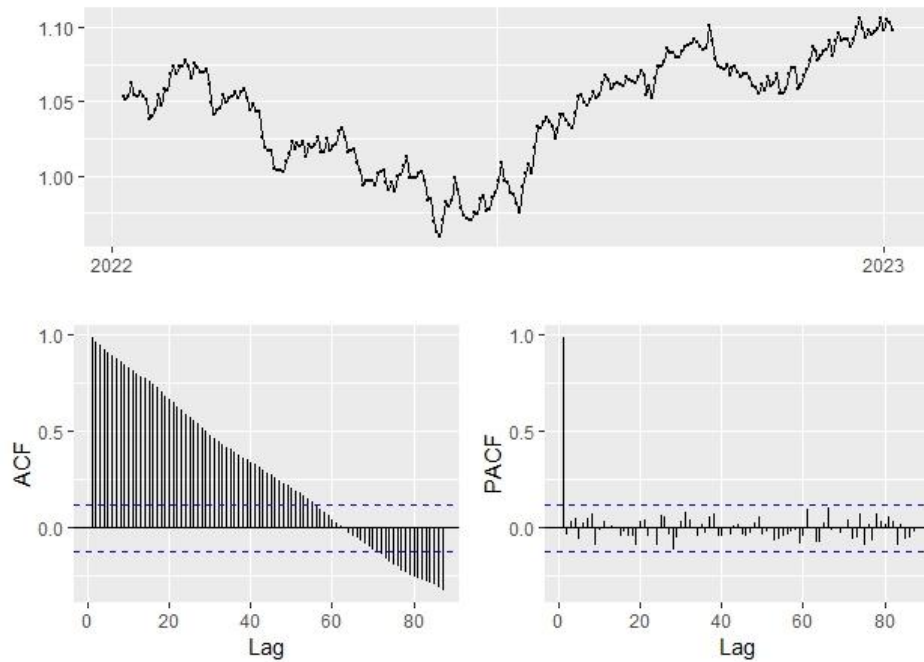


Рис. 4.3. Обмінний курс EUR/USD у період 1.05.2022 - 1.05.2023 разом із автокореляційною функцією (ACF) та частинною автокореляційною функцією (PACF)

Нас більше буде цікавити саме оцінка моделі на даних, а не врахування кількості параметрів у моделі, тому для перевірки точності моделей на тренувальних даних та порівняння прогнозів моделей на тестових даних нами буде використано три основні показники

$$RMSE = \sqrt{\frac{1}{T} \sum_{t=1}^T (y_t - \hat{y}_t)^2},$$

$$MAE = \frac{1}{T} \sum_{t=1}^T |y_t - \hat{y}_t|,$$

$$MAPE = \frac{1}{T} \sum_{t=1}^T \left| \frac{y_t - \hat{y}_t}{y_t} \right| * 100\%,$$

де T – кількість спостережень, y_t – реальні значення процесу у час t , \hat{y}_t – наближення y_t за моделлю або прогноз у час t .

Для аналізу нами вибрано оптимальну ARIMA моделі та дві GARCH специфікації: спрощену мультиваріантну sGARCH та узагальнену із

розривними дисперсіями gjrGARCH. Дані дві GARCH специфікації виявилися найбільш оптимальними, оскільки показують найменші значення $RMSE, MAE, MAPE$. Дані розділені на два інтервали – тренувальний у період з 1.05.2022 по 31.03.2023 та тестовий з 1.04.2023 по 1.05.2023. Результати порівняння моделей не тренувальних даних ми можемо бачити в Таб. 4.1. Як ми можемо спостерігати із даної таблиці на тренувальних даних всі три моделі дають приблизно однаковий результат, що може бути інтерпритоване як специфічний вибір моделі для середнього в GARCH моделі, що співпадає з оптимальною ARIMA моделлю. Проте аналіз показує [76, 151], що на тренувальних даних значення $RMSE, MAE, MAPE$ відрізняються більш суттєво.

Таб. 4.1. Порівняння числових характеристик для трьох процесів на тренувальних вибірках

	Apple			
	RMSE	MAE	MAPE(%)	logLik
ARIMA(4,1,1)	3.014875	2.234123	1.51676	-587.93
sGARCH(4,1,1)(1,1)	3.025881	2.222184	1.510279	-586.3425
gjrGARCH(4,1,1)(1,1)	3.025102	2.223523	1.511729	-585.1905
	BTC/USD			
	RMSE	MAE	MAPE(%)	logLik
ARIMA(0,2,1)	741.3196	463.484	2.030672	-2690.52
sGARCH(0,2,1)(1,1)	745.9248	458.2985	2.014739	-2662.567
gjrGARCH(0,2,1)(1,1)	744.017	457.5174	2.012519	-2661.796
	EUR/USD			
	RMSE	MAE	MAPE(%)	logLik
ARIMA(0,1,0)	0.006248042	0.004683876	0.4534139	855.18
sGARCH(0,1,0)(1,1)	0.006246513	0.004674132	0.4535177	855.2357
gjrGARCH(0,1,0)(1,1)	0.006246636	0.004676206	0.4537088	856.1199

Розглянемо тепер аналіз моделей, розглянутих вище, та нової запропонованої моделі (4.6) при $L_{n,m}$ та $G_{n_t, m_t, n_s, m_s, k}$, причому будемо вважати що всі константи в даних просторах рівні, тобто

$$n = m = n_t = m_t = n_s = m_s = k \in \{1, 2, 3, 4\}.$$

Результати оцінки показників якості моделі $RMSE, MAE, MAPE$ на тестовій вибірці показано в Таб. 5.2. Як ми можемо бачити із даної Табл., за допомогою нового методу прогнози є більш точними з використанням нового методу, описаного вище. Ще одним спостереженням є те, що оцінки точності при $n = 3$ та $n = 4$ практично співпадають, тобто використання просторів $L_{n,m}$ та $G_{n_t, m_t, n_s, m_s, k}$ із одночасним збільшенням всіх розмірностей є недоцільним. Крім того, для першого процесу (вартість акцій компанії Apple) при $n = 2$, точність є значно гіршою ніж для класичних лінійних моделей. Це можна пояснити тим фактом, що оптимальна ARIMA модель та її GARCH модифікації мають високий рівень залежності від передісторії, тобто порядку автокореляції $p = 4$. Тому для досягнення точності класичних моделей слід використати простори більш високих розмірностей. Проте при $n = 4$, точність у всіх розглянутих реальних процесах перевищує точність класичних ARIMA та GARCH моделей.

Таб. 4.2. Порівняння числових характеристик для трьох процесів на тренувальних вибірках

	Apple		
	RMSE	MAE	MAPE(%)
ARIMA(4,1,1)	3.683801	3.238859	1.9497963
sGARCH(4,1,1)(1,1)	2.253259	1.712294	1.0455451
gjrGARCH(4,1,1)(1,1)	2.161535	1.620379	0.9864241
Нова модель, $n = 1$	4.4890	4.21678	2.45854
Нова модель, $n = 2$	4.00785	2.148923	1.843941
Нова модель, $n = 3$	2.047652	1.634870	0.978273

Нова модель, $n = 4$	1.622162	1.100354	0.835698
BTC/USD			
	RMSE	MAE	MAPE(%)
ARIMA(0,2,1)	1771.8733	1438.0472	5.063041
sGARCH(0,2,1)(1,1)	1963.9042	1590.3372	5.613335
gjrGARCH(0,2,1)(1,1)	989.8264	886.2989	3.067350
Нова модель, $n = 1$	2851.3414	2051.3746	11.35465
Нова модель, $n = 2$	845.3452	765.4656	2.93745
Нова модель, $n = 3$	841.2938	753.3984	2.87374
Нова модель, $n = 4$	841.2938	753.3984	2.87374
EUR/USD			
	RMSE	MAE	MAPE(%)
ARIMA(0,1,0)	0.01297121	0.011561217	1.0525709
sGARCH(0,1,0)(1,1)	0.01089572	0.009608173	0.8746997
gjrGARCH(0,1,0)(1,1)	0.01151610	0.010181460	0.9268951
Нова модель, $n = 1$	0.164759	0.163545	3.4657567
Нова модель, $n = 2$	0.009845	0.009154	0.852643
Нова модель, $n = 3$	0.008754	0.008765	0.828991
Нова модель, $n = 4$	0.008754	0.008765	0.828991

4.2. Одне узагальнення LSTM нейронних мереж

4.2.1. Історія питання, мотивація

Використання стохастичних моделей як математичного апарату для побудови прогностичних систем набуло бурхливого розвитку у другій половині XX століття. Це пов'язано насамперед із швидким розвитком теорії стохастичних диференціальних рівнянь [106, 109, 143], часових рядів [17, 25, 26] та прикладних задач, що зумовили розвиток даних теоретичних

здобутків. Особливу увагу автори звертають на дослідження часових рядів, оскільки дані моделі володіють рядом переваг. По-перше, дані системи описують дискретні системи, що значно спрощує дослідження реальних систем та побудову прогнозів. По-друге, велика кількість реальних даних є дискретними, що описуються рекурентними формулами або послідовностями випадкових величин.

В останні два десятиліття широкого вжитку в прикладних задачах набули методи штучного інтелекту, глибинного інтелекту, машинного навчання. Цей феномен можна пояснити збільшенням обчислювальних потужностей для розв'язання прикладних задач, що дало можливість використовувати складні системи, такі як нейронні мережі та генетичні алгоритми. В даній роботі увагу буде звернено на узагальнення однієї із відомих нейронних мереж, а саме Long Short-Term Memory (LSTM) нейронної мережі [36, 57, 92, 96].

Теорія нейронних мереж є однією із найбільш розвинутих напрямків штучного інтелекту, машинного навчання та глибинного навчання. Поза сумнівом є той факт, що саме нейронні мережі є найбільш точними та якісними моделями в задачах класифікації реальних процесів різної природи. Слід зауважити, що вибір топології нейронної мережі буде залежати від конкретних даних та важко описується на основі тих чи інших характеристик даних. Тому процес підбору топології нейронної мережі є скоріше задачею дослідницького, ніж задачею теоретичного характеру. Даний факт насамперед диктується складністю опису характеристик вихідного сигналу (вихідних параметрів) для нейронних мереж із складною структурою, причому основна проблема полягає у наявності функції активації, яка нелінійно змінює сигнал між шарами нейронної мережі. У роботі буде розглянуто використання розширення LSTM нейронної мережі для побудови прогнозу фінансових активів. Слід зауважити, що дослідженню побудови прогнозів присвячена велика кількість робіт, які використовують різні припущення щодо властивостей реальних процесів.

У роботі [145] розглянуто прогноз динаміки захворюваності на COVID-19 у м. Києві та Чехії на основі марковських ланцюгів Маркова для основного процесу

$$S(t) = S(t - 1) - i(t)$$

та допоміжного процесу $i(t)$, що описуються нелінійним часом рядом, де $S(t)$ – кількість сприйнятливих до хвороби осіб в день t , $i(t)$ – кількість нових інфікованих осіб в день t . Слід зауважити, що саме допоміжний процес $i(t)$ будується на основі нелінійного часового ряду, що враховує передісторію процесу. Іншим підходом побудови прогнозів для цілого набору випадкових процесів є використання часових рядів [146]. В даній роботі розглянуто оптимізаційна задача для оцінки параметрів випадкового поля.

У роботі [142] використовується множинний підхід для оптимізації портфелю цінних паперів. Слід зауважити, що основним припущенням даної роботи є велика кількість допустимих стратегій, що унеможлиблює пряму перевірку всіх стратегій.

В даному розділі буде досліджено одне розширення LSTM нейронної мережі, які є потужним засобом для аналізу задач прогнозування часових рядів [41, 62], задач класифікації [1, 93], задач розпізнавання мовлення [90], задач розпізнавання рукописних текстів [133] тощо. Слід зауважити, що LSTM нейронні мережі є частинним випадком більш загальних рекурентних нейронних мереж. Останні в свою чергу будуються на принципі послідовного наближення вихідного сигналу із використанням ітераційного процесу із тими ж самими вагами на кожному кроці. Ґрунтуючись на роботі [57], топологію LSTM нейронної мережі можна описати наступними співвідношеннями

$$\begin{cases} f_t = \sigma_g(W_f x_t + U_f h_{t-1} + b_f), \\ i_t = \sigma_g(W_i x_t + U_i h_{t-1} + b_i), \\ o_t = \sigma_g(W_o x_t + U_o h_{t-1} + b_o), \\ \tilde{c}_t = \sigma_c(W_c x_t + U_c h_{t-1} + b_c), \\ c_t = f_t \odot c_{t-1} + i_t \odot \tilde{c}_t, \\ h_t = o_t \odot \sigma_h(c_t), \end{cases} \quad (4.13)$$

де

- $x_t \in R^d$ – значення вхідного сигналу,
- $f_t, i_t, o_t \in (0,1)^h$ – допоміжні вектори (вектори вентилів),
- $\tilde{c}_t \in (-1,1)^h$ – вхідний вектор активації клітинки,
- $c_t \in R^h$ – вихідний вектор комірки (вектор стану комірки),
- $h_t \in (-1,1)^h$ – вихідний сигнал,
- $W_f, W_i, W_o, W_c \in R^{h \times d}$, $U_x, U_i, U_o, U_c \in R^{h \times h}$, $b_f, b_i, b_o, b_c \in R^h$ - ваги нейронної мережі,
- d – розмірність вхідного сигналу, h - розмірність сигналу із прихованих шарів нейронної мережі,
- σ_g, σ_c – функції активації нейронної мережі,
- \odot – множення Адамара.

Формальне представлення кроку реалізації LSTM нейронної мережі при фіксованому часі t представлено на Рис. 4.4.

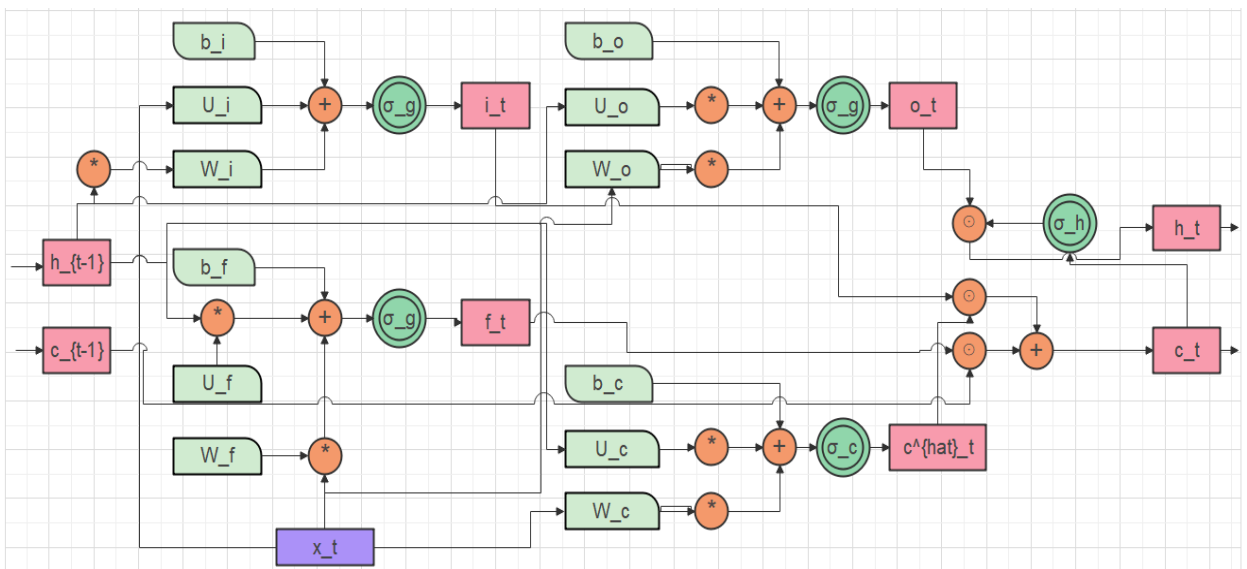


Рис 4.4. Діаграма LSTM нейронної мережі

У роботах [36, 92] розглянуто розширення LSTM нейронної мережі на випадок залежності вихідного сигналу від деяких зовнішніх чинників. В цьому випадку залежна змінна (вихідний сигнал) залежить від рівня деякої факторної змінної. Наприклад, у роботі [112], автори розглядають декілька дескрипторів, які впливають на залученість дітей під час перегляду відео із супроводом текстової інформації. А саме, цими дескрипторами є положення голови дитини, гістограму орієнтованих градієнтів і гістограму орієнтованого потоку навколо верхньої частини тіла дитини. В цьому випадку, топологія розширеної нейронної моделі (Multi View LSTM, MV-LSTM) формально описується наступним чином

$$\left\{ \begin{array}{l} f_t^{(v)} = \sigma_g \left(W_f^{(v)} x_t + U_f A h_{t-1}^{(v)} + \sum_{k \neq v} U_f^{(k)} B h_{t-1}^{(k)} + b_f \right), \\ i_t = \sigma_g \left(W_i^{(v)} x_t + U_i A h_{t-1}^{(v)} + \sum_{k \neq v} U_i^{(k)} B h_{t-1}^{(k)} + b_i \right), \\ o_t = \sigma_g \left(W_o^{(v)} x_t + U_o A h_{t-1}^{(v)} + \sum_{k \neq v} U_o^{(k)} B h_{t-1}^{(k)} + b_o \right), \\ \tilde{c}_t = \sigma_c \left(W_c^{(v)} x_t + U_c A h_{t-1}^{(v)} + \sum_{k \neq v} U_c^{(k)} B h_{t-1}^{(k)} + b_c \right), \\ c_t = f_t \odot c_{t-1} + i_t \odot \tilde{c}_t, \\ h_t = o_t \odot \sigma_h(c_t), \end{array} \right. \quad (4.14)$$

де $v \in \{1, \dots, V\}$ – рівень факторної змінної, матриці A та B визначаються із наступних співвідношень

$$A_{ij} = \begin{cases} 1, & i = j \leq \alpha d, \\ 0, & \text{в іншому випадку} \end{cases}$$

$$B_{ij} = \begin{cases} 1, & i = j \geq (1 - \beta)d, \\ 0, & \text{в іншому випадку} \end{cases}$$

$\alpha, \beta \in [0, 1]$ – гіперпараметри нейронної мережі MV-LSTM, які вказують на тісноту зв'язку між різними рівнями вихідного сигналу $v \in \{1, \dots, V\}$. Слід зауважити, що модель (4.14) можна інтерпретувати як панельну модель, де

одночасно відслідковуються декілька реалізацій одного процесу або явища із можливо різними значеннями вихідного сигналу.

4.2.2. Узагальнена модель

Розглянемо переваги недоліки LSTM чи MV-LSTM нейронних мереж, які являються класичним прикладом нейронних мереж для однопотокових та багатопотокових процесів. Перевагами даних топологій насамперед являється простота задання та врахування скінченної передісторії у вигляді одного лагу. Основним недоліком моделей (4.13) та (4.14) являється використання лише першого лагу в моделі LSTM чи MV-LSTM. При цьому виникає проблема у аналізі часових рядів із чітко визначеною періодичністю. Наприклад, аналіз часового ряду

$$x_t = \phi x_{t-k} + \sum_{i=1}^{k-1} \phi_i x_{t-i} + \varepsilon_t,$$

де $|\phi| < 1$, $\sum_{i=1}^{k-1} |\phi_i| \ll |\phi|$, $\varepsilon_t \sim i.i.d. N(0, \sigma^2)$ вказує на низьку точність прогнозу на основі LSTM нейронної мережі. В задачах побудови прогнозу для часових рядів в класичних моделях припускається, що має стан процесу у теперішній час залежить від стану процесу в попередні моменти часу, тобто

$$x_t = f(x_{t-1}, x_{t-2}, \dots).$$

Для подальшого аналізу будемо використовувати топологію розширеної LSTM нейронної мережі (extended LSTM, E-LSTM) [153], яка буде визначатися наступними співвідношеннями

$$\left\{ \begin{array}{l} f_t = \sigma_g(W_f x_t + U_f h_{t-1} + b_f), \\ i_t = \sigma_g(W_i x_t + U_i h_{t-1} + b_i), \\ o_t = \sigma_g(W_o x_t + U_o h_{t-1} + b_o), \\ \tilde{c}_t = \sigma_c(W_c x_t + U_c h_{t-1} + b_c), \\ c_t = \sum_{k=0}^p f_{t-k} \odot c_{t-1-k} + i_t \odot \tilde{c}_t, \\ h_t = o_t \odot \sigma_h(c_t), \end{array} \right. \quad (4.15)$$

Як ми можемо бачити із опису, різниця між LSTM та E-LSTM (Рис. 4.5) полягає у використанні AR моделі для випадкового процесу c_t , де в якості коефіцієнтів даної AR моделі виступають наближення $f_t, t \geq 0$.

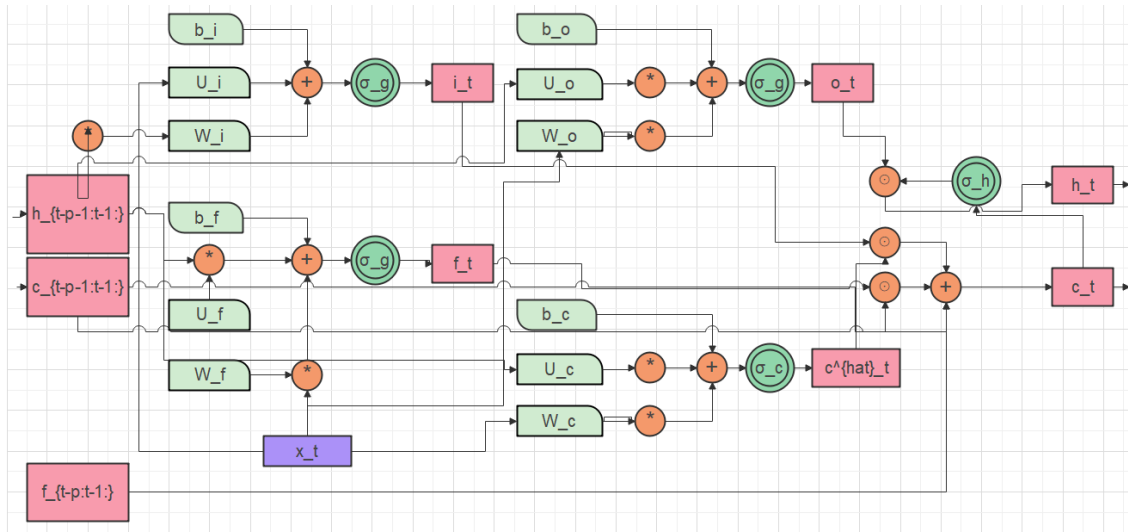


Рис 4.5. Діаграма узагальненої LSTM (E-LSTM) нейронної мережі

4.2.3. Оцінка коефіцієнтів

В даному розділі розглянемо метод зворотного поширення (backward propagation) для моделі (4.15), який використовується для оцінки параметрів моделі

$$Par = (W_f, U_f, b_f, W_i, U_i, b_i, W_o, U_o, b_o, W_c, U_c, b_c)$$

із застосування ітераційних методів. Для оцінки параметрів моделі (4.15), визначимо функціонал якості

$$Q(Par) = Q(Par; Y),$$

де

$$Y = (y_1, y_2, \dots, y_T)$$

визначає значення вихідного сигналу, причому на основі означення моделі (4.13), $y_t \in (-1, 1)^h$. Відзначимо, що на основі параметрів Par однозначно визначаються значення оцінок

$$\hat{Y} = (\hat{y}_1, \hat{y}_2, \dots, \hat{y}_T) = (h_1, \dots, h_T).$$

Функція Q являється цільовою функцією, яка відображає якість оцінки моделі. Зазвичай у якості функції Q вибирається функція правдоподібності, функція втрат або сума квадратів залишків.

Розглянемо спочатку частинні похідні \hat{Y} за параметрами Par . Оскільки схема (4.15) пропонує покрокове обчислення h_t , то обчислення частинних похідних будемо проводити з кінця до початку (backpropagation procedure):

$$\begin{aligned}
\frac{\partial \hat{y}_t}{\partial b_c} &= \frac{\partial h_t}{\partial b_c} = \frac{\partial}{\partial b_c} \left(o_t \odot \sigma_h \left(\sum_{k=0}^p f_{t-k} \odot c_{t-1-k} + i_t \odot \tilde{c}_t \right) \right) = \\
&= \frac{\partial}{\partial b_c} \sigma_g(W_o x_t + U_o h_{t-1} + b_o) \odot \sigma_h \left(\sum_{k=0}^p f_{t-k} \odot c_{t-1-k} + i_t \odot \tilde{c}_t \right) + \\
&+ \sigma_g(W_o x_t + U_o h_{t-1} + b_o) \odot \frac{\partial}{\partial b_c} \sigma_h \left(\sum_{k=0}^p f_{t-k} \odot c_{t-1-k} + i_t \odot \tilde{c}_t \right) = \\
&= \sigma'_g(W_o x_t + U_o h_{t-1} + b_o) U_o \frac{\partial}{\partial b_c} h_{t-1} \odot \sigma_h \left(\sum_{k=0}^p f_{t-k} \odot c_{t-1-k} + i_t \odot \tilde{c}_t \right) + \\
&\quad \sigma_g(W_o x_t + U_o h_{t-1} + b_o) \odot \sigma'_h \left(\sum_{k=0}^p f_{t-k} \odot c_{t-1-k} + i_t \odot \tilde{c}_t \right) * \\
&\quad * \frac{\partial}{\partial b_c} \left(\sum_{k=0}^p f_{t-k} \odot c_{t-1-k} + i_t \odot \tilde{c}_t \right) = \\
&= \sigma'_g(o_t) U_o \frac{\partial}{\partial b_c} h_{t-1} \odot \sigma_h(c_t) + \\
&\quad \sigma_g(o_t) \odot \sigma'_h(c_t) \frac{\partial}{\partial b_c} (c_t),
\end{aligned}$$

де

$$\frac{\partial}{\partial b_c} (c_t) = \frac{\partial}{\partial b_c} \left(\sum_{k=0}^p f_{t-k} \odot c_{t-1-k} + i_t \odot \tilde{c}_t \right) =$$

$$\begin{aligned}
&= \sum_{k=0}^p \frac{\partial}{\partial b_c} (f_{t-k}) \odot c_{t-1-k} + \sum_{k=0}^p f_{t-k} \odot \frac{\partial}{\partial b_c} (c_{t-1-k}) + \\
&\quad + \frac{\partial}{\partial b_c} (i_t) \odot \tilde{c}_t + i_t \odot \frac{\partial}{\partial b_c} (\tilde{c}_t) = \\
&= \sum_{k=0}^p \sigma'_g(W_f x_t + U_f h_{t-1} + b_f) U_f \frac{\partial}{\partial b_c} (h_{t-1}) \odot c_{t-1-k} + \\
&+ \sum_{k=0}^p f_{t-k} \odot \frac{\partial}{\partial b_c} (c_{t-1-k}) + \sigma'_g(W_i x_t + U_i h_{t-1} + b_i) U_i \frac{\partial}{\partial b_c} (h_{t-1}) \odot \tilde{c}_t + \\
&\quad + i_t \odot \sigma'_c(W_c x_t + U_c h_{t-1} + b_c) U_c \frac{\partial}{\partial b_c} (h_{t-1}).
\end{aligned}$$

Зауважимо, що для $t \leq 0$ всі наближення будуть рівні нульовим векторам відповідних розмірам, тобто

$$\frac{\partial}{\partial b_c} (c_t), \frac{\partial}{\partial b_c} (h_t)$$

є нульовими векторами. Аналогічно можемо визначити і похідні за змінними (W_c, U_c)

$$\frac{\partial \hat{y}_t}{\partial U_c} = \sigma'_g(o_t) U_o \frac{\partial}{\partial U_c} h_{t-1} \odot \sigma_h(c_t) + \sigma_g(o_t) \odot \sigma'_h(c_t) \frac{\partial}{\partial U_c} (c_t),$$

де

$$\begin{aligned}
\frac{\partial}{\partial U_c} (c_t) &= \sum_{k=0}^p \sigma'_g(W_f x_t + U_f h_{t-1} + b_f) U_f \frac{\partial}{\partial U_c} (h_{t-1}) \odot c_{t-1-k} + \\
&+ \sum_{k=0}^p f_{t-k} \odot \frac{\partial}{\partial U_c} (c_{t-1-k}) + \sigma'_g(W_i x_t + U_i h_{t-1} + b_i) U_i \frac{\partial}{\partial U_c} (h_{t-1}) \odot \tilde{c}_t + \\
&\quad + i_t \odot \sigma'_c(W_c x_t + U_c h_{t-1} + b_c) U_c \frac{\partial}{\partial U_c} (h_{t-1}) + \\
&\quad i_t \odot \sigma'_c(W_c x_t + U_c h_{t-1} + b_c) h_{t-1}
\end{aligned}$$

та

$$\frac{\partial \hat{y}_t}{\partial W_c} = \sigma'_g(o_t) U_o \frac{\partial}{\partial W_c} h_{t-1} \odot \sigma_h(c_t) + \sigma_g(o_t) \odot \sigma'_h(c_t) \frac{\partial}{\partial W_c} (c_t),$$

де

$$\begin{aligned} \frac{\partial}{\partial W_c}(c_t) &= \sum_{k=0}^p \sigma'_g(W_f x_t + U_f h_{t-1} + b_f) U_f \frac{\partial}{\partial W_c}(h_{t-1}) \odot c_{t-1-k} + \\ &+ \sum_{k=0}^p f_{t-k} \odot \frac{\partial}{\partial W_c}(c_{t-1-k}) + \sigma'_g(W_i x_t + U_i h_{t-1} + b_i) U_i \frac{\partial}{\partial W_c}(h_{t-1}) \odot \tilde{c}_t + \\ &+ i_t \odot \sigma'_c(W_c x_t + U_c h_{t-1} + b_c) x_t. \end{aligned}$$

Розглянемо тепер також похідні за іншими змінними

$$\begin{aligned} \frac{\partial \hat{y}_t}{\partial b_o} &= \frac{\partial}{\partial b_o} \left(o_t \odot \sigma_h \left(\sum_{k=0}^p f_{t-k} \odot c_{t-1-k} + i_t \odot \tilde{c}_t \right) \right) = \\ &= \frac{\partial}{\partial b_o}(o_t) \odot \sigma_h(c_t) + o_t \odot \frac{\partial}{\partial b_o}(\sigma_h(c_t)) = \\ &= \sigma'_g(W_o x_t + U_o h_{t-1} + b_o) \odot \sigma_h(c_t) + \\ &+ o_t \odot \sigma'_h(c_t) \sum_{k=0}^p \left(\begin{array}{l} \frac{\partial}{\partial b_o}(f_{t-k} \odot) c_{t-1-k} + f_{t-k} \odot \frac{\partial}{\partial b_o}(c_{t-1-k}) + \\ \frac{\partial}{\partial b_o}(i_t) \odot \tilde{c}_t + i_t \odot \frac{\partial}{\partial b_o}(\tilde{c}_t) \end{array} \right), \end{aligned}$$

де $\frac{\partial}{\partial b_o}(f_{t-k} \odot), \frac{\partial}{\partial b_o}(c_{t-1-k})$ задаються рекурентно, аналогічно формулам вище.

$$\begin{aligned} \frac{\partial \hat{y}_t}{\partial U_o} &= \sigma'_g(W_o x_t + U_o h_{t-1} + b_o) \left(h_{t-1} + U_o \frac{\partial h_{t-1}}{\partial U_o} \right) \odot \sigma_h(c_t) + \\ &+ o_t \odot \sigma'_h(c_t) \sum_{k=0}^p \left(\begin{array}{l} \frac{\partial}{\partial U_o}(f_{t-k}) \odot c_{t-1-k} + f_{t-k} \odot \frac{\partial}{\partial U_o}(c_{t-1-k}) + \\ \frac{\partial}{\partial U_o}(i_t) \odot \tilde{c}_t + i_t \odot \frac{\partial}{\partial U_o}(\tilde{c}_t) \end{array} \right); \\ \frac{\partial \hat{y}_t}{\partial W_o} &= \sigma'_g(W_o x_t + U_o h_{t-1} + b_o) x_t \odot \sigma_h(c_t) + \\ &+ o_t \odot \sigma'_h(c_t) \sum_{k=0}^p \left(\begin{array}{l} \frac{\partial}{\partial W_o}(f_{t-k}) \odot c_{t-1-k} + f_{t-k} \odot \frac{\partial}{\partial W_o}(c_{t-1-k}) + \\ \frac{\partial}{\partial W_o}(i_t) \odot \tilde{c}_t + i_t \odot \frac{\partial}{\partial W_o}(\tilde{c}_t) \end{array} \right); \end{aligned}$$

$$\begin{aligned}
\frac{\partial \hat{y}_t}{\partial b_i} &= \frac{\partial}{\partial b_i} \left(o_t \odot \sigma_h \left(\sum_{k=0}^p f_{t-k} \odot c_{t-1-k} + i_t \odot \tilde{c}_t \right) \right) = \\
&\quad \sigma'_g(W_o x_t + U_o h_{t-1} + b_o) U_o \frac{\partial}{\partial b_i} (h_{t-1}) \odot \sigma_h(c_t) + \\
&\quad o_t \odot \sigma'_h(c_t) \sum_{k=0}^p \left(\begin{aligned} &\frac{\partial}{\partial b_i} (f_{t-k}) \odot c_{t-1-k} + f_{t-k} \odot \frac{\partial}{\partial b_i} (c_{t-1-k}) + \\ &+ \sigma'_g(W_i x_t + U_i h_{t-1} + b_i) \odot \tilde{c}_t + \\ &+ i_t \odot \sigma'_c(W_c x_t + U_c h_{t-1} + b_c) U_c \frac{\partial}{\partial b_i} (h_{t-1}) \end{aligned} \right); \\
\frac{\partial \hat{y}_t}{\partial U_i} &= \frac{\partial}{\partial U_i} \left(o_t \odot \sigma_h \left(\sum_{k=0}^p f_{t-k} \odot c_{t-1-k} + i_t \odot \tilde{c}_t \right) \right) = \\
&\quad \sigma'_g(W_o x_t + U_o h_{t-1} + b_o) U_o \frac{\partial}{\partial U_i} (h_{t-1}) \odot \sigma_h(c_t) + \\
&\quad o_t \odot \sigma'_h(c_t) \sum_{k=0}^p \left(\begin{aligned} &\frac{\partial}{\partial U_i} (f_{t-k}) \odot c_{t-1-k} + f_{t-k} \odot \frac{\partial}{\partial U_i} (c_{t-1-k}) + \\ &+ \sigma'_g(W_i x_t + U_i h_{t-1} + b_i) \left(h_{t-1} + U_i \frac{\partial}{\partial U_i} (h_{t-1}) \right) \odot \tilde{c}_t + \\ &+ i_t \odot \sigma'_c(W_c x_t + U_c h_{t-1} + b_c) U_c \frac{\partial}{\partial U_i} (h_{t-1}) \end{aligned} \right); \\
\frac{\partial \hat{y}_t}{\partial W_i} &= \frac{\partial}{\partial W_i} \left(o_t \odot \sigma_h \left(\sum_{k=0}^p f_{t-k} \odot c_{t-1-k} + i_t \odot \tilde{c}_t \right) \right) = \\
&\quad \sigma'_g(W_o x_t + U_o h_{t-1} + b_o) U_o \frac{\partial}{\partial W_i} (h_{t-1}) \odot \sigma_h(c_t) + \\
&\quad o_t \odot \sigma'_h(c_t) \sum_{k=0}^p \left(\begin{aligned} &\frac{\partial}{\partial W_i} (f_{t-k}) \odot c_{t-1-k} + f_{t-k} \odot \frac{\partial}{\partial W_i} (c_{t-1-k}) + \\ &+ \sigma'_g(W_i x_t + U_i h_{t-1} + b_i) \left(x_t + U_i \frac{\partial}{\partial W_i} (h_{t-1}) \right) \odot \tilde{c}_t + \\ &+ i_t \odot \sigma'_c(W_c x_t + U_c h_{t-1} + b_c) U_c \frac{\partial}{\partial W_i} (h_{t-1}) \end{aligned} \right);
\end{aligned}$$

$$\begin{aligned}
\frac{\partial \hat{y}_t}{\partial b_f} &= \frac{\partial}{\partial b_f} \left(o_t \odot \sigma_h \left(\sum_{k=0}^p f_{t-k} \odot c_{t-1-k} + i_t \odot \tilde{c}_t \right) \right) = \\
&\quad \sigma'_g(W_o x_t + U_o h_{t-1} + b_o) U_o \frac{\partial}{\partial b_f} (h_{t-1}) \odot \sigma_h(c_t) + \\
&\quad o_t \odot \sigma'_h(c_t) \sum_{k=0}^p \left(\begin{aligned} &\sigma'_g(W_f x_t + U_f h_{t-1} + b_f) c_{t-1-k} + \\ &+ f_{t-k} \odot \frac{\partial}{\partial b_f} (c_{t-1-k}) + \\ &+ \sigma'_g(W_i x_t + U_i h_{t-1} + b_i) U_i \frac{\partial}{\partial b_f} (h_{t-1}) \odot \tilde{c}_t + \\ &+ i_t \odot \sigma'_c(W_c x_t + U_c h_{t-1} + b_c) U_c \frac{\partial}{\partial b_f} (h_{t-1}) \end{aligned} \right); \\
\frac{\partial \hat{y}_t}{\partial U_f} &= \frac{\partial}{\partial U_f} \left(o_t \odot \sigma_h \left(\sum_{k=0}^p f_{t-k} \odot c_{t-1-k} + i_t \odot \tilde{c}_t \right) \right) = \\
&\quad \sigma'_g(W_o x_t + U_o h_{t-1} + b_o) U_o \frac{\partial}{\partial U_f} (h_{t-1}) \odot \sigma_h(c_t) + \\
&\quad o_t \odot \sigma'_h(c_t) \sum_{k=0}^p \left(\begin{aligned} &\sigma'_g(W_f x_t + U_f h_{t-1} + b_f) \left(h_{t-1} + U_f \frac{\partial}{\partial U_f} (h_{t-1}) \right) c_{t-1-k} + \\ &+ f_{t-k} \odot \frac{\partial}{\partial U_f} (c_{t-1-k}) + \\ &+ \sigma'_g(W_i x_t + U_i h_{t-1} + b_i) U_i \frac{\partial}{\partial U_f} (h_{t-1}) \odot \tilde{c}_t + \\ &+ i_t \odot \sigma'_c(W_c x_t + U_c h_{t-1} + b_c) U_c \frac{\partial}{\partial U_f} (h_{t-1}) \end{aligned} \right); \\
\frac{\partial \hat{y}_t}{\partial W_f} &= \frac{\partial}{\partial W_f} \left(o_t \odot \sigma_h \left(\sum_{k=0}^p f_{t-k} \odot c_{t-1-k} + i_t \odot \tilde{c}_t \right) \right) = \\
&\quad \sigma'_g(W_o x_t + U_o h_{t-1} + b_o) U_o \frac{\partial}{\partial W_f} (h_{t-1}) \odot \sigma_h(c_t) +
\end{aligned}$$

$$o_t \odot \sigma'_h(c_t) \sum_{k=0}^p \left(\begin{aligned} & \sigma'_g(W_f x_t + U_f h_{t-1} + b_f) \left(x_t + U_f \frac{\partial}{\partial W_f} (h_{t-1}) \right) c_{t-1-k} + \\ & + f_{t-k} \odot \frac{\partial}{\partial W_f} (c_{t-1-k}) + \\ & + \sigma'_g(W_i x_t + U_i h_{t-1} + b_i) U_i \frac{\partial}{\partial W_f} (h_{t-1}) \odot \tilde{c}_t + \\ & + i_t \odot \sigma'_c(W_c x_t + U_c h_{t-1} + b_c) U_c \frac{\partial}{\partial W_f} (h_{t-1}) \end{aligned} \right).$$

На основі рекурентних формул, визначених для частинних похідних, аналогічно до класичного результату, неважко побудувати оновлення параметрів нейронної мережі, яке будується на градієнтно - орієнтованих методах. Іншим підходом до оцінки параметрів моделі є використання генетичних алгоритмів, в яких оновлення не потребує наявності градієнту. Недоліком використання генетичних алгоритмів являється висока ресурсозатратність зумовлена потребою великої кількості обчислень цільової функції Q для різних значень параметру Par .

Додатковою проблемою, яка виникає при дослідженні розширеної моделі (4.15) є оцінка невідомого параметру p , що вказує на довжину передісторії, яка повинна бути використана при аналізі. Для оцінки даного параметру у розширеній моделі буде використано оцінку на основі $AR(p)$ моделі [8], оскільки вихідний сигнал являється функцією від того ж сигналу в попередні моменти часу, тобто має місце наступна залежність

$$g(x_t) = f(x_{t-1}, x_{t-2}, \dots, x_{t-p}),$$

де g – відома функція (перетворення), яка становить інтерес при дослідженні. Наприклад, в деяких прикладних задачах функція $g(x)$ визначається співвідношенням

$$g(x) = \begin{cases} 1, & x > 0, \\ 0, & x = 0, \\ -1, & x < 0 \end{cases}$$

або

$$g(x) = \begin{cases} 1, & x > h, \\ 0, & x \leq h. \end{cases}$$

Таким чином, підсумовуючи всі вищевказані кроки можна узагальнити на наступні кроки:

Крок 1. Оцінка гіперпараметра p на основі AR моделі із відомою функцією g із мінімізацією залишків або інформаційного критерію.

Крок 2. Оцінка параметрів моделі (4.15) на основі градієнтно-орієнтованих методів або на основі генетичного алгоритму.

Крок 3. Побудова прогнозу для

$$h_t = g(x_t)$$

на основі цільової функції Q .

Висновки до розділу 4

Розділ 4 присвячено дослідженню двох гібридних моделей, які ґрунтуються на моделях часових рядів. Перша модель являє собою неперервний та нелінійний аналог класичної лінійної ARMA моделі, друга модель являє собою узагальнення класичної LSTM нейронної мережі із врахуванням довільної скінченної передісторії.

Перша модель базується на новому методі оцінки прогнозу процесів фінансових ринків в основі якого лежать моделі стохастичних диференціальних рівнянь із запізненням. У теорії часових рядів на основі даного підходу будуються нелінійні моделі та неперервні моделі, що дає змогу будувати більш точний прогноз для даних із замірюваннями через нерівномірні інтервали часу. Для основних класичних моделей часових рядів розроблені алгоритми оцінки параметрів, проте для нелінійних моделей розробка аналогічних алгоритмів ускладнюється формою нелінійності. Тому для оцінки параметрів моделі будемо використовувати генетичний алгоритм, а саме метод рою частинок, оскільки цей метод виявився найбільш

оптимальним в нашій задачі. Зауважимо, що методи послідовних наближень (див. наприклад [22, 23]), не можуть бути використані в даній задачі.

Також вказано один із підходів переходу від оптимізаційної задачі по класу функцій до параметричної оптимізаційної задачі, що дозволяє значно ефективніше використовувати генетичні алгоритми пошуку оптимальних параметрів. Крім того, на основі Теорема 1 можна проводити оцінку параметрів функціоналів L та G за наступним алгоритмом: спочатку оцінити функціонал L , використовуючи детерміноване рівняння з $g = 0$, після цього оцінювати функціонал G з використанням дисперсії реального процесу.

Отримані теоретичні результати протестовано на реальних даних із використанням динаміки трьох реальних фінансових процесів. Результати тестування показали, що нова модель краще апроксимує реальні фінансові процеси, оскільки враховує як нелінійності моделей, так і нерегулярність між замірами реальних процесів. Основним із недоліків моделі полягає у переході від дискретної до неперервної моделі часто залежить на додаткових припущеннях щодо властивостей теоретичного наближення. Іншим недолік моделі полягає у відсутності ітераційного процесу оцінки параметрів. Тому у роботі використано генетичні алгоритм для оцінки параметрів, що не може бути самим ефективним методом оцінки.

Друга модель даної роботи присвячено одному узагальненню LSTM нейронних мереж із врахуванням довжини передісторії нейронної мережі. Основна відмінність від класичної LSTM моделі або багатовимірного варіанту MV-LSTM моделі полягає у наявності гіперпараметру p , який відображає глибину залежності вхідного сигналу (x_t, \dots, x_{t+1-p}) та вихідного сигналу $g(x_{t+1})$ або в більш загальному випадку $g(x_{t+h})$ для $h > 0$. Як було вказано у цьому розділі, часові ряди із явною періодичною залежністю погано оцінюються за допомогою класичної LSTM нейронної мережі, проте дана проблема вирішується на основі узагальненої моделі (4.15).

Для оцінки параметрів нейронних мереж найбільш вживаними являються градієнтно-орієнтовані методи, наприклад метод зворотного

поширення. Для реалізації даного методу необхідним являється обчислення похідних па параметрах нейронної мережі. В підрозділі 4.2.3 здійснено теоретичні викладки щодо обчислення градієнту параметрів нейронної мережі.

ВИСНОВКИ

Як було зазначено в працях багатьох авторів, поєднання різних методим оцінки параметрів найчастіше дає більш кращий результат, ніж використання кожної із методик окремо. Це нашоує на розвиток методів машинного навчання та штучного інтелекту у напрямку побудови гібридних алгоритмів, за допомогою яких основна задача побудови прогнозу “розбивається” на частинки, одна із яких буде відповідати за аналіз основного показника (вартості випадкового активу S_t або його ренти R_t). Одними з найпростіших прототипів гібридних алгоритмів є ARCH моделі часових рядів та їх численні модифікації, в яких основний процес описується за допомогою моделі часового ряду без участі екзогенних змінних, проте дисперсія часового ряду описується за допомогою моделей часових рядів із врахуванням екзогенних змінних включаючи і основну досліджувану змінну.

Даний підхід можна розширювати з використанням гібридних моделей, де в якості основної моделі залишається модель часового ряду, проте опис залишків або врахування специфічних властивостей часового ряду здійснюється за допомогою більш складних моделей. В даній роботі в якості допоміжних моделей вибрано використання нейронних мереж та динамічних хаотичних систем, причому вибір останнього ґрунтується на швидкому розвитку теорії детермінованого хаосу та її використанні в описі часових рядів.

Отже, основними здобутками роботи, представленими в трьох основних розділах 2 – 4, є наступне:

1. Доведено, що часові ряди фондових індексів містять як лінійну, так і не лінійну складову, а отже, окремо лінійна модель ARIMA та нелінійна модель ANN не можуть дати точну оцінку таких часових рядів.
2. Розроблено гібридну модель моделі авторегресії-ковзної середньої та штучної нейронної мережі прогнозування часових рядів фондових

індексів. Побудована модель показала більшу точність прогнозів, отриманих за допомогою алгоритму гібридної інтегрованої моделі авторегресії-ковзної середньої та штучної нейронної мережі, для якого значення відповідних оцінок похибок нижчі.

3. Доведено, якщо часовий ряд є строго стаціонарним, то ARIMA модель є більш адекватною для гауссівських рядів. В загальному випадку, з метою перевірки гіпотези про нормальність розподілу використовується тест Жарка-Бера, який, зокрема, передбачає перевірку значення ексцесу. Даний теоретичний результат протестовано на реальних даних.
4. Доведено практичність використання теорії хаотичних динамічних систем із використанням різних показників хаотичності для аналізу як реальних часових рядів, так і класичних теоретичних моделей. Для визначення існування хаосу в системі (часовому ряді) були представлені чотири методи: перетворення Фур'є, показник Ляпунова, показник Герста і фрактальна розмірність.
5. Використовуючи підхід оцінки наявності хаосу в динамічній системі, розроблено ряд гібридних моделей на основі класичної моделі нейронної мережі прямого поширення та нейронної мережі Елмана, які дозволили більш точно оцінювати прогноз для моделі Маккея-Глесса, логістичного рівняння та моделі Ено.
6. Проілюстровано, що використання гібридних алгоритмів для підбору стартової точки оцінки параметрів моделі, може значно зменшити як обчислювальну складність моделі, так і підвищити точність прогнозу.
7. Розроблено загальний підхід побудови гібридних моделей на основі узагальненої моделі LSTM із врахуванням періодичних складових
8. У четвертому розділі розглянуто модель, яка базується на новому методі оцінки прогнозу процесів фінансових ринків в основі якого лежать моделі стохастичних диференціальних рівнянь із запізненням. У теорії часових рядів на основі даного підходу будуються нелінійні

моделі та неперервні моделі, що дає змогу будувати більш точний прогноз для даних із замірюваннями через нерівномірні інтервали часу. Для основних класичних моделей часових рядів розроблені алгоритми оцінки параметрів, проте для нелінійних моделей розробка аналогічних алгоритмів ускладнюється формою нелінійності. Тому для оцінки параметрів моделі будемо використовувати генетичний алгоритм, а саме метод рою частинок, оскільки цей метод виявився найбільш оптимальним в нашій задачі. Зауважимо, що методи послідовних наближень (див. наприклад [22, 23]), не можуть бути використані в даній задачі.

9. Вказано один із підходів переходу від оптимізаційної задачі по класу функцій до параметричної оптимізаційної задачі, що дозволяє значно ефективніше використовувати генетичні алгоритми пошуку оптимальних параметрів. Крім того, на основі Теорема 4.1 можна проводити оцінку параметрів функціоналів L та G за наступним алгоритмом: спочатку оцінити функціонал L , використовуючи детерміноване рівняння з $g = 0$, після цього оцінювати функціонал G з використанням дисперсії реального процесу.
10. Отримані теоретичні результати протестовано на реальних даних із використанням динаміки трьох реальних фінансових процесів. Результати тестування показали, що нова модель краще апроксимує реальні фінансові процеси, оскільки враховує як нелінійності моделей, так і нерегулярність між замірами реальних процесів. Основним із недоліків моделі полягає у переході від дискретної до неперервної моделі часто залежить на додаткових припущеннях щодо властивостей теоретичного наближення. Іншим недолік моделі полягає у відсутності ітераційного процесу оцінки параметрів. Тому у роботі використано генетичні алгоритм для оцінки параметрів, що не може бути самим ефективним методом оцінки.

11. Друга модель розділу 4 присвячено одному узагальненню LSTM нейронних мереж із врахуванням довжини передісторії нейронної мережі. Основна відмінність від класичної LSTM моделі або багатовимірного варіанту MV-LSTM моделі полягає у наявності гіперпараметру p , який відображає глибину залежності вхідного сигналу (x_t, \dots, x_{t+1-p}) та вихідного сигналу $g(x_{t+1})$ або в більш загальному випадку $g(x_{t+h})$ для $h > 0$. Як було вказано у цьому розділі, часові ряди із явною періодичною залежністю погано оцінюються за допомогою класичної LSTM нейронної мережі, проте дана проблема вирішується на основі узагальненої моделі.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

- [1] A Deep Learning-Based Phishing Detection System Using CNN, LSTM, and LSTM-CNN / Z. Alshingiti et al. *Electronics*. 2023. Vol. 12, Iss. 1. P. 232. URL: <https://doi.org/10.3390/electronics12010232>.
- [2] A two-stage architecture for stock price forecasting by integrating self-organizing map and support vector regression / S.-H. Hsu et al. *Expert Systems with Applications*. 2009. Vol. 36, Iss. 4. P. 7947–7951. URL: <https://doi.org/10.1016/j.eswa.2008.10.065>.
- [3] Afolabi M., Olude O. Predicting Stock Prices Using a Hybrid Kohonen Self Organizing Map (SOM). *2007 40th Annual Hawaii International Conference on System Sciences (HICSS'07)*, Waikoloa, HI, USA, 3–6 January 2007. 2007. URL: <https://doi.org/10.1109/hicss.2007.441>.
- [4] Aghabozorgi S., Teh Y. W. Stock market co-movement assessment using a three-phase clustering method. *Expert Systems with Applications*. 2014. Vol. 41, Iss. 4. P. 1301–1314. URL: <https://doi.org/10.1016/j.eswa.2013.08.028>.
- [5] Aguilar-Rivera R., Valenzuela-Rendón M., Rodríguez-Ortiz J. J. Genetic algorithms and Darwinian approaches in financial applications: A survey. *Expert Systems with Applications*. 2015. Vol. 42, Iss. 21. P. 7684–7697. URL: <https://doi.org/10.1016/j.eswa.2015.06.001>.
- [6] Andrea Leccese. Machine Learning in Finance: Why You Should Not Use LSTM's to Predict the Stock Market. веб-сайт. URL: <https://towardsdatascience.com/machine-learning-in-finance-why-you-should-not-use-lstms-to-predict-the-stock-market-3f7239b9832c>.
- [7] Application of high-dimensional feature selection: evaluation for genomic prediction in man / M. L. Bermingham et al. *Scientific Reports*. 2015. Vol. 5, Iss. 1. URL: <https://doi.org/10.1038/srep10312>.
- [8] Ari Y. Continuous Autoregressive Moving Average Models. *Methodologies and Applications of Computational Statistics for*

- Machine Intelligence*. 2021. P. 118–141.
URL: <https://doi.org/10.4018/978-1-7998-7701-1.ch007>.
- [9] Armillotta M., Fokianos K., Krikidis I. Bootstrapping Network Autoregressive Models for Testing Linearity. *Studies in Computational Intelligence*. Cham, 2023. P. 99–116. URL: https://doi.org/10.1007/978-3-031-24453-7_6.
- [10] Atsalakis G. S., Valavanis K. P. Surveying stock market forecasting techniques – Part II: Soft computing methods. *Expert Systems with Applications*. 2009. Vol. 36, Iss. 3. P. 5932–5941. URL: <https://doi.org/10.1016/j.eswa.2008.07.006>.
- [11] Barro D., Consigli G., Varun V. A stochastic programming model for dynamic portfolio management with financial derivatives. *Journal of Banking & Finance*. 2022. P. 106445. URL: <https://doi.org/10.1016/j.jbankfin.2022.106445>.
- [12] Billah M., Waheed S., Hanifa A. Predicting Closing Stock Price using Artificial Neural Network and Adaptive Neuro Fuzzy Inference System (ANFIS: The Case of the Dhaka Stock Exchange. *International Journal of Computer Applications*. 2015. Vol. 129, Iss. 11. P. 1–5. URL: <https://doi.org/10.5120/ijca2015906952>.
- [13] Blechschmidt J. Numerical Methods for Stochastic Control Problems with Applications in Financial Mathematics. *Chemnitz*. 2022. P. 223.
- [14] Bowden G. J., Maier H. R., Dandy G. C. Optimal division of data for neural network models in water resources applications. *Water Resources Research*. 2002. Vol. 38, Iss. 2. P. 2–1–2–11. URL: <https://doi.org/10.1029/2001wr000266> .
- [15] Box G. E. P. Time series analysis: Forecasting and control. San Francisco : Holden-Day, 1970. 553 p.
- [16] Boyacioglu M. A., Avci D. An Adaptive Network-Based Fuzzy Inference System (ANFIS) for the prediction of stock market return: The case of the

- Istanbul Stock Exchange. *Expert Systems with Applications*. 2010. Vol. 37, Iss. 12. P. 7908–7912. URL: <https://doi.org/10.1016/j.eswa.2010.04.045>.
- [17] Brockwell P.J., Davis R.A. Introduction to Time Series and Forecasting: Springer (3rd edition). 2016. 439 p.
- [18] Bryant P. H. Extensional singularity dimensions for strange attractors. *Physics Letters A*. 1993. Vol. 179, Iss. 3. P. 186–190. URL: [https://doi.org/10.1016/0375-9601\(93\)91136-s](https://doi.org/10.1016/0375-9601(93)91136-s).
- [19] Bryant P., Brown R., Abarbanel H. D. I. Lyapunov exponents from observed time series. *Physical Review Letters*. 1990. Vol. 65, Iss. 13. P. 1523–1526. URL: <https://doi.org/10.1103/physrevlett.65.1523> .
- [20] Campbell J. Y. The econometrics of financial markets. Princeton, N.J : Princeton University Press, 1997. 611 p.
- [21] Cencini M. Chaos: From simple models to complex systems. Hackensack, NJ : World Scientific, 2010. 460 p.
- [22] Chande T. S. Beyond Technical Analysis: How to Develop and Implement a Winning Trading System, 2nd Edition. Wiley, 2001. 336 p.
- [23] Chen W.-S., Du Y.-K. Using neural networks and data mining techniques for the financial distress prediction model. *Expert Systems with Applications*. 2009. Vol. 36, Iss. 2. P. 4075–4086. URL: <https://doi.org/10.1016/j.eswa.2008.03.020>.
- [24] Clustering of financial time series / P. D’Urso et al. *Physica A: Statistical Mechanics and its Applications*. 2013. Vol. 392, Iss. 9. P. 2114–2129. URL: <https://doi.org/10.1016/j.physa.2013.01.027>.
- [25] Continuous-time autoregressive models excited by semi-Lévy process for cyclostationary signal analysis / M. Mohammadi et al. *Digital Signal Processing*. 2021. Vol. 118. P. 103195. URL: <https://doi.org/10.1016/j.dsp.2021.103195> .
- [26] Costacurta J., Duncker L., Sheffer B., Williams A., Gillis W., Weinreb C., Markowitz J., Datta S., Linderman S. Distinguishing discrete and

- continuous behavioral variability using warped autoregressive HMMs. 2022. P. 19. URL:<https://doi.org/10.1101/2022.06.10.495690>.
- [27] Dablemont S., Simon G., Lendasse A., Ruttiens A., Blayo F. and Verleysen M. Time series forecasting with SOM and local non-linear models-Application to the DAX30 index prediction. *Proceedings of the workshop on self-organizing maps, Kitakyushu, Japan, Citeseer*. 2003. URL: <https://research.cs.aalto.fi/aml/Publications/Publication33.pdf>
- [28] Dai W., Wu J.-Y., Lu C.-J. Combining nonlinear independent component analysis and neural network for the prediction of Asian stock market indexes. *Expert Systems with Applications*. 2012. Vol. 39, Iss. 4. P. 4444–4452. URL: <https://doi.org/10.1016/j.eswa.2011.09.145>.
- [29] Dhar S., Mukherjee T., Ghoshal A.K. Performance evaluation of neural network approach in financial prediction: Evidence from Indian market. *International conference on communication and computational intelligence, IEEE*. 2010. P. 597-602.
- [30] Dinh-Tuan Pham. Estimation of continuous-time autoregressive model from finely sampled data. *IEEE Transactions on Signal Processing*. 2000. Vol. 48, Iss. 9. P. 2576–2584. URL: <https://doi.org/10.1109/78.863060> .
- [31] Durbin J., Koopman S. J. Linear state space models. *Time Series Analysis by State Space Methods*. 2012. P. 43–75. URL: <https://doi.org/10.1093/acprof:oso/9780199641178.003.0003>.
- [32] Dynamical generalized Hurst exponent as a tool to monitor unstable periods in financial time series / R. Morales et al. *Physica A: Statistical Mechanics and its Applications*. 2012. Vol. 391, Iss. 11. P. 3180–3189. URL: <https://doi.org/10.1016/j.physa.2012.01.004> .
- [33] Eckmann J. P., Ruelle D. Ergodic theory of chaos and strange attractors. *Reviews of Modern Physics*. 1985. Vol. 57, Iss. 3. P. 617–656. URL: <https://doi.org/10.1103/revmodphys.57.617> .

- [34] Evaluating multiple classifiers for stock price direction prediction / M. Ballings et al. *Expert Systems with Applications*. 2015. Vol. 42, Iss. 20. P. 7046–7056. URL: <https://doi.org/10.1016/j.eswa.2015.05.013>.
- [35] Evans C., Pappas K., Xhafa F. Utilizing artificial neural networks and genetic algorithms to build an algo-trading model for intra-day foreign exchange speculation. *Mathematical and Computer Modelling*. 2013. Vol. 58, Iss. 5-6. P. 1249–1266. URL: <https://doi.org/10.1016/j.mcm.2013.02.002>.
- [36] Extending Long Short-Term Memory for Multi-View Structured Learning / S. S. Rajagopalan et al. *Computer Vision – ECCV 2016*. Cham, 2016. P. 338–353. URL: https://doi.org/10.1007/978-3-319-46478-7_21.
- [37] Falconer K. *Fractal Geometry*. Chichester, UK : John Wiley & Sons, Ltd, 2003. URL: <https://doi.org/10.1002/0470013850> .
- [38] Farmer J. D., Ott E., Yorke J. A. The dimension of chaotic attractors. *Physica D: Nonlinear Phenomena*. 1983. Vol. 7, Iss. 1-3. P. 153–180. URL: [https://doi.org/10.1016/0167-2789\(83\)90125-2](https://doi.org/10.1016/0167-2789(83)90125-2) .
- [39] Feo F., Federico S., Świąch A. Optimal control of stochastic delay differential equations and applications to path-dependent financial and economic models. *arXiv preprint arXiv:2302.08809*. 2023. P. 23. URL: <https://doi.org/10.48550/arXiv.2302.08809>.
- [40] Ferreira C. Gene Expression Programming: A New Adaptive Algorithm for Solving Problems. *Complex Systems*, 13. 2001. P 87-129. URL: <https://doi.org/10.48550/arXiv.cs/0102027>.
- [41] F-LSTM: FPGA-Based Heterogeneous Computing Framework for Deploying LSTM-Based Algorithms / B. Liang et al. *Electronics*. 2023. Vol. 12, Iss. 5. P. 1139. URL: <https://doi.org/10.3390/electronics12051139> .
- [42] Fogel D. B. *Evolutionary Computation: The Fossil Record*. Wiley-IEEE Press, 1998. 656 p.

- [43] Forecasting stock indices with back propagation neural network / J.-Z. Wang et al. *Expert Systems with Applications*. 2011. URL: <https://doi.org/10.1016/j.eswa.2011.04.222> .
- [44] From an artificial neural network to a stock market day-trading system: A case study on the BM&F BOVESPA / L. C. Martinez et al. *2009 International Joint Conference on Neural Networks (IJCNN 2009 - Atlanta)*, Atlanta, Ga, USA, 14–19 June 2009. 2009. URL: <https://doi.org/10.1109/ijcnn.2009.5179050> .
- [45] Fuller R., Neural Fuzzy System. *Abo Akademic University*. 1995.
- [46] Geometry from a Time Series / N. H. Packard et al. *Physical Review Letters*. 1980. Vol. 45, Iss. 9. P. 712–716. URL: <https://doi.org/10.1103/physrevlett.45.712> .
- [47] Gheyas I. A., Smith L. S. A novel neural network ensemble architecture for time series forecasting. *Neurocomputing*. 2011. Vol. 74, Iss. 18. P. 3855–3864. URL: <https://doi.org/10.1016/j.neucom.2011.08.005> .
- [48] Ghosh I., Sanyal M. K., Jana R. K. Fractal Inspection and Machine Learning-Based Predictive Modelling Framework for Financial Markets. *Arabian Journal for Science and Engineering*. 2017. Vol. 43, Iss. 8. P. 4273–4287. URL: <https://doi.org/10.1007/s13369-017-2922-3> .
- [49] Gómez V. Automatic Model Identification in the Presence of Missing Observations and Outliers. Working paper D-98009. *Ministerio de Econom'ia y Hacienda, Dirección General de Análisis y Programación Presupuestaria*. 1998.
- [50] Gómez V., Maravall A. Programs TRAMO and SEATS, Instructions for the Users. Working paper 97001. *Ministerio de Econom'ia y Hacienda, Dirección General de Análisis y Programación Presupuestaria*. 1998.
- [51] Guyon I., Elisseeff A. An Introduction to Variable and Feature Selection. *JMLR* 3. 2003. P 1157-1182. URL: 10.1162/153244303322753616

- [52] Hannan E. J., Rissanen J. Recursive estimation of mixed autoregressive-moving average order. *Biometrika*. 1982. Vol. 69, Iss. 1. P. 81–94. URL: <https://doi.org/10.1093/biomet/69.1.81>.
- [53] Hanson J. E., Crutchfield J. P. The attractor-basin portrait of a cellular automaton. *Journal of Statistical Physics*. 1992. Vol. 66, Iss. 5-6. P. 1415–1462. URL: <https://doi.org/10.1007/bf01054429> .
- [54] Hasselblatt B., Katok A. A First Course in Dynamics: With a Panorama of Recent Developments. Cambridge University Press, 2003. 436 p
- [55] Hénon M. (1983) Numerical exploration of Hamiltonian Systems. *Chaotic Behaviour of Deterministic Systems*, Elsevier Science Ltd. 1983. P. 53–170.
- [56] Hentschel H. G. E., Procaccia I. The infinite number of generalized dimensions of fractals and strange attractors. *Physica D: Nonlinear Phenomena*. 1983. Vol. 8, Iss. 3. P. 435–444. URL: [https://doi.org/10.1016/0167-2789\(83\)90235-x](https://doi.org/10.1016/0167-2789(83)90235-x) .
- [57] Hochreiter S., Schmidhuber J. Long Short-Term Memory. *Neural Computation*. 1997. Vol. 9, Iss. 8. P. 1735–1780. URL: <https://doi.org/10.1162/neco.1997.9.8.1735>.
- [58] V. Hryhorkiv, L. Buiak, A. Verstiak, M. Hryhorkiv, O. Verstiak and K. Tokarieva, "Forecasting Financial Time Series Using Combined ARIMA-ANN Algorithm," 2020 10th International Conference on Advanced Computer Information Technologies (ACIT), Deggendorf, Germany, 16-18 September 2020, P. 455-458. ISBN: 978-172816760-2 (Scopus), (URL: <https://www.scopus.com/record/display.uri?eid=2-s2.0-85094125855&origin=resultslist>)
- [59] Hurst H. E. Long-Term Storage Capacity of Reservoirs. *Transactions of the American Society of Civil Engineers*. 1951. Vol. 116, Iss. 1. P. 770–799. URL: <https://doi.org/10.1061/taceat.0006518> .

- [60] Hyndman R. J., Khandakar Y. Automatic Time Series Forecasting: TheforecastPackage forR. *Journal of Statistical Software*. 2008. Vol. 27, Iss. 3. URL: <https://doi.org/10.18637/jss.v027.i03> .
- [61] Improving option price forecasts with neural networks and support vector regressions / X. Liang et al. *Neurocomputing*. 2009. Vol. 72, Iss. 13-15. P. 3055–3065. URL: <https://doi.org/10.1016/j.neucom.2009.03.015> .
- [62] Ismanto Edi. LSTM Network Hyperparameter Optimization for Stock Price Prediction Using the Optuna Framework. *Jurnal Ilmiah Teknik Elektro Komputer dan Informatika (JITEKI)*. 2023. Vol. 9, Iss. 1. P. 22-35. URL: [10.26555/jiteki.v9i1.24944](https://doi.org/10.26555/jiteki.v9i1.24944).
- [63] Jang J. S. R. ANFIS: adaptive-network-based fuzzy inference system. *IEEE Transactions on Systems, Man, and Cybernetics*. 1993. Vol. 23, Iss. 3. P. 665–685. URL: <https://doi.org/10.1109/21.256541> .
- [64] Julien C. Sprott. Do We Need More Chaos Examples? *CHAOS*. Vol . 2 Iss. 2. 2020. P. 1-3. URL: <https://www.researchgate.net/publication/341293656>.
- [65] Kamaruzzaman Z., Zaidi I., Ismail Mohd Tahir. Mixtures of Normal Distributions: Application to Bursa Malaysia Stock Market Indices. *World Applied Sciences Journal*. 2012. 16. P. 781-790.
- [66] Kantz H. *Nonlinear time series analysis*. 2nd ed. Cambridge, UK : Cambridge University Press, 2004. 369 p.
- [67] Karimi A., Paul M. R. Extensive chaos in the Lorenz-96 model. *Chaos: An Interdisciplinary Journal of Nonlinear Science*. 2010. Vol. 20, Iss. 4. P. 043105. URL: <https://doi.org/10.1063/1.3496397> .
- [68] Kennedy J., Eberhart R. Particle swarm optimization. *ICNN'95 - International Conference on Neural Networks*, Perth, WA, Australia. URL: <https://doi.org/10.1109/icnn.1995.488968> .
- [69] Khan E. Neural Fuzzy Based Intelligent Systems and Applications. *Fusion of Neural Networks, Fuzzy Systems, and Genetic Algorithms*. 2020. P. 105–140. URL: <https://doi.org/10.1201/9780367811464-5> .

- [70] Khashei M., Bijari M. A novel hybridization of artificial neural networks and ARIMA models for time series forecasting. *Applied Soft Computing*. 2011. Vol. 11, Iss. 2. P. 2664–2675. URL: <https://doi.org/10.1016/j.asoc.2010.10.015> .
- [71] Khashei M., Bijari M. An artificial neural network (p,d,q) model for timeseries forecasting. *Expert Systems with Applications*. 2010. Vol. 37, Iss. 1. P. 479–489. URL: <https://doi.org/10.1016/j.eswa.2009.05.044> .
- [72] Khashei M., Bijari M., Raissi Ardali G. A. Improvement of Auto-Regressive Integrated Moving Average models using Fuzzy logic and Artificial Neural Networks (ANNs). *Neurocomputing*. 2009. Vol. 72, Iss. 4-6. P. 956–967. URL: <https://doi.org/10.1016/j.neucom.2008.04.017> .
- [73] Kirshner H., Unser M., Ward J. P. On the Unique Identification of Continuous-Time Autoregressive Models From Sampled Data. *IEEE Transactions on Signal Processing*. 2014. Vol. 62, Iss. 6. P. 1361–1376. URL: <https://doi.org/10.1109/tsp.2013.2296879> .
- [74] Kohonen T. Self-organized formation of topologically correct feature maps. *Biological Cybernetics*. 1982. Vol. 43, Iss. 1. P. 59–69. URL: <https://doi.org/10.1007/bf00337288> .
- [75] Kumar D. A., Murugan S. Performance analysis of Indian stock market index using neural network time series model. *2013 International Conference on Pattern Recognition, Informatics and Mobile Engineering (PRIME)*, Salem, 21–22 February 2013. 2013. URL: <https://doi.org/10.1109/icprime.2013.6496450> .
- [76] Kushnir, M., Tokarieva, K. A Generalization of the Arima Model to the Nonlinear and Continuous Cases. *Cybernetics and Systems Analysis*. 2023. Vol. 59. Iss. 6. P. 900–909 (Scopus), (Q3 – <https://link.springer.com/article/10.1007/s10559-023-00625-8>)
- [77] M. Kushnir, A. Komarnytskyi, K. Tokarieva, N. Savchyn, P. Kroialo and V. Toronchuk, Technological and Legal Aspects of the Use of Machine Learning Elements in Chaotic Information Processing Systems. *2020 IEEE*

International Conference on Problems of Infocommunications. Science and Technology (PIC S&T), Kharkiv, Ukraine, 6-9 October 2020, P. 713-716. ISBN: 978-172819177-5 (Scopus)

- [78] Lakhmi C.J., Martin N.M. The CRC Press International Series on Computational Intelligence. 2000. P. 107–139.
- [79] Lam M. Neural network techniques for financial performance prediction: integrating fundamental and technical analysis. *Decision Support Systems*. 2004. Vol. 37, Iss. 4. P. 567–581. URL: [https://doi.org/10.1016/s0167-9236\(03\)00088-5](https://doi.org/10.1016/s0167-9236(03)00088-5) .
- [80] Lee S. Mouse Genetics Concepts and Applications. 2023 URL:10.1093/oso/9780195075540.001.0001.
- [81] Lee M.-C. Using support vector machine with a hybrid feature selection method to the stock trend prediction. *Expert Systems with Applications*. 2009. Vol. 36, Iss. 8. P. 10896–10904. URL: <https://doi.org/10.1016/j.eswa.2009.02.038> .
- [82] Lei Wu, Shahidehpour M. A Hybrid Model for Day-Ahead Price Forecasting. *IEEE Transactions on Power Systems*. 2010. Vol. 25, Iss. 3. P. 1519–1530. URL: <https://doi.org/10.1109/tpwrs.2009.2039948> .
- [83] Li Y., Ma W. Applications of Artificial Neural Networks in Financial Economics: A Survey. *2010 3rd International Symposium on Computational Intelligence and Design (ISCID)*, Hangzhou, China, 29–31 October 2010. 2010. URL: <https://doi.org/10.1109/iscid.2010.70> .
- [84] Lin C.-S., Chiu S.-H., Lin T.-Y. Empirical mode decomposition–based least squares support vector regression for foreign exchange rate forecasting. *Economic Modelling*. 2012. Vol. 29, Iss. 6. P. 2583–2590. URL: <https://doi.org/10.1016/j.econmod.2012.07.018> .
- [85] Liu L.-M. Identification of seasonal arima models using a filtering method. *Communications in Statistics - Theory and Methods*. 1989. Vol. 18, Iss. 6. P. 2279–2288. URL: <https://doi.org/10.1080/03610928908830035> .

- [86] Localized support vector regression for time series prediction / H. Yang et al. *Neurocomputing*. 2009. Vol. 72, Iss. 10-12. P. 2659–2669. URL: <https://doi.org/10.1016/j.neucom.2008.09.014> .
- [87] Logistic Map: Stability and Entrance to Chaos / S. Chen et al. *Journal of Physics: Conference Series*. 2021. Vol. 2014, Iss. 1. P. 012009. URL: <https://doi.org/10.1088/1742-6596/2014/1/012009> .
- [88] Lorenz E. N. Deterministic Nonperiodic Flow. *Journal of the Atmospheric Sciences*. 1963. Vol. 20, Iss. 2. P. 130–141. URL: [https://doi.org/10.1175/1520-0469\(1963\)020%3C0130:dnf%3E2.0.co;2](https://doi.org/10.1175/1520-0469(1963)020%3C0130:dnf%3E2.0.co;2) .
- [89] Lu C.-J., Lee T.-S., Chiu C.-C. Financial time series forecasting using independent component analysis and support vector regression. *Decision Support Systems*. 2009. Vol. 47, Iss. 2. P. 115–125. URL: <https://doi.org/10.1016/j.dss.2009.02.001> .
- [90] M G. E. The goal: A process of ongoing improvement. 2nd ed. Aldershot, Hampshire, England : Gower, 1993. 333 p.
- [91] Majhi R., Panda G., Sahoo G. Efficient prediction of exchange rates with low complexity artificial neural network models. *Expert Systems with Applications*. 2009. Vol. 36, Iss. 1. P. 181–189. URL: <https://doi.org/10.1016/j.eswa.2007.09.005> .
- [92] Memory Fusion Network for Multi-view Sequential Learning / A. Zadeh et al. *Proceedings of the AAAI Conference on Artificial Intelligence*. 2018. Vol. 32, Iss. 1. URL: <https://doi.org/10.1609/aaai.v32i1.12021>.
- [93] Memristor-based LSTM Network for Text Classification / G. Dou et al. *Fractals*. 2023. URL: <https://doi.org/10.1142/s0218348x23400406>.
- [94] Meyler A. Forecasting Irish inflation using ARIMA models. Dublin : Central Bank of Ireland, Economic Analysis, Research and Publications Department, 1998. 46 p.

- [95] Mixtures of t-distributions for finance and forecasting / R. Giacomini et al. *Journal of Econometrics*. 2008. Vol. 144, Iss. 1. P. 175–192. URL: <https://doi.org/10.1016/j.jeconom.2008.01.004> .
- [96] Monner D., Reggia J. A. A generalized LSTM-like training algorithm for second-order recurrent neural networks. *Neural Networks*. 2012. Vol. 25. P. 70–83. URL: <https://doi.org/10.1016/j.neunet.2011.07.003> .
- [97] Murphy J.J. Technical analysis of the financial markets. *New York Institute of Finance*. 1999.
- [98] Neural network stochastic differential equation models with applications to financial data forecasting / L. Yang et al. *Applied Mathematical Modelling*. 2022. URL: <https://doi.org/10.1016/j.apm.2022.11.001> .
- [99] Nikfarjam A., Emadzadeh E., Muthaiyah S. Text mining approaches for stock market prediction. *2nd International Conference on Computer and Automation Engineering (ICCAE 2010)*, Singapore, 26–28 February 2010. 2010. URL: <https://doi.org/10.1109/iccae.2010.5451705>.
- [100] Nonlinear forecasting for the classification of natural time series. *Philosophical Transactions of the Royal Society of London. Series A: Physical and Engineering Sciences*. 1994. Vol. 348, Iss. 1688. P. 477–495. URL: <https://doi.org/10.1098/rsta.1994.0106> .
- [101] Non-stationary and stationary prediction of financial time series using dynamic ridge polynomial neural network / R. Ghazali et al. *Neurocomputing*. 2009. Vol. 72, Iss. 10-12. P. 2359–2367. URL: <https://doi.org/10.1016/j.neucom.2008.12.005> .
- [102] Palit A. K., Popovic D. *Computational Intelligence in Time Series Forecasting: Theory and Engineering Applications*. Springer London, Limited, 2005.
- [103] Qian B., Rasheed K. Hurst exponent and financial market predictability. IASTED conference on Financial Engineering and Applications. 2004. P. 203–209.

- [104] Robert M. Simple mathematical models with very complicated dynamics. *Nature* 261. 1976. P. 459–467.
- [105] Roshan W. D. S., Gopura R. A. R. C., Jayasekara A. G. B. P. Financial forecasting based on artificial neural networks: Promising directions for modeling. *2011 IEEE 6th International Conference on Industrial and Information Systems (ICIIS)*, Kandy, Sri Lanka, 16–19 August 2011. 2011. URL: <https://doi.org/10.1109/iciinfs.2011.6038088>
- [106] Rüschemdorf L. *Stochastic Processes and Financial Mathematics*. Berlin, Heidelberg : Springer Berlin Heidelberg, 2023. URL: <https://doi.org/10.1007/978-3-662-64711-0> .
- [107] S&P Dow Jones Indices. веб-сайт. URL: <https://us.spindices.com/>.
- [108] Samir Mohamed. Comparison between ANN and RNN. веб-сайт. URL: <http://www.irisa.fr/armor/lesmembres/Mohamed/Thesis/node151.html>.
- [109] Schilling R. L., Böttcher B. *Brownian Motion: A Guide to Random Processes and Stochastic Calculus*. de Gruyter GmbH, Walter, 2021.
- [110] Shahpazov V. L., Velez V. B., Doukovska L. A. Design and application of Artificial Neural Networks for predicting the values of indexes on the Bulgarian Stock market. *2013 Signal Processing Symposium (SPS)*, Serock, Poland, 5–7 June 2013. 2013. URL: <https://doi.org/10.1109/sps.2013.6623604> .
- [111] Si Y.-W., Yin J. OBST-based segmentation approach to financial time series. *Engineering Applications of Artificial Intelligence*. 2013. Vol. 26, Iss. 10. P. 2581–2596. URL: <https://doi.org/10.1016/j.engappai.2013.08.015> .
- [112] Snoek C. G. M., Worring M., Smeulders A. W. M. Early versus late fusion in semantic video analysis. *the 13th annual ACM international conference*, Hilton, Singapore, 6–11 November 2005. New York, New York, USA, 2005. URL: <https://doi.org/10.1145/1101149.1101236> .

- [113] Soni S. Applications of ANNs in stock market prediction: a survey. *International Journal of Computer Science and Engineering Technology*. 2011. 2 (3). P. 71-83.
- [114] Sprott J. C. Introduction. *Chaos and Time-Series Analysis*. 2003. P. 1–19. URL: <https://doi.org/10.1093/oso/9780198508397.003.0001> .
- [115] Steves B., Maciejewski A. The Restless Universe Applications of Gravitational N-Body Dynamics to Planetary Stellar and Galactic Systems. Taylor & Francis, 2001. URL: <https://doi.org/10.1201/9781420033304>.
- [116] Stewart I. The Lorenz attractor exists. *Nature*. 2000. Vol. 406, Iss. 6799. P. 948–949. URL: <https://doi.org/10.1038/35023206> .
- [117] Stock index forecasting based on a hybrid model / J.-J. Wang et al. *Omega*. 2012. Vol. 40, Iss. 6. P. 758–766. URL: <https://doi.org/10.1016/j.omega.2011.07.008> .
- [118] Su F. The Chaos Theory and its Application. *Journal of Physics: Conference Series*. 2021. Vol. 2012, Iss. 1. P. 012118. URL: <https://doi.org/10.1088/1742-6596/2012/1/012118> .
- [119] Takens F. Detecting strange attractors in turbulence. *Lecture Notes in Mathematics*. Berlin, Heidelberg, 1981. P. 366–381. URL: <https://doi.org/10.1007/bfb0091924> .
- [120] Tay F. E. H., Cao L. Application of support vector machines in financial time series forecasting. *Omega*. 2001. Vol. 29, Iss. 4. P. 309–317. URL: [https://doi.org/10.1016/s0305-0483\(01\)00026-3](https://doi.org/10.1016/s0305-0483(01)00026-3) .
- [121] Teixeira L. A., de Oliveira A. L. I. A method for automatic stock trading combining technical analysis and nearest neighbor classification. *Expert Systems with Applications*. 2010. Vol. 37, Iss. 10. P. 6885–6890. URL: <https://doi.org/10.1016/j.eswa.2010.03.033>
- [122] The Nonlinear Dynamics Characteristics of Stock Market and Its Variation / J. Meng et al. *2nd International Conference on Computer Science and Electronics Engineering (ICCSEE 2013)*, China, 22–23

March 2013. Paris, France, 2013.
URL: <https://doi.org/10.2991/iccsee.2013.116> .

- [123] The recognition of time-compressed speech as a function of age in listeners with cochlear implants or normal hearing / A. R. Tinnemore et al. *Frontiers in Aging Neuroscience*. 2022. Vol. 14.
URL: <https://doi.org/10.3389/fnagi.2022.887581> .
- [124] The visibility graph: A new method for estimating the Hurst exponent of fractional Brownian motion / L. Lacasa et al. *EPL (Europhysics Letters)*. 2009. Vol. 86, Iss. 3. P. 30001. URL: <https://doi.org/10.1209/0295-5075/86/30001> .
- [125] Ticknor J. L. A Bayesian regularized artificial neural network for stock market forecasting. *Expert Systems with Applications*. 2013. Vol. 40, no. 14. P. 5501–5506. URL: <https://doi.org/10.1016/j.eswa.2013.04.013> .
- [126] Tkáč M., Verner R. Artificial neural networks in business: Two decades of research. *Applied Soft Computing*. 2016. Vol. 38. P. 788–804.
URL: <https://doi.org/10.1016/j.asoc.2015.09.040> .
- [127] Tsai C.-F., Hsiao Y.-C. Combining multiple feature selection methods for stock prediction: Union, intersection, and multi-intersection approaches. *Decision Support Systems*. 2010. Vol. 50, Iss. 1. P. 258–269.
URL: <https://doi.org/10.1016/j.dss.2010.08.028> .
- [128] Vanstone B., Finnie G. An empirical methodology for developing stockmarket trading systems using artificial neural networks. *Expert Systems with Applications*. 2009. Vol. 36, Iss. 3. P. 6668–6680.
URL: <https://doi.org/10.1016/j.eswa.2008.08.019> .
- [129] Vanstone B., Finnie G. Enhancing stockmarket trading performance with ANNs. *Expert Systems with Applications*. 2010. Vol. 37, Iss. 9. P. 6602–6610. URL: <https://doi.org/10.1016/j.eswa.2010.02.124> .
- [130] Wang B., Huang H., Wang X. A novel text mining approach to financial time series forecasting. *Neurocomputing*. 2012. Vol. 83. P. 136–145.
URL: <https://doi.org/10.1016/j.neucom.2011.12.013> .

- [131] Webb A. R. Statistical Pattern Recognition, 2nd Edition. Wiley, 2002. 514 p.
- [132] Wilding, Richard D. The Detection and Quantification of Chaos in Supply Chains.1999.
- [133] Xu H., Hu B. Legal Text Recognition Using LSTM-CRF Deep Learning Model. *Computational Intelligence and Neuroscience*. 2022. Vol. 2022. P. 1–10. URL: <https://doi.org/10.1155/2022/9933929> .
- [134] Yoo P. D., Kim M. H., Jan T. Machine Learning Techniques and Use of Event Information for Stock Market Prediction: A Survey and Evaluation. *International Conference on Computational Intelligence for Modelling, Control and Automation and International Conference on Intelligent Agents, Web Technologies and Internet Commerce (CIMCA-IAWTIC'06)*, Vienna, Austria. URL: <https://doi.org/10.1109/cimca.2005.1631572> .
- [135] Yuan Y. Forecasting the movement direction of exchange rate with polynomial smooth support vector machine. *Mathematical and Computer Modelling*. 2013. Vol. 57, Iss. 3-4. P. 932–944. URL: <https://doi.org/10.1016/j.mcm.2012.10.004> .
- [136] Zaslavsky G. Hamiltonian Chaos and Fractional Dynamics. 2004. URL:10.1093/oso/9780198526049.001.0001.
- [137] Zhang G. P. Time series forecasting using a hybrid ARIMA and neural network model. *Neurocomputing*. 2003. Vol. 50. P. 159–175. URL: [https://doi.org/10.1016/s0925-2312\(01\)00702-0](https://doi.org/10.1016/s0925-2312(01)00702-0) .
- [138] Zhang G., Eddy Patuwo B., Y. Hu M. Forecasting with artificial neural networks:. *International Journal of Forecasting*. 1998. Vol. 14, Iss. 1. P. 35–62. URL: [https://doi.org/10.1016/s0169-2070\(97\)00044-7](https://doi.org/10.1016/s0169-2070(97)00044-7) .
- [139] Zheng Y. Chaotic Phenomenal and One-dimensional Logistic Map. *Highlights in Science, Engineering and Technology*. 2023. Vol. 72. P. 638–644. URL: <https://doi.org/10.54097/wd5v9r36> .

- [140] Zhou D., Li J., Ma W. Clustering Based on LLE For Financial Multivariate Time Series. *2009 International Conference on Management and Service Science (MASS)*, Beijing, China, 20–22 September 2009. 2009. URL: <https://doi.org/10.1109/icmss.2009.5305089>.
- [141] Атоєв К.Л., Кнопов П.С., Пепеляєва Т.В. Розробка нових моделей оцінювання ефективності природокористування за умов змін клімату та зростання невизначеності. *Теорія оптимальних рішень*. 2017. С. 72-77.
- [142] Гаращенко Ф.Г., Кулян В.Р., Петрович В.Н., Юнькова Е.А. Моделирование динамики и диверсификация портфеля акций. *Проблемы управления и информатики*. 2016. № 4. С. 124-135.
- [143] Гихман И. И., Скороход А. В. Стохастические дифференциальные уравнения. К.: *Наукова думка*, 1968. 355 с.
- [144] Демченко С.С., Кнопов П.С., Чорней Р.К. Оптимальні стратегії для напівмарковської системи запасів. *Кибернетика и системный анализ*. 2002. № 1. С. 146-160.
- [145] Кнопов П.С., Богданов А.В. Использование стохастической модели для прогнозирования длительных эпидемий. *Проблемы управления и информатики*. 2021. № 3. С. 50-57.
- [146] Кнопов П.С., Касицька Є.Й. Состоятельность и свойства больших уклонений эмпирических оценок в задаче стохастической оптимизации для однородного случайного поля при неоднородных и однородных наблюдениях. *Кибернетика и системный анализ*. 2021. № 1. С. 21-34.
- [147] Кнопов П.С., Пепеляева Т.В., Демченко И.Ю. Об одной полумарковской модели управления запасами. *Кибернетика и системный анализ*. 2016. Т. 52, № 5. С. 81-88.
- [148] Кнопов П.С., Пепеляева Т.В. Про деякі торгові стратегії на ринку цінних паперів. *Кибернетика и системный анализ*. 2002. № 5. С. 117-121.

- [149] Кнопова В.П., Пепеляєва Т.В. Про деякі стохастичні моделі фінансової математики. *Кибернетика и системный анализ*. 2001. № 3. С. 152-158.
- [150] Кушнір М.Я., Токарева К.А. Використання Систем Штучного Інтелекту У Задачах Прогнозування Фінансових Індексів: Огляд Наукових Джерел / Artificial Intelligence Systems In The Financial Market Predictions: Literature Review. *Radioelectronic and Computer Systems*. 2020. Is. 3(95). С. 108-117.
- [151] Кушнір М.Я., Токарева К.А. Одне узагальнення ARIMA-моделі на нелінійний та неперервний випадки / A Generalization of the Arima Model to the Nonlinear and Continuous Cases. *Cybernetics and Systems Analysis*. 2023. Vol. 59. Iss. 6. P. 900 – 909. (Scopus), (Q3 - URL: <https://www.scimagojr.com/journalsearch.php?q=12933&tip=sid&clean=0>)
- [152] Кушнір М.Я., Токарева К.А. Гібридна модель самоорганізаційної карти кохонена та адаптивної нейро-нечіткої системи виводу у задачах прогнозування цін фондових індексів. *Буковинський математичний журнал*. 2021. Т. 9. № 2. С. 70–80. (URL: <https://doi.org/10.31861/bmj2021.02.05>)
- [153] Кушнір, М. Я., Токарева, К. А. Одне узагальнення LSTM-нейронних мереж. *International Scientific Technical Journal "Problems of Control and Informatics"*. 2023. 68(4). С. 111–118. (URL: <https://jais.net.ua/index.php/files/article/view/118>)
- [154] Норкін Б.В. Застосування методу послідовних наближень для знаходження ймовірності нерозорення страхової компанії за наявності випадкових премій. *Кибернетика и системный анализ*. 2006. № 1. С.112-127.
- [155] Норкін Б.В. Стохастичний метод послідовних наближень для оцінки ризику неплатоспроможності страхової компанії. *Кибернетика и системный анализ*. 2008. № 6. С.116-130.

- [156] Пепеляєва Т.В. Про оптимальні моменти переключення між портфелями цінних паперів. *Кибернетика и системный анализ*. 2002. № 1. С. 130-137.
- [157] Токарева К.А. Використання систем штучного інтелекту у прогнозуванні фінансових ринків. Міжнародний науковий симпозиум «Інтелектуальні рішення». Обчислювальний інтелект (результати, проблеми, перспективи): праці міжнар. наук.-практ. конф., 15-20 квітня 2019 р., Ужгород / М-во освіти і науки України, ДВНЗ «Ужгородський національний університет», та [ін.]; наук. ред. В.Є. Снитюк. С. 54-55. (URL: <https://er.chdtu.edu.ua/bitstream/ChSTU/3192/1/OI-2019.pdf>)
- [158] Токарева К.А. Лог-рентабельність активів у дослідженні фінансових часових рядів. Математичні методи, моделі та інформаційні технології в економіці: Матеріали VI Міжнародної науково-методичної конференції. Чернівці: Друк Арт, 6 вересня 2019. С. 154-155.

ДОДАТОК

СПИСОК ПУБЛІКАЦІЙ ЗА ТЕМОЮ ДИСЕРТАЦІЇ

Наукові праці у періодичних наукових виданнях, проіндексованих у наукометричній базі даних Scopus:

1. Kushnir, M., Tokarieva, K. A Generalization of the Arima Model to the Nonlinear and Continuous Cases. *Cybernetics and Systems Analysis*. 2023. Vol. 59. Iss. 6. P. 900–909 (Scopus), (Q3 – <https://link.springer.com/article/10.1007/s10559-023-00625-8>)
2. Кушнір М.Я., Токарева К.А. Використання Систем Штучного Інтелекту У Задачах Прогнозування Фінансових Індексів: Огляд Наукових Джерел / Artificial Intelligence Systems In The Financial Market Predictions: Literature Review. *Radioelectronic and Computer Systems*. 2020. Is. 3(95). С. 108-117 (Scopus), (URL: <https://www.scopus.com/record/display.uri?eid=2-s2.0-85101514809&origin=resultslist>)

Наукові праці у виданнях, включених до переліку наукових фахових видань України:

3. Кушнір М.Я., Токарева К.А. Одне узагальнення ARIMA-моделі на нелінійний та неперервний випадки / A Generalization of the Arima Model to the Nonlinear and Continuous Cases. *Cybernetics and Systems Analysis*. 2023. Vol. 59. Iss. 6. P. 900 – 909. (Scopus), (Q3 - URL: <https://www.scimagojr.com/journalsearch.php?q=12933&tip=sid&clean=0>)
4. Кушнір М.Я., Токарева К.А. Гібридна модель самоорганізаційної карти Кохонена та адаптивної нейро-нечіткої системи виводу у задачах прогнозування цін фондових індексів. *Буковинський математичний журнал*. 2021. Т. 9. № 2. С. 70–80. (URL: <https://doi.org/10.31861/bmj2021.02.05>)
5. Кушнір, М. Я., Токарева, К. А. Одне узагальнення LSTM-нейронних мереж. *International Scientific Technical Journal "Problems of Control and Informatics"*. 2023. 68(4). С. 111–118. (URL:

<https://jais.net.ua/index.php/files/article/view/118>)

Наукові праці, які засвідчують апробацію матеріалів дисертації:

6. M. Kushnir, A. Komarnytskyi, K. Tokarieva, N. Savchyn, P. Kroialo and V. Toronchuk, Technological and Legal Aspects of the Use of Machine Learning Elements in Chaotic Information Processing Systems. *2020 IEEE International Conference on Problems of Infocommunications. Science and Technology (PIC S&T)*, Kharkiv, Ukraine, 6-9 October 2020, P. 713-716. ISBN: 978-172819177-5 (Scopus) (URL: <https://www.scopus.com/record/display.uri?eid=2-s2.0-85114410889&origin=resultslist>)
7. V. Hryhorkiv, L. Buiak, A. Verstiak, M. Hryhorkiv, O. Verstiak and K. Tokarieva, "Forecasting Financial Time Series Using Combined ARIMA-ANN Algorithm," *2020 10th International Conference on Advanced Computer Information Technologies (ACIT)*, Deggendorf, Germany, 16-18 September 2020, P. 455-458. ISBN: 978-172816760-2 (Scopus), (URL: <https://www.scopus.com/record/display.uri?eid=2-s2.0-85094125855&origin=resultslist>)
8. Токарева К.А. Використання систем штучного інтелекту у прогнозуванні фінансових ринків. *Міжнародний науковий симпозиум «Інтелектуальні рішення»*. Обчислювальний інтелект (результати, проблеми, перспективи): праці міжнар. наук.-практ. конф., 15-20 квітня 2019 р., Ужгород / М-во освіти і науки України, ДВНЗ «Ужгородський національний університет», та [ін.]; наук. ред. В.Є. Снитюк. С. 54-55. (URL: <https://er.chdtu.edu.ua/bitstream/ChSTU/3192/1/OI-2019.pdf>)
9. Токарева К.А. Лог-рентабельність активів у дослідженні фінансових часових рядів. *Математичні методи, моделі та інформаційні технології в економіці*: Матеріали VI Міжнародної науково-методичної конференції. Чернівці: Друк Арт, 6 вересня 2019. С. 154-155.