МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ ВОДНОГО ГОСПОДАРСТВА ТА ПРИРОДОКОРИСТУВАННЯ

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ ВОДНОГО ГОСПОДАРСТВА ТА ПРИРОДОКОРИСТУВАННЯ

Кваліфікаційна наукова
праця на правах рукопису

КУБАЙ ОЛЕКСАНДР ВАСИЛЬОВИЧ

УДК 004.8:004.62

ДИСЕРТАЦІЯ

МЕТОДИ ІНТЕЛЕКТУАЛЬНОГО АНАЛІЗУ ТА ПРОГНОЗУВАННЯ ЧАСОВИХ РЯДІВ БІРЖОВОЇ ДИНАМІКИ НА ОСНОВІ ПОЛІНОМІАЛЬНОГО ПІДХОДУ

122 – Комп'ютерні науки

12 – Інформаційні технології

Подається на здобуття ступеня доктора філософії

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

_____/О.В. Кубай/

Науковий керівник Турбал Юрій Васильович, доктор технічних наук, професор

Рівне – 2026

АНОТАЦІЯ

Кубай О.В. – Методи інтелектуального аналізу та прогнозування часових рядів біржової динаміки на основі поліноміального підходу. – Кваліфікаційна праця на правах рукопису.

Дисертація на здобуття доктора філософії за спеціальністю 122 «Комп'ютерні науки». – Національний університет водного господарства та природокористування, Рівне, 2026.

Зміст анотації. Дисертаційна робота присвячена вирішенню актуального наукового завдання – розробці та вдосконаленню методів інтелектуального аналізу і прогнозування часових рядів біржової динаміки на основі поліноміального підходу. Актуальність дослідження зумовлена потребою у створенні адаптивних, інтерпретованих і придатних до застосування в умовах обмеженого обсягу даних прогнозних моделей, які дають змогу враховувати локальну структуру фінансових часових рядів, високу волатильність, шумовий характер біржових процесів і недетермінованість ринкової динаміки.

У роботі інтегровано підходи математичного моделювання, поліноміальної екстраполяції, інтелектуального аналізу даних, кластерного аналізу, машинного навчання та нейромережевого моделювання. Запропонований підхід ґрунтується на переході від побудови одиничного поліноміального прогнозу за фіксованого степеня полінома до формування та аналізу послідовності поліноміальних прогнозних значень PPS (Polynomial Predictions Sequence), яка розглядається як окремий інформаційний об'єкт, що характеризує локальну поведінку часового ряду біржової динаміки.

У вступі обґрунтовано актуальність теми дослідження, її зв'язок із науково-дослідними темами, сформульовано мету, завдання, об'єкт і предмет дослідження. Наведено опис застосованих наукових методів, зокрема методів системного аналізу, математичного моделювання часових рядів, поліноміальної екстраполяції, скінченних різниць, кластерного аналізу, машинного навчання, нейромережевого моделювання та статистичного оцінювання точності прогнозування. Визначено наукову новизну, практичне значення одержаних

результатів, особистий внесок здобувача, а також засвідчено апробацію результатів дослідження на конференціях і у фахових виданнях.

У першому розділі «Теоретичні засади інтелектуального аналізу та прогнозування часових рядів біржової динаміки» проведено системне дослідження сучасного стану задачі прогнозування фінансових часових рядів. Розглянуто часові ряди біржової динаміки як об’єкт інтелектуального аналізу, визначено їхні ключові властивості, зокрема шум, волатильність, нестаціонарність, фрактальність, локальні тренди, короткочасні структурні зміни та специфіку внутрішньоденних біржових даних. Проаналізовано класичні статистичні методи прогнозування, методи машинного навчання та нейромережеві підходи до моделювання фінансової динаміки.

На основі аналізу наукової літератури встановлено, що значна частина існуючих підходів орієнтована на використання великих навчальних вибірок, фіксованої структури моделі або наперед заданих параметрів прогнозування. Виявлено недостатню опрацьованість питань адаптивного вибору параметрів прогнозної моделі в умовах малого обсягу історичних даних і швидкої зміни локальної структури біржового часового ряду. Обґрунтовано доцільність використання поліноміального підходу, аналізу послідовності поліноміальних прогнозних значень PPS та методів машинного навчання для підвищення точності короткострокового прогнозування біржової динаміки.

У другому розділі «Методичні основи поліноміального підходу до аналізу часових рядів біржової динаміки» формалізовано задачу короткострокового прогнозування біржового часового ряду. Розглянуто локальний фрагмент ряду як основу для побудови прогнозу та введено поняття послідовності поліноміальних прогнозних значень PPS. Показано, що PPS формується як множина прогнозів, отриманих для різних значень параметра m, який визначає глибину локальної передісторії та відповідний степінь поліноміальної екстраполяції.

Обґрунтовано перехід від одиничного прогнозного значення до комплексного аналізу всієї послідовності PPS. Встановлено, що така послідовність відображає локальні властивості часового ряду, зокрема збіжність

або розбіжність прогнозних значень, наявність локальних екстремумів, монотонних ділянок і областей згущення. Досліджено властивості PPS як окремого інформаційного об’єкта, що дає змогу оцінювати чутливість прогнозу до зміни параметрів поліноміальної екстраполяції та формувати підґрунтя для адаптивного вибору прогнозної моделі.

У третьому розділі «Методи адаптивного прогнозування часових рядів біржової динаміки на основі послідовності поліноміальних прогнозів» розроблено та вдосконалено методи адаптивного формування прогнозу на основі інтелектуального аналізу PPS. Запропоновано підходи до вибору оптимального значення параметра m або інформативної підмножини поліноміальних прогнозних значень. Розглянуто методи, що ґрунтуються на аналізі локальних екстремумів PPS, виявленні областей згущення прогнозних значень, кластеризації, DBSCAN-аналізі та усередненні відібраних елементів послідовності.

Особливу увагу приділено процедурі формування остаточного прогнозного значення на основі внутрішніх характеристик PPS без залучення додаткових зовнішніх факторів. Такий підхід підвищує придатність розроблених методів до використання в умовах обмеженого обсягу даних, коли складні багатофакторні моделі можуть бути неефективними або нестабільними. Визначено межі застосовності запропонованих методів, зокрема випадки швидкої розбіжності PPS, відсутності виражених областей згущення або нестабільної локальної структури прогнозних значень.

У четвертому розділі «Розроблення власної архітектури нейронної мережі для прогнозування біржових часових рядів на основі поліноміального підходу» запропоновано власну архітектуру нейронної мережі PPS-Net, орієнтовану на використання послідовності поліноміальних прогнозних значень. Архітектура включає спеціальний блок формування PPS, блок нейромережевого аналізу її структури та блок адаптивного зважування поліноміальних прогнозів для різних значень параметра m.

На відміну від традиційних нейромережевих моделей, запропонована архітектура використовує не лише початкові значення локального фрагмента часового ряду, а й множину поліноміально обчислених прогнозних значень. Остаточний прогноз формується як зважена комбінація елементів PPS, що дає змогу поєднати аналітичні переваги поліноміальної екстраполяції з навчальною здатністю нейронної мережі. Ваги значущості поліноміальних прогнозів забезпечують додаткову інтерпретованість моделі, оскільки дозволяють визначити, які саме елементи PPS найбільше впливають на формування прогнозного результату.

У п’ятому розділі «Експериментальне дослідження ефективності розроблених методів» сформовано науково-прикладні засади оцінювання запропонованих методів і нейромережевої архітектури. Експериментальне дослідження виконано на модельних, стохастичних і реальних часових рядах біржової динаміки. Для перевірки ефективності методів використано процедури walk-forward перевірки та one-step-ahead прогнозування, які відповідають реальним умовам послідовного надходження біржових даних.

Оцінювання точності прогнозування здійснено за допомогою метрик MAE, MSE, RMSE і MAPE. Проведено порівняння запропонованих поліноміальних методів, адаптивних алгоритмів аналізу PPS, кластерного підходу, DBSCAN-аналізу та нейромережевої архітектури PPS-Net із класичними статистичними, машинними та нейромережевими методами прогнозування. Показано, що аналіз структури PPS, наявності кластерів прогнозних значень, локальних екстремумів і ваг нейромережевої архітектури дає змогу не лише оцінювати числову похибку прогнозу, а й інтерпретувати поведінку моделі в різних режимах біржової динаміки.

Результати дисертаційного дослідження, включаючи запропоновані методи адаптивної поліноміальної екстраполяції, алгоритми аналізу PPS та нейромережеву архітектуру PPS-Net, можуть бути використані під час розроблення програмних засобів інтелектуального аналізу фінансових часових рядів, систем підтримки прийняття інвестиційних рішень, модулів

короткострокового прогнозування біржових котирувань, а також у навчальному процесі під час викладання дисциплін, пов’язаних із машинним навчанням, аналізом даних, інтелектуальними системами та прогнозуванням часових рядів.

Практична цінність роботи полягає у створенні методичного та алгоритмічного інструментарію, що забезпечує адаптивне формування прогнозу на основі локальної інформації часового ряду, дає змогу працювати з обмеженими вибірками та підвищує інтерпретованість результатів прогнозування. Запропоновані підходи можуть становити наукову основу для подальшого розвитку інформаційних технологій обробки фінансових даних, інтелектуального аналізу біржової динаміки та побудови гібридних моделей прогнозування, які поєднують математичну формалізацію, поліноміальну екстраполяцію та нейромережеве навчання.

Основні наукові результати дисертації опубліковано у 11 працях, зокрема у 6 статтях у наукових фахових виданнях України, та 5 публікаціях у матеріалах міжнародних і всеукраїнських науково-практичних конференцій.

Ключові слова: інформаційні технології, інтелектуальний аналіз даних, часові ряди, біржова динаміка, фінансові часові ряди, короткострокове прогнозування, поліноміальний підхід, поліноміальна екстраполяція, послідовність поліноміальних прогнозних значень, PPS, адаптивне прогнозування, машинне навчання, нейронні мережі, PPS-Net, кластерний аналіз, DBSCAN, walk-forward перевірка, one-step-ahead прогнозування.

ABSTRACT

Kubai O. V. – Methods of Intelligent Analysis and Forecasting of Time Series of Stock Exchange Dynamics Based on a Polynomial Approach. – Qualification scientific paper as a manuscript.

Dissertation for the degree of Doctor of Philosophy in specialty 122 «Computer Science». – National University of Water and Environmental Engineering, Rivne, 2026.

Abstract content. The dissertation is devoted to solving an urgent scientific problem, namely the development and improvement of methods for intelligent analysis and forecasting of time series of stock exchange dynamics based on a polynomial approach. The relevance of the study is determined by the need to create adaptive, interpretable forecasting models suitable for application under conditions of limited data availability. Such models make it possible to take into account the local structure of financial time series, high volatility, the noisy nature of stock exchange processes, and the nondeterministic character of market dynamics.

The research integrates approaches of mathematical modelling, polynomial extrapolation, intelligent data analysis, cluster analysis, machine learning, and neural network modelling. The proposed approach is based on the transition from constructing a single polynomial forecast with a fixed polynomial degree to forming and analyzing a sequence of polynomial forecast values, PPS (Polynomial Predictions Sequence), which is considered as a separate information object characterizing the local behavior of a stock exchange time series.

The introduction substantiates the relevance of the research topic and its connection with research themes, and formulates the aim, objectives, object, and subject of the study. It presents the scientific methods applied in the research, including system analysis, mathematical modelling of time series, polynomial extrapolation, finite differences, cluster analysis, machine learning, neural network modelling, and statistical evaluation of forecasting accuracy. The scientific novelty, practical significance of the obtained results, personal contribution of the applicant, and

approbation of the research results at conferences and in professional publications are also specified.

The first chapter, “Theoretical Foundations of Intelligent Analysis and Forecasting of Time Series of Stock Exchange Dynamics”, presents a systematic study of the current state of the problem of financial time series forecasting. Time series of stock exchange dynamics are considered as an object of intelligent analysis, and their key properties are identified, including noise, volatility, nonstationarity, fractality, local trends, short-term structural changes, and the specificity of intraday stock exchange data. Classical statistical forecasting methods, machine learning methods, and neural network approaches to modelling financial dynamics are analyzed.

Based on the analysis of scientific literature, it is established that a significant part of existing approaches is focused on the use of large training samples, a fixed model structure, or predefined forecasting parameters. Insufficient elaboration has been identified regarding the adaptive selection of forecasting model parameters under conditions of a small amount of historical data and rapid changes in the local structure of a stock exchange time series. The expediency of using a polynomial approach, analyzing the sequence of polynomial forecast values PPS, and applying machine learning methods to improve the accuracy of short-term forecasting of stock exchange dynamics is substantiated.

The second chapter, “Methodological Foundations of the Polynomial Approach to the Analysis of Time Series of Stock Exchange Dynamics”, formalizes the problem of short-term forecasting of a stock exchange time series. A local fragment of the time series is considered as the basis for constructing a forecast, and the concept of the sequence of polynomial forecast values PPS is introduced. It is shown that PPS is formed as a set of forecasts obtained for different values of the parameter (m), which determines the depth of the local historical fragment and the corresponding degree of polynomial extrapolation.

The transition from a single forecast value to a comprehensive analysis of the entire PPS sequence is substantiated. It is established that such a sequence reflects local properties of the time series, including convergence or divergence of forecast values,

the presence of local extrema, monotonic segments, and areas of concentration. The properties of PPS as a separate information object are investigated, which makes it possible to assess the sensitivity of the forecast to changes in the parameters of polynomial extrapolation and to form a basis for the adaptive selection of the forecasting model.

The third chapter, “Methods of Adaptive Forecasting of Time Series of Stock Exchange Dynamics Based on the Sequence of Polynomial Forecasts”, develops and improves methods of adaptive forecast formation based on intelligent analysis of PPS. Approaches to selecting the optimal value of the parameter (m) or an informative subset of polynomial forecast values are proposed. Methods based on the analysis of local extrema of PPS, detection of areas of concentration of forecast values, clustering, DBSCAN analysis, and averaging of selected elements of the sequence are considered.

Special attention is paid to the procedure for forming the final forecast value based on the internal characteristics of PPS without involving additional external factors. This approach increases the applicability of the developed methods under conditions of limited data, when complex multifactor models may be ineffective or unstable. The limits of applicability of the proposed methods are defined, in particular cases of rapid PPS divergence, absence of pronounced areas of concentration, or an unstable local structure of forecast values.

The fourth chapter, “Development of an Original Neural Network Architecture for Forecasting Stock Exchange Time Series Based on a Polynomial Approach”, proposes an original neural network architecture, PPS-Net, focused on the use of the sequence of polynomial forecast values. The architecture includes a special PPS formation block, a neural network block for analyzing its structure, and a block for adaptive weighting of polynomial forecasts for different values of the parameter (m).

Unlike traditional neural network models, the proposed architecture uses not only the initial values of a local fragment of the time series but also a set of polynomially calculated forecast values. The final forecast is formed as a weighted combination of PPS elements, which makes it possible to combine the analytical advantages of polynomial extrapolation with the learning capability of a neural

network. The significance weights of polynomial forecasts provide additional interpretability of the model, since they allow determining which PPS elements have the greatest influence on the formation of the forecasting result.

The fifth chapter, “Experimental Study of the Efficiency of the Developed Methods”, forms the scientific and applied foundations for evaluating the proposed methods and neural network architecture. The experimental study is carried out on model, stochastic, and real time series of stock exchange dynamics. To verify the efficiency of the methods, walk-forward validation and one-step-ahead forecasting procedures are used, which correspond to real conditions of the sequential arrival of stock exchange data.

Forecasting accuracy is evaluated using the MAE, MSE, RMSE, and MAPE metrics. A comparison is conducted between the proposed polynomial methods, adaptive PPS analysis algorithms, the clustering approach, DBSCAN analysis, and the PPS-Net neural network architecture, on the one hand, and classical statistical, machine learning, and neural network forecasting methods, on the other. It is shown that the analysis of the PPS structure, the presence of clusters of forecast values, local extrema, and the weights of the neural network architecture makes it possible not only to evaluate the numerical forecasting error but also to interpret the behavior of the model under different modes of stock exchange dynamics.

The results of the dissertation research, including the proposed methods of adaptive polynomial extrapolation, PPS analysis algorithms, and the PPS-Net neural network architecture, can be used in the development of software tools for intelligent analysis of financial time series, investment decision support systems, modules for short-term forecasting of stock exchange quotations, as well as in the educational process when teaching disciplines related to machine learning, data analysis, intelligent systems, and time series forecasting.

The practical value of the work lies in the creation of methodological and algorithmic tools that provide adaptive forecast formation based on local information of a time series, make it possible to work with limited samples, and increase the interpretability of forecasting results. The proposed approaches may form a scientific

basis for the further development of information technologies for financial data processing, intelligent analysis of stock exchange dynamics, and the construction of hybrid forecasting models that combine mathematical formalization, polynomial extrapolation, and neural network learning.

The main scientific results of the dissertation have been published in 13 papers, including 6 articles in scientific professional journals of Ukraine and 7 publications in the proceedings of international and all-Ukrainian scientific and practical conferences.

Keywords: information technologies, intelligent data analysis, time series, stock exchange dynamics, financial time series, short-term forecasting, polynomial approach, polynomial extrapolation, Polynomial Predictions Sequence, PPS, adaptive forecasting, machine learning, neural networks, PPS-Net, cluster analysis, DBSCAN, walk-forward validation, one-step-ahead forecasting.

Список публікацій здобувача за темою дисертації

Статті у фахових наукових виданнях України

1. Турбал Ю., Турбал М., Кубай О., Смірнов Д., Мельничук М. (2025) Метод прогнозування на основі усереднення поліноміальної екстраполяційної послідовності. Моделювання, керування та інформаційні технології. № 8. С. 326–329. DOI: 10.31713/MCIT.2025.102. *(0,4 / 0,1 д.а.* авторський внесок: *розробка алгоритму та чисельний експеримент).*

2. Турбал Ю., Кубай О. (2025) Теорема Ковера та задача відокремлення множин за допомогою поліноміальних порогових функцій (PTF – Polynomial Threshold Functions). Вісник Національного університету водного господарства та природокористування. Технічні науки. №2. С. 178–191. DOI: https://doi.org/10.31713/vt2202516. *(0,58 / 0,49 д.а.;* авторський внесок: *ідея застосування нелінійного відображення координати 3-вимірного простору у 6-вимірний простір, в якому квадратична роздільність відповідає лінійній роздільності в просторі ознак (теорема Ковера та трюк ядра (kernel trick), чисельний експеримент).*

3. Turbal Y., Kubai O. (2026) A forecasting method based on clustering the polynomial extrapolation sequence. Computer Systems and Information Technologies. No. 2. P. 48–60. DOI: 10.31891/csit-2026-2-5. *(1,27 / 0,96 д.а.;* авторський внесок: *ідея методу прогнозування часових рядів на основі кластеризації поліноміальних екстраполяційних послідовностей, застосування алгоритму, чисельний експеримент).*

4. Турбал Ю., Кубай О. (2026). Про особливості використання послідовностей поліноміальних прогнозів у алгоритмах інтелектуального аналізу даних за умов існування збіжності. Вимірювальна та обчислювальна техніка в технологічних процесах, (2), С. 431–441. https://doi.org/10.31891/2219-9365-2026-86-51. *(0,93 / 0,4 д.а.;* авторський внесок: *розробка алгоритму, чисельний експеримент).*

5. Турбал Ю., Кубай, О. (2026). Аналіз ступеня стохастичності процесів на основі послідовності поліноміальних прогнозів (PPS). Комп’ютерно-

інтегровані технології: освіта, наука, виробництво, (63), 277-286. https://doi.org/10.36910/6775-2524-0560-2026-63-31. *(0,72 / 0,58 д.а.;* авторський внесок: *ідея кількісно оцінювати перехід від переважно детермінованої до поведінки процесу з ознаками стохастичності, розробка алгоритмів, чисельний експеримент).*

6. Турбал Ю., Кубай О. (2026) Архітектура спеціалізованої нейронної мережі для прогнозування фінансових часових рядів на основі послідовності поліноміальних прогнозів. Вісник Національного університету водного господарства та природокористування. Технічні науки. №1. С. 365–384. DOI: https://doi.org/10.31713/vt2202516. *(0,84 / 0,63 д.а.;* авторський внесок: *ідея інтеграції послідовності поліноміальних прогнозів у нейромережеву селекторну архітектуру, використання структурних ознак PPS і регуляризованого знаходження ваг для адаптивного вибору підсумкового прогнозу, розробка алгоритмів, чисельний експеримент).*

Публікації в матеріалах конференції, тези доповідей

1. Кубай О.В Особливості цифровізації правових відносин в умовах воєнного стану. Матеріали регіональної науково-практичної конференції «Актуальні проблеми національного адміністративного права в умовах воєнного стану», 2.03.2023 р. *(0,1 д.а).*

2. Кубай О.В. Роль та використання інформаційно-правових цифрових сервісів в умовах воєнного стану. Матеріали регіональної науково-практичної конференції «Актуальні проблеми національного адміністративного права в умовах воєнного стану», 2.03.2023 р. *(0,1 д.а).*

3. Кубай О.В. Використання штучних нейронних мереж в бізнес процесах: передумови, проблеми, перспективи Міжнародна науково-практична конференція молодих науковців, аспірантів і здобувачів вищої освіти «Проблеми та перспективи розвитку сучасної науки», 11-12 травня 2023 р, м. Рівне, 532 с. *(0,12 д.а).*

4. Кубай О. Регулювання використання штучного інтелекту в США як модель для України. Системи та засоби штучного інтелекту: тези доповідей

Міжнародної наукової конференції «Штучний інтелект: досягнення, виклики та ризики». – Київ: ІПШІ «Наука і освіта», 15-16.03.2024. – 550 с. *(0,1 д.а).*

5. Кубай О.В. Верифікація та валідація різнотипних даних як ключові аспекти для їх подальшої автоматизованої обробки. Міжнародна науково-практична конференція молодих науковців, аспірантів і здобувачів вищої освіти «Проблеми та перспективи розвитку сучасної науки», 9-10 травня 2024 р, м. Рівне. *(0,11 д.а).*

ЗМІСТ

ПЕРЕЛІК УМОВНИХ СКОРОЧЕНЬ

ACF	Autocorrelation Function	Автокореляційна функція
ADF	Augmented Dickey-Fuller Test	Розширений тест Дікі – Фуллера
AI	Artificial Intelligence	Штучний інтелект
AIC	Akaike Information Criterion	Інформаційний критерій Акаіке
ANN	Artificial Neural Network	Штучна нейронна мережа
AR	Autoregressive Model	Авторегресійна модель
ARCH	Autoregressive Conditional Heteroskedasticity	Авторегресійна умовна гетероскедастичність
ARIMA	Autoregressive Integrated Moving Average	Авторегресійна інтегрована модель ковзного середнього
ARMA	Autoregressive Moving Average	Авторегресійна модель ковзного середнього
BIC	Bayesian Information Criterion	Байєсівський інформаційний критерій
CNN	Convolutional Neural Network	Згорткова нейронна мережа
DBSCAN	Density-Based Spatial Clustering of Applications with Noise	Просторова кластеризація застосувань із шумом на основі щільності
DL	Deep Learning	Глибоке навчання
ETS	Error-Trend-Seasonal Model	Модель помилки, тренду та сезонності
GARCH	Generalized Autoregressive Conditional Heteroskedasticity	Узагальнена авторегресійна умовна гетероскедастичність
GBM	Gradient Boosting Machine	Машина градієнтного бустингу
GRU	Gated Recurrent Unit	Рекурентний блок із вентильним механізмом
LSTM	Long Short-Term Memory	Нейронна мережа з довгою короткочасною пам’яттю
MA	Moving Average Model	Модель ковзного середнього
MAE	Mean Absolute Error	Середня абсолютна похибка
MAPE	Mean Absolute Percentage Error	Середня абсолютна відсоткова похибка
ML	Machine Learning	Машинне навчання
MLP	Multilayer Perceptron	Багатошаровий перцептрон
MSE	Mean Squared Error	Середньоквадратична похибка
N-BEATS	Neural Basis Expansion Analysis for Time Series Forecasting	Нейронний базисний розклад для прогнозування часових рядів
OHLC	Open, High, Low, Close	Ціна відкриття, максимум, мінімум і ціна закриття
OHLCV	Open, High, Low, Close, Volume	Ціна відкриття, максимум, мінімум, ціна закриття та обсяг торгів

PACF	Partial Autocorrelation Function	Часткова автокореляційна функція
PPS	Polynomial Prediction Sequence	Послідовність поліноміальних прогнозних значень
PPS-Net	Polynomial Prediction Sequence Network	Нейромережева архітектура на основі послідовності поліноміальних прогнозних значень
RF	Random Forest	Випадковий ліс
RMSE	Root Mean Squared Error	Корінь із середньоквадратичної похибки
RNN	Recurrent Neural Network	Рекурентна нейронна мережа
SARIMA	Seasonal Autoregressive Integrated Moving Average	Сезонна авторегресійна інтегрована модель ковзного середнього
sMAPE	Symmetric Mean Absolute Percentage Error	Симетрична середня абсолютна відсоткова похибка
SVM	Support Vector Machine	Метод опорних векторів
SVR	Support Vector Regression	Регресія на основі опорних векторів
TFT	Temporal Fusion Transformer	Темпоральний трансформер злиття ознак
XGBoost	Extreme Gradient Boosting	Екстремальний градієнтний бустинг

ВСТУП

Актуальність теми дослідження. Сучасний етап розвитку інформаційних технологій, цифрової економіки та фінансових ринків характеризується істотним зростанням обсягів даних, швидкості їх оновлення та складності процесів, що підлягають аналізу. Особливої актуальності набувають задачі інтелектуального аналізу та прогнозування часових рядів біржової динаміки, оскільки фінансові часові ряди відображають складні, недетерміновані, волатильні та шумові процеси, пов’язані з функціонуванням фондових, валютних, товарних та інших біржових ринків. Ефективне прогнозування таких рядів є важливою складовою підтримки прийняття рішень в інвестиційній діяльності, алгоритмічній торгівлі, управлінні ризиками, фінансовому моніторингу та побудові аналітичних інформаційних систем.

Часові ряди біржової динаміки мають низку специфічних властивостей, які ускладнюють побудову точних прогнозних моделей. До таких властивостей належать висока мінливість значень, наявність випадкових коливань, короткочасних трендів, локальних екстремумів, змін волатильності, нестаціонарності та можливих структурних зсувів. Окремої уваги потребують внутрішньоденні біржові дані, для яких характерними є обмежена довжина локального фрагмента, висока чутливість до короткочасних змін ринку та необхідність формування прогнозу в режимі послідовного надходження нової інформації. У таких умовах класичні статистичні методи, машинне навчання та нейромережеві підходи не завжди забезпечують достатню точність прогнозування, особливо за обмеженого обсягу історичних даних або за відсутності стабільної глобальної структури ряду.

Значний внесок у розвиток теорії та практики аналізу часових рядів, прогнозування фінансової динаміки, машинного навчання та нейромережевого моделювання здійснено низкою провідних зарубіжних науковців. Зокрема, фундаментальні основи класичного статистичного аналізу та лінійного моделювання послідовностей закладено у працях Дж. Бокса (G. Box) та Г. Дженкінса (G. Jenkins). Розвиток методів дослідження волатильності та

нелінійних фінансових процесів пов'язаний із моделями авторегресійної умовної гетероскедастичності, запропонованими лауреатом Нобелівської премії Р. Енглом (R. Engle) та Т. Боллерслевом (T. Bollerslev), а також економетричними працями Дж. Гамільтона (J. Hamilton) та Р. Цея (R. Tsay). Сучасні концепції статистичного прогнозування систематизовано у дослідженнях Р. Гайндмана (R. Hyndman) та Г. Атанасопулоса (G. Athanasopoulos). Фундамент класичного машинного навчання та некореляційних ансамблів сформовано завдяки працям В. Вапника (V. Vapnik) та Л. Бреймана (L. Breiman). Вагомий прорив у сфері глибокого навчання та побудови рекурентних архітектур із керованою пам'яттю забезпечили З. Гохрайтер (S. Hochreiter), Ю. Шмідгубер (J. Schmidhuber) та А. Грейвс (A. Graves), тоді як еволюція сучасних увагових та інтерпретованих моделей часового об'єднання представлена фундаментальними розробками А. Васвані (A. Vaswani), Б. Ліма (B. Lim) та інших дослідників. Вагомими для розвитку відповідного наукового напряму є також праці українських учених, зокрема О. Івахненка, В. Степашка, П. Бідюка, М. Згуровського, В. Вітлінського, О. Черняка, А. Матвійчука та інших, у яких досліджено питання інтелектуального аналізу даних, економіко-математичного моделювання, прогнозування складних процесів, групового методу обробки даних, машинного навчання та моделювання економічної динаміки.

Аналіз наукових праць свідчить, що сучасна теорія прогнозування часових рядів біржової динаміки охоплює широкий спектр підходів: класичні статистичні моделі, регресійні методи, методи машинного навчання, глибокі нейронні мережі, рекурентні та attention-based архітектури. Водночас значна частина наявних методів орієнтована на використання достатньо великих навчальних вибірок, фіксованої структури моделі або наперед заданих параметрів прогнозування. Недостатньо опрацьованими залишаються питання побудови адаптивних методів прогнозування, які здатні ефективно використовувати локальну інформацію часового ряду, враховувати зміну його структури та автоматизовано обирати параметри прогнозної моделі в умовах обмеженого обсягу даних.

У цьому контексті перспективним є застосування поліноміального підходу, який дає змогу будувати прогноз на основі локального фрагмента часового ряду та досліджувати зміну прогнозного значення залежно від параметрів поліноміальної екстраполяції. Особливого значення набуває не лише побудова одиничного прогнозу за фіксованого степеня полінома, а й формування та аналіз послідовності поліноміальних прогнозних значень PPS (Polynomial Predictions Sequence). Така послідовність може розглядатися як окремий інформаційний об’єкт, що відображає локальну структуру біржового часового ряду, зокрема збіжність, розбіжність, наявність екстремумів, монотонних ділянок та областей згущення прогнозних значень.

Отже, актуальним науковим завданням є вдосконалення методів інтелектуального аналізу та прогнозування часових рядів біржової динаміки шляхом розроблення адаптивних алгоритмів на основі поліноміального підходу, аналізу послідовності PPS та побудови нейромережевої архітектури, здатної поєднувати аналітичні властивості поліноміальної екстраполяції з навчальною здатністю методів машинного навчання. Зазначене обумовлює актуальність теми дисертаційної роботи, її теоретичну та практичну значущість.

Зв’язок роботи з науковими програмами, планами, темами. Результати дисертаційного дослідження використовувались при виконанні наукової теми кафедри комп’ютерних наук та прикладної математики Національного університету водного господарства та природокористування «Математичне та комп’ютерне моделювання процесів та систем» (номер державної реєстрації 0125U001313), де автором запропоновано архітектуру нейронної мережі для прогнозування часових рядів біржової динаміки, яка характеризується введенням спеціального блоку формування послідовності поліноміальних прогнозних значень, блоку нейромережевого аналізу її структури та блоку адаптивного зважування поліноміальних прогнозів. Це дає можливість використовувати не лише початкові значення вихідного часового ряду, а й множину поліноміально обчислених прогнозних значень, забезпечуючи таким чином інтерпретоване формування прогнозу.

Крім того, результати досліджень використовувалися при виконанні науково-дослідних робіт в науково-дослідній частині Національного університету водного господарства та природокористування, а саме: «Вдосконалення інформаційних технологій обробки даних на основі оптимізованих комп'ютерних систем» (номер державної реєстрації 0121U111219), де автором удосконалено інформаційну технологію інтелектуальної обробки та прогнозування часових рядів біржової динаміки. Запропонований підхід передбачає формування послідовності поліноміальних прогнозних значень, аналіз її структурних властивостей, виділення інформативної підмножини прогнозів, адаптивний вибір оптимальної кількості її елементів та обчислення результуючого прогнозного значення. Використання розроблених алгоритмів дає змогу підвищити ефективність обробки фінансових даних у комп'ютерних системах шляхом автоматизованого налаштування параметрів поліноміальної екстраполяції відповідно до локальної поведінки часового ряду, зменшення впливу випадкових коливань і підвищення обґрунтованості вибору прогнозної моделі в умовах невизначеності та високої волатильності біржових процесів.

Мета і завдання дослідження. Метою дисертаційної роботи є вдосконалення методів інтелектуального аналізу часових рядів біржової динаміки шляхом розроблення нових алгоритмів прогнозування фінансових часових рядів на основі поліноміального підходу, що забезпечують підвищення точності короткострокового прогнозування в умовах обмеженого обсягу даних та недетермінованості біржових процесів.

Відповідно до окресленої мети в роботі визначено сукупність завдань, спрямованих на її досягнення:

− провести системне дослідження сучасного стану теоретичних засад інтелектуального аналізу та прогнозування часових рядів біржової динаміки;

− формалізувати методичні підходи до прогнозування часових рядів біржової динаміки на основі поліноміальної екстраполяції, ввести поняття

послідовності поліноміальних прогнозних значень PPS та дослідити її властивості як окремого інформаційного об’єкта;

− удосконалити сучасні методи адаптивного прогнозування часових рядів біржової динаміки на основі інтелектуального аналізу послідовності поліноміальних прогнозних значень PPS;

− розробити власну архітектуру нейронної мережі для прогнозування часових рядів біржової динаміки на основі поліноміального підходу;

− сформувати науково-прикладні засади оцінювання запропонованих методів і нейромережевої архітектури на модельних, стохастичних та реальних часових рядах біржової динаміки, а також виконати порівняння з класичними статистичними, машинними та нейромережевими методами прогнозування.

Об’єктом дослідження є процеси інтелектуального аналізу, обробки та прогнозування часових рядів біржової динаміки.

Предмет дослідження – моделі, методи та алгоритми інтелектуального аналізу часових рядів біржової динаміки на основі поліноміальної екстраполяції, адаптивного вибору параметрів прогнозування та аналізу послідовності поліноміальних прогнозних значень.

Методи дослідження. Теоретичну та методичну основу дисертаційної роботи становлять методи системного та порівняльного аналізу, які використано для обґрунтування актуальності дослідження, аналізу сучасного стану задачі прогнозування біржових часових рядів і постановки наукового завдання дисертаційної роботи. Методи математичного моделювання часових рядів застосовано для формалізації біржової динаміки та опису локального фрагмента ряду як основи для побудови прогнозу.

Методи інтерполяції та екстраполяції функцій, апарат скінченних різниць і поліноміального моделювання використано для побудови послідовності поліноміальних прогнозних значень PPS та дослідження залежності прогнозного результату від параметра поліноміальної екстраполяції. Методи інтелектуального аналізу даних і машинного навчання застосовано для виявлення закономірностей

у структурі PPS, аналізу локальних екстремумів, областей згущення прогнозних значень та адаптивного вибору інформативної підмножини прогнозів.

Методи кластерного аналізу, зокрема DBSCAN-аналіз, використано для виявлення груп поліноміальних прогнозних значень у PPS. Методи нейромережевого моделювання застосовано для розроблення архітектури PPS-орієнтованої нейронної мережі, у якій спеціальний блок формування PPS поєднується з блоком нейромережевого аналізу її структури та блоком адаптивного зважування поліноміальних прогнозів. Методи статистичного та експериментального аналізу використано для оцінювання точності та меж застосовності запропонованих алгоритмів. Для кількісного оцінювання результатів застосовано метрики MAE, MSE, RMSE і MAPE, а також процедури walk-forward перевірки та one-step-ahead прогнозування.

Інформаційною базою дослідження стали наукові праці вітчизняних і зарубіжних учених з питань аналізу часових рядів, прогнозування фінансової динаміки, поліноміальної екстраполяції, машинного навчання та нейромережевого моделювання; модельні детерміновані часові ряди; стохастичні ряди з різними типами випадкових збурень; реальні часові ряди біржової динаміки, зокрема внутрішньоденні дані біржових котирувань із фіксацією значень ціни закриття. Обчислювальні експерименти виконувалися в середовищі Python із використанням бібліотек для обробки даних, математичного моделювання, машинного навчання та візуалізації результатів.

Наукова новизна одержаних результатів полягає у вдосконаленні теоретичних, методичних і прикладних засад інтелектуального аналізу та прогнозування часових рядів біржової динаміки на основі поліноміального підходу. Найбільш вагомими результатами, які характеризують наукову новизну дисертаційної роботи та відображають особистий внесок автора, є такі:

вперше:

− запропоновано архітектуру нейронної мережі для прогнозування часових рядів біржової динаміки, яка характеризується введенням спеціального блоку формування послідовності поліноміальних прогнозних значень PPS, блоку

нейромережевого аналізу її структури та блоку адаптивного зважування поліноміальних прогнозів. На відміну від традиційних нейромережевих підходів, запропонована архітектура використовує не лише початкові значення локального фрагмента часового ряду, а й множину поліноміально обчислених прогнозних значень, що забезпечує інтерпретоване формування прогнозу через ваги значущості для різних значень параметра *m*;

удосконалено:

− методи адаптивного прогнозування часових рядів біржової динаміки, які базуються на побудові послідовності PPS, аналізі її структурних властивостей, виборі оптимального значення параметра m або інформативної підмножини поліноміальних прогнозів та формуванні остаточного прогнозного значення. На відміну від підходів із наперед заданим степенем полінома, запропоновані методи дають змогу адаптивно визначати параметри поліноміальної екстраполяції відповідно до локальної поведінки біржового часового ряду;

− науково-прикладні засади оцінювання методів прогнозування часових рядів біржової динаміки, які поєднують walk-forward перевірку, one-step-ahead прогнозування, аналіз точності за класичними метриками MAE, MSE, RMSE і MAPE, а також дослідження структури PPS у різних режимах детермінованості та стохастичності. На відміну від традиційного оцінювання лише за числовими показниками похибки, запропонований підхід враховує наявність кластерів PPS, характер розподілу поліноміальних прогнозних значень і поведінку ваг нейромережевої архітектури;

набули подальшого розвитку:

− теоретичні засади інтелектуального аналізу та прогнозування часових рядів біржової динаміки, які відрізняються поєднанням поліноміальної екстраполяції, аналізу послідовності поліноміальних прогнозних значень PPS та методів машинного навчання. Це дало змогу сформувати наукову основу для розроблення адаптивних методів прогнозування, орієнтованих на використання локальної інформації часового ряду в умовах обмеженої вибірки;

− формалізація методичних підходів до прогнозування часових рядів біржової динаміки на основі поліноміальної екстраполяції, що відрізняється переходом від побудови одиничного прогнозу за фіксованого степеня полінома до комплексного аналізу послідовності PPS як окремого інформаційного об’єкта. Це забезпечує можливість досліджувати збіжність, розбіжність, екстремуми, монотонні ділянки та області згущення прогнозних значень, а також обґрунтовано вибирати параметр m для короткострокового прогнозування біржової динаміки.

Практичне значення одержаних результатів полягає в можливості використання запропонованих методів, алгоритмів і нейромережевої архітектури для розв’язання задач короткострокового прогнозування часових рядів біржової динаміки в умовах обмеженого обсягу вхідних даних, високої волатильності та недетермінованості фінансових процесів. Розроблені підходи можуть бути застосовані в інформаційних системах підтримки прийняття інвестиційних рішень, системах фінансового моніторингу, програмних засобах аналізу біржових котирувань, навчальному процесі під час викладання дисциплін, пов’язаних з аналізом даних, машинним навчанням, інтелектуальними системами та прогнозуванням часових рядів.

Практична цінність роботи полягає також у можливості інтерпретації результатів прогнозування через аналіз структури PPS та ваг поліноміальних прогнозів у запропонованій нейромережевій архітектурі. Це дає змогу не лише отримувати числовий прогноз, а й оцінювати внесок окремих поліноміальних прогнозних значень у формування остаточного результату, виявляти випадки швидкої розбіжності PPS, відсутності виражених областей згущення або зниження надійності прогнозної моделі.

Результати дисертаційного дослідження впроваджено у діяльність ТОВ «ВІВА СОФТ» під час розроблення та вдосконалення програмних засобів аналізу й прогнозування часових рядів.

Особистий внесок здобувача. Усі наукові результати, представлені в дисертаційній роботі, одержані автором особисто. Автором виконано системний

аналіз сучасних методів прогнозування часових рядів біржової динаміки, формалізовано поліноміальний підхід до побудови послідовності прогнозних значень PPS, розроблено адаптивні методи вибору параметрів прогнозування, запропоновано нейромережеву архітектуру на основі PPS, виконано програмну реалізацію алгоритмів та проведено експериментальне дослідження їх ефективності. Із наукових праць, опублікованих у співавторстві, у дисертації використано лише ті положення, ідеї та висновки, які належать автору особисто.

Апробація результатів дисертації. Основні положення та результати дисертаційної роботи обговорювалися й доповідалися на науково-практичних конференціях, семінарах і наукових заходах, присвячених проблемам комп'ютерних наук, інтелектуального аналізу даних, машинного навчання, інформаційних технологій та прогнозування часових рядів. Зокрема:

Міжнародна науково-практична конференція молодих науковців, аспірантів і здобувачів вищої освіти «Проблеми та перспективи розвитку сучасної науки», 11-12 травня 2023 р, м. Рівне, Міжнародна наукова конференція «Штучний інтелект: досягнення, виклики та ризики», 15-16.03.2024 р, м. Київ, Міжнародна науково-практична конференція молодих науковців, аспірантів і здобувачів вищої освіти "Проблеми та перспективи розвитку сучасної науки», 9-10 травня 2024 р, м. Рівне, Міжнародна конференція Modelling, Control &Information Technologies. НУВГП, м. Рівне, 6-7.11.2025 р.

Публікації. За результатами дисертаційного дослідження опубліковано 11 наукових праць, з яких 6 статей у наукових фахових виданнях України, та 5 тез доповідей у матеріалах міжнародних і всеукраїнських науково-практичних конференцій. Загальний обсяг публікацій – 5,263 д.а., в тому числі авторський внесок здобувача – 3,69 д.а.

Структура та обсяг дисертації. Дисертація складається зі вступу, п'яти розділів, висновків, списку використаних джерел і додатків. У першому розділі досліджено теоретичні засади інтелектуального аналізу та прогнозування часових рядів біржової динаміки. У другому розділі формалізовано методичні основи поліноміального підходу та введено поняття послідовності

поліноміальних прогнозних значень PPS. У третьому розділі розроблено методи адаптивного прогнозування на основі аналізу структури PPS. У четвертому розділі запропоновано нейромережеву архітектуру прогнозування часових рядів біржової динаміки на основі поліноміального підходу. У п’ятому розділі наведено результати експериментального дослідження ефективності запропонованих методів на модельних, стохастичних та реальних часових рядах біржової динаміки.

Загальний обсяг дисертації становить 206 сторінок. Робота містить 7 таблиць, 22 рисунки, 7 додатків. Список використаних джерел налічує 167 найменування.

РОЗДІЛ 1. ТЕОРЕТИЧНІ ЗАСАДИ ІНТЕЛЕКТУАЛЬНОГО АНАЛІЗУ ТА ПРОГНОЗУВАННЯ ЧАСОВИХ РЯДІВ БІРЖОВОЇ ДИНАМІКИ

1.1. Часові ряди біржової динаміки як об'єкт інтелектуального аналізу

Функціонування фінансових ринків супроводжується безперервною зміною вартості активів, обсягів операцій, інтенсивності торгів, співвідношення попиту і пропозиції, рівня ліквідності та очікувань учасників ринку. Сукупність таких змін доцільно розглядати як біржову динаміку – впорядкований в часі процес формування та зміни кількісних характеристик біржового ринку під впливом інформаційних, економічних, поведінкових і мікроструктурних чинників.

У вузькому розумінні біржова динаміка може бути представлена зміною ціни окремого фінансового інструмента. У ширшому розумінні вона охоплює множину взаємопов'язаних показників: ціни відкриття, максимуму, мінімуму і закриття інтервалу; найкращі ціни купівлі та продажу; величину спреду; обсяг торгів; кількість укладених угод; показники ліквідності; біржові індекси; похідні індикатори технічного аналізу; характеристики волатильності; параметри книги заявок тощо.

Відповідно до підходу А. Мадгаван (A. Madhavan), мікроструктура ринку характеризує процес, у межах якого приховані наміри учасників щодо купівлі та продажу активів трансформуються у спостережувані ціни, обсяги та послідовності угод [17]. Отже, біржові котирування не є ізольованими числовими значеннями: вони відображають складний механізм взаємодії учасників ринку та інформаційного середовища.

У загальному випадку часовий ряд біржових даних можна подати у вигляді впорядкованої послідовності

$$X = \mathbf{x}(t_i)_{(i=1)}^{N} \tag{1.1}$$

де t_i – момент спостереження або межа часового інтервалу; $\mathbf{x}(t_i)$ – вектор характеристик ринку в момент часу t_i; N – кількість доступних спостережень.

Для агрегованих біржових даних, сформованих у вигляді інтервальних свічок, вектор спостережень може бути визначений як

$$\mathbf{x}(t_i) = (O_i, H_i, L_i, C_i, V_i)$$

де O_i, H_i, L_i, C_i – відповідно ціни відкриття, максимуму, мінімуму та закриття інтервалу; V_i – обсяг торгів. Такий формат часто позначають абревіатурою OHLCV.

У задачах прогнозування можуть використовуватися як абсолютні значення ціни P_t, так і логарифмічні ціни $p_t = \ln P_t$ та логарифмічні дохідності

$$r_t = p_t - p_{t-1} = ln\left(\frac{P_t}{P_{t-1}}\right). \quad (1.2)$$

Перехід від рівнів цін до дохідностей дає змогу аналізувати відносні зміни вартості активів, порівнювати ряди з різними ціновими масштабами та частково зменшувати вплив трендової компоненти. Водночас вибір цільової змінної залежить від постановки задачі. Об’єктом прогнозування можуть бути майбутня ціна, дохідність, напрям зміни котирування, волатильність, імовірність локального екстремуму або клас ринкового стану.

У роботі основну увагу зосереджено на прогнозуванні значень біржового часового ряду на основі обмеженого локального фрагмента спостережень. Така постановка відповідає необхідності формування прогнозу в умовах швидкої зміни ринкового середовища та обмеженої інформативності віддалених історичних даних.

Фінансові часові ряди суттєво відрізняються від детермінованих сигналів і багатьох класичних статистичних процесів. Їм притаманні нерегулярність, підвищена мінливість, залежність від часових масштабів, зміна статистичних характеристик, наявність екстремальних коливань і складна взаємодія закономірних та випадкових складових.

Систематизацію найбільш поширених емпіричних властивостей фінансових рядів здійснено у праці Р. Конта (R. Cont) [1]. Автор узагальнив результати численних досліджень і виокремив сукупність стилізованих фактів – характеристик, які регулярно спостерігаються для різних фінансових інструментів, ринків і часових масштабів. До них належать важкі хвости розподілів дохідностей, слабка лінійна автокореляція дохідностей, кластеризація

волатильності, залежність характеристик від масштабу агрегування та нелінійні залежності між абсолютними або квадратичними значеннями дохідностей.

Водночас стилізовані факти не слід інтерпретувати як універсальні закони, що однаково проявляються для кожного активу, періоду та способу формування вибірки. У сучасному дослідженні Е. Ретліффа-Крейна, С. М. Ван Оорта, М. Т. К. Келера, Б. Ф. Тівнана, (E. Ratliff-Crain, C. M. Van Oort, M. T. K. Koehler і B. F. Tivnan) повторно перевірили властивості, описані Р. Контом, на внутрішньоденних даних акцій, що входили до промислового індексу Доу-Джонса (Dow Jones Industrial Average) [2]. Автори підтвердили вісім із одинадцяти досліджуваних властивостей. Це свідчить про необхідність емпіричної перевірки припущень для конкретного набору даних, частоти спостережень і прогнозного горизонту.

Однією з основних особливостей біржових часових рядів є високий рівень шуму. Зміна ціни формується під впливом значної кількості чинників, частина яких не може бути безпосередньо спостережена або формалізована: новинних повідомлень, макроекономічних очікувань, поведінкових реакцій учасників, алгоритмічних стратегій, змін ліквідності, великих заявок і випадкових локальних дисбалансів попиту та пропозиції.

У спрощеному вигляді спостережуване значення біржового ряду можна представити як

$$y_t = s_t + \varepsilon_t \,, \tag{1.3}$$

де s_t – умовна інформаційна складова, що відображає локальні закономірності процесу; ε_t – випадкова або слабко передбачувана складова.

Для високочастотних даних додатково враховують мікроструктурний шум. У такому разі спостережувану ціну можна подати як

$$P_t^{(\text{obs})} = P_t^* + \eta_t, \tag{1.4}$$

де P_t^* – латентна «ефективна» ціна активу; η_t – шум, зумовлений механізмом торгів. До джерел такого шуму належать коливання між цінами попиту та пропозиції, дискретність кроку ціни, асинхронність угод, відмінності

ліквідності, повторення котирувань і нерівномірність надходження транзакцій [22; 23].

П. Р. Гансен і А. Лунде (P. R. Hansen, A. Lunde) показали, що за підвищення частоти спостережень вплив мікроструктурного шуму може істотно ускладнювати оцінювання реалізованої дисперсії [22]. Я. Аїт-Сагалія та Дж. Ю (Y. Aït-Sahalia, J. Yu) встановили зв'язок між величиною шуму й ліквідністю фінансового інструмента: для більш ліквідних акцій співвідношення шумової та інформаційної складових, як правило, є нижчим [23]. Таким чином, збільшення кількості внутрішньоденних спостережень не завжди означає пропорційне зростання обсягу корисної інформації.

Ранні емпіричні дослідження біржових рядів засвідчили, що розподіли дохідностей не завжди можуть бути адекватно описані нормальним законом. У праці Б. Б. Мандельброта (B. B. Mandelbrot) показано, що зміни спекулятивних цін характеризуються значною кількістю великих відхилень, які трапляються частіше, ніж передбачає гаусівська модель [3]. Е. Е. Фама (E. F. Fama) також виявив істотні відмінності емпіричних розподілів дохідностей від нормального розподілу [4].

Така властивість отримала назву важких хвостів розподілу. Її практичне значення полягає в тому, що екстремальні коливання не можна розглядати лише як виняткові випадки. Вони є важливою складовою реальної біржової динаміки та мають ураховуватися під час побудови прогнозних алгоритмів, оцінювання похибок і дослідження стійкості моделей.

Волатильність характеризує інтенсивність коливань вартості фінансового активу. На відміну від випадкового шуму з постійною дисперсією, фінансові ряди демонструють часову мінливість амплітуди коливань: періоди відносно спокійної поведінки змінюються періодами різких цінових рухів. Високі за абсолютною величиною зміни часто групуються поблизу інших великих змін, а низькі – поблизу малих. Це явище називають кластеризацією волатильності.

Формалізація змінної умовної дисперсії стала основою ARCH-моделі (Autoregressive Conditional Heteroskedasticity – авторегресивна умовна

гетероскедастичність), запропонованої Р. Ф. Енглом (R. F. Engle) [6]. Подальшим розвитком цього підходу стала узагальнена модель авторегресивної умовної гетероскедастичності GARCH-модель (Generalized Autoregressive Conditional Heteroskedasticity) Т. Боллерслева (T. Bollerslev) [7], у якій поточна умовна дисперсія залежить як від попередніх похибок, так і від попередніх оцінок дисперсії. Огляд методів прогнозування волатильності та проблем їх порівняння наведено у праці С.-Х. Пун і К. В. Дж. Грейнджера (S.-H. Poon, C. W. J. Granger) [10].

Для внутрішньоденних даних важливим є також поняття реалізованої волатильності, яка оцінюється на основі сукупності високочастотних спостережень. Економетричні основи її застосування для оцінювання стохастичної волатильності розглянуто О. Е. Барндорфф-Нільсеном і Н. Шепардом (O. E. Barndorff-Nielsen, N. Shephard) [21].

Стаціонарний часовий ряд передбачає сталість основних статистичних характеристик процесу в часі. Для біржових даних це припущення часто виконується лише наближено та в межах обмежених інтервалів. Рівні цін можуть містити трендові зміни, тоді як розподіли дохідностей, дисперсії та кореляційні зв’язки здатні змінюватися залежно від ринкового стану.

Т. А. Шмітт, Д. Четалова, Р. Шефер і Т. Гур (T. A. Schmitt, D. Chetalova, R. Schäfer, T. Guhr) показали, що нестаціонарність є суттєвою властивістю фінансових рядів: дисперсії та кореляції можуть істотно залежати від обраного часового вікна, а зміни режимів впливають на форму емпіричних розподілів [11].

З позицій інтелектуального аналізу даних такі зміни можуть бути інтерпретовані як прояв дрейфу концепції (concept drift). У загальному випадку дрейф концепції означає зміну в часі залежностей між вхідними ознаками та цільовою змінною. Огляд методів виявлення та адаптації до таких змін наведено Ж. Гамою, І. Жліобайте, А. Біфетом, М. Пєченізкієм та А. Бушашією (J. Gama, I. Žliobaitė, A. Bifet, M. Pechenizkiy, A. Bouchachia) [12].

Нестаціонарність ускладнює прогнозування, оскільки закономірності, виявлені на віддаленій історичній вибірці, можуть втратити актуальність у

поточному ринковому режимі. Тому для короткострокового прогнозування доцільно досліджувати локальні фрагменти ряду та використовувати адаптивні алгоритми, параметри яких можуть змінюватися залежно від поточної структури даних.

Окремий напрям аналізу фінансових часових рядів пов’язаний із дослідженням масштабних властивостей, самоподібності, довготривалої пам’яті та мультифрактальності. У класичній праці Б. Мандельброта [3] було започатковано підхід до аналізу фінансових процесів із важкими хвостами та нетривіальною масштабною поведінкою. Подальший розвиток ці ідеї отримали в роботі Л. Е. Кальве та А. Дж. Фішера (L. E. Calvet, A. J. Fisher), у якій запропоновано мультифрактальний підхід до моделювання дохідностей активів [13].

Т. Ді Маттео, Т. Асте та М. М. Дакоронья (T. Di Matteo, T. Aste, M. M. Dacorogna) використали масштабний аналіз і узагальнений показник Герста для порівняння розвинених ринків і ринків, що формуються [14]. Результати таких досліджень свідчать, що характеристики фінансових рядів можуть залежати від масштабу спостереження та рівня розвитку ринку.

Для волатильності сучасним напрямом є концепція «нерегулярної» волатильності. Дж. Гатерал, Т. Жейссон і М. Розенбаум (J. Gatheral, T. Jaisson, M. Rosenbaum) показали, що логарифмічна волатильність на досліджуваних даних може поводитися подібно до фрактального броунівського руху з малим значенням показника Герста [15]. Водночас М. Фукасава, Т. Такабатаке і Р. Вестфаль (M. Fukasawa, T. Takabatake, R. Westphal) звернули увагу на необхідність обережного оцінювання ступеня нерегулярності латентної волатильності, оскільки результати можуть залежати від процедури її відновлення [16].

Отже, фрактальність доцільно трактувати не як безумовну властивість кожного біржового ряду, а як сукупність емпірично досліджуваних масштабних характеристик. Їх прояв залежить від активу, часового горизонту, частоти спостережень, способу агрегування даних і методу оцінювання.

Фінансові часові ряди можуть містити локальні інтервали спрямованого руху, фази корекції, короткочасні тренди, коливальні фрагменти та періоди відносної стабільності. Проте такі закономірності не обов'язково зберігаються протягом тривалого часу.

У працях Е. Фами розглянуто поведінку біржових цін у контексті випадкових блукань і ефективності ринку [4; 5]. Відповідно до цієї концепції, загальнодоступна інформація швидко враховується у вартості активів, а можливості систематичного прогнозування дохідності на основі минулих значень є обмеженими. Це не виключає наявності локальних залежностей, але вимагає обережності під час їх інтерпретації.

У контексті прогнозування доцільно виходити з того, що локальний тренд є тимчасовою структурною властивістю фрагмента ряду, а не сталою глобальною закономірністю. Саме тому прогнозний алгоритм повинен адаптивно визначати довжину релевантного часового вікна та складність моделі.

Внутрішньоденні дані містять спостереження, сформовані впродовж однієї або кількох торговельних сесій із часовим кроком, меншим за один день. Залежно від джерела та задачі аналізу використовують тикові дані, хвилинні котирування або агреговані інтервали тривалістю 5, 15, 30 чи 60 хвилин.

Тикові дані (tick data) є найбільш деталізованою формою подання інформації про біржову динаміку. На відміну від часових рядів із фіксованим кроком, новий елемент тикового ряду формується в момент настання ринкової події: укладення угоди, зміни котирування, додавання або скасування заявки. Тому такі спостереження розміщуються нерівномірно в часі. Тикові дані містять значний обсяг інформації про мікроструктуру ринку, однак характеризуються підвищеним рівнем шуму та потребують додаткової попередньої обробки. Для зменшення обчислювальної складності їх часто агрегують у хвилинні, 15-хвилинні або 30-хвилинні інтервальні свічки формату OHLCV.

Внутрішньоденна біржова динаміка характеризується низкою специфічних властивостей. По-перше, інтенсивність торгів і волатильність змінюються протягом сесії. Р. А. Вуд, Т. Г. Мак-Ініш і Дж. К. Орд (R. A. Wood, T. H. McInish,

J. K. Ord) встановили, що для даних Нью-Йоркської фондової біржі підвищена мінливість спостерігається на початку та наприкінці торговельного дня [18]. Т. Г. Андерсен і Т. Боллерслев (T. G. Andersen) показали, що внутрішньоденна періодичність істотно впливає на динамічні властивості високочастотних дохідностей і має враховуватися під час моделювання волатильності [17]. П. К. Джейн і Г.-Х. Джо (P. C. Jain, G.-H. Joh) виявили відмінності обсягів торгів і дохідностей залежно від години торговельної сесії та дня тижня [20].

По-друге, необхідно розрізняти календарний час і час подій. У календарному поданні спостереження формуються через рівні проміжки часу. У подієвому поданні новий елемент ряду з'являється після кожної угоди, зміни котирування або накопичення визначеної кількості транзакцій. Е. Ретліфф-Крейн (E. Ratliff-Crain) та співавтори встановили, що прояв окремих стилізованих фактів може залежати від того, чи аналізуються дані в календарному або подієвому часі [2].

По-третє, межі торговельної сесії створюють специфічні розриви. Ціна відкриття може суттєво відрізнятися від ціни закриття попереднього дня через надходження інформації в неробочий час. Тому об'єднання декількох сесій у єдиний часовий ряд потребує окремого врахування нічних розривів.

По-четверте, якість внутрішньоденних даних залежить від ліквідності активу. Для низьколіквідних інструментів можуть виникати повторювані значення котирувань, пропущені інтервали, нерівномірні обсяги торгів і підвищений вплив спреду. У високочастотному режимі додатковим джерелом похибки стає мікроструктурний шум [22; 23].

По-п'яте, підготовка внутрішньоденних даних повинна враховувати календарні та технічні особливості. До основних процедур попередньої обробки належать:

1) упорядкування спостережень за часовими мітками;
2) вилучення дублікатів і перевірка коректності часових інтервалів;
3) виявлення пропущених значень;

4) відокремлення основної торговельної сесії від премаркету та післяринкових торгів;
5) урахування вихідних, святкових днів і змін часових поясів;
6) коригування цін у разі дроблення акцій та інших корпоративних подій;
7) формування фіксованої часової сітки;
8) перевірка впливу нічних розривів;
9) збереження хронологічного порядку під час поділу вибірки на навчальну й тестову частини.

Таким чином, внутрішньоденні дані містять значний обсяг інформації про локальну поведінку ринку, але водночас характеризуються високим рівнем шуму, внутрішньосесійною сезонністю та залежністю від режиму торгів. Це потребує застосування методів, здатних працювати з короткими локальними фрагментами ряду та адаптуватися до зміни його структури.

Для багатьох методів машинного навчання збільшення обсягу навчальної вибірки є необхідною умовою підвищення точності та стійкості моделей. Однак у задачах прогнозування біржової динаміки використання довгої історії не завжди забезпечує покращення результату. Внаслідок нестаціонарності ринкового процесу віддалені спостереження можуть відповідати іншому ринковому режиму та втрачати релевантність для поточного моменту.

Виникає суперечність між двома вимогами:

- збільшення обсягу вибірки зменшує статистичну невизначеність оцінювання параметрів;
- скорочення локального часового вікна підвищує актуальність даних для поточного режиму ринку.

Ця суперечність особливо помітна для внутрішньоденного прогнозування, коли рішення необхідно формувати на основі невеликої кількості останніх спостережень. Крім того, через залежність сусідніх значень і мікроструктурний шум номінальна кількість елементів ряду може перевищувати його ефективний інформаційний обсяг.

Дж. Бонтемпі, С. Бен Таєб і Й.-А. Ле Борн (G. Bontempi, S. Ben Taieb, Y.-A. Le Borgne) розглянули часовий ряд як послідовність задач навчання з учителем і показали доцільність локальних стратегій прогнозування, орієнтованих на часову структуру даних [24]. Р. П. Мазіні, М. К. Медейрос і Е. Ф. Мендес (R. P. Masini, M. C. Medeiros, E. F. Mendes) систематизували сучасні методи машинного навчання для часових рядів, зокрема нелінійні моделі, ансамблі та нейронні мережі [26]. Водночас складні моделі з великою кількістю параметрів потребують достатнього обсягу репрезентативних даних і належного контролю перенавчання.

Окремою проблемою є вибір процедури експериментальної перевірки. Випадкове перемішування елементів фінансового часового ряду може призвести до потрапляння інформації з майбутнього до навчальної вибірки. В. Серкейра, Л. Торго та І. Мозетич (V. Cerqueira, L. Torgo, I. Mozetič) показали важливість методів оцінювання, які зберігають часовий порядок спостережень, особливо для нестаціонарних рядів [25]. Для однокрокового прогнозування доцільно використовувати послідовну схему walk-forward validation, за якої на кожному кроці прогноз формується лише на основі доступної на цей момент історії.

За умов малого обсягу даних перспективними є моделі з обмеженою кількістю параметрів, зрозумілою структурою та можливістю адаптивного налаштування. До таких моделей належить локальна поліноміальна екстраполяція. Її застосування не передбачає припущення про існування глобальної поліноміальної закономірності біржового процесу. Поліном у цьому разі використовується як локальна апроксимаційна модель для опису короткого фрагмента ряду.

Нехай для формування прогнозу використовуються останні m значень

$$y_{n-m+1}, y_{n-m+2}, \ldots, y_n.$$

На їх основі будується локальна поліноміальна модель степеня d, після чого визначається прогноз $y_{n+1}^{(m,d)}$. Якість такого прогнозу залежить від вибору довжини вікна m, степеня полінома d, рівня шуму та локальної структури даних. Надто коротке вікно підвищує чутливість до випадкових коливань. Надто довге вікно

може охоплювати застарілі спостереження або декілька різних режимів. Використання полінома високого степеня збільшує ризик неефективної екстраполяції.

Отже, ключове значення має не лише побудова окремого прогнозу, а й аналіз сукупності прогнозних значень, отриманих для різних параметрів локальної моделі. Такий підхід створює основу для формування послідовності поліноміальних прогнозів (Polynomial Predictions Sequence, PPS), дослідження її структури та адаптивного вибору інформативної підмножини прогнозних значень.

1.2. Класичні статистичні методи прогнозування фінансових часових рядів

Прогнозування часових рядів біржової динаміки є складовою задачею інтелектуального аналізу фінансових даних. Класичні статистичні методи формують теоретичну основу розв’язання цієї задачі та залишаються важливим інструментом порівняльного оцінювання нових алгоритмів. До цієї групи належать наївні прогнозні моделі, методи ковзного середнього та експоненціального згладжування, авторегресійні моделі, моделі авторегресії – ковзного середнього ARMA (Autoregressive Moving Average) та авторегресії – інтегрованого ковзного середнього ARIMA (Autoregressive Integrated Moving Average), моделі умовної гетероскедастичності, багатовимірні економетричні моделі, моделі простору станів і моделі з перемиканням режимів.

Методи експоненціального згладжування та ARIMA належать до найбільш поширених класичних підходів до прогнозування часових рядів. Водночас вони ґрунтуються на різних принципах: експоненціальне згладжування описує локальний рівень, трендову та сезонну складові ряду, тоді як ARIMA-моделі відтворюють структуру автокореляційних залежностей між його значеннями [27].

Особливістю застосування статистичних методів до біржових даних є необхідність урахування нестаціонарності котирувань, мінливості волатильності, наявності структурних зрушень, короткочасних трендів і значного

рівня випадкових коливань. У праці Р. Конта систематизовано характерні емпіричні властивості фінансових часових рядів, зокрема відсутність сталої лінійної автокореляції дохідностей, важкі хвости розподілів, кластеризацію волатильності та залежність статистичних властивостей від часового масштабу [1].

Нехай P_t ціна фінансового інструменту в момент часу *t*. У практичних задачах прогнозування можуть використовуватися безпосередньо ціни, абсолютні прирости цін або дохідності. Логарифмічна дохідність визначається формулою (1.2)

Перехід від цін до дохідностей є поширеним етапом попередньої обробки фінансових даних. Цінові ряди зазвичай демонструють стохастичний тренд і нестаціонарну поведінку, тоді як послідовності дохідностей часто є ближчими до стаціонарних процесів. Це спрощує застосування авторегресійних моделей та аналіз умовної дисперсії [36; 5].

Водночас вибір прогнозованої величини залежить від прикладної мети дослідження. Якщо потрібно оцінити напрям зміни ринку або ризик, доцільно моделювати дохідності. Якщо метою є отримання безпосереднього прогнозного значення котирування на наступному часовому кроці, як у межах цієї дисертаційної роботи, прогнозування може виконуватися для послідовності цін P_t із подальшим оцінюванням похибки у вихідних одиницях вимірювання.

Найпростішою базовою моделлю короткострокового прогнозування біржового котирування є наївна модель, відповідно до якої наступне значення ряду дорівнює останньому відомому значенню. Незважаючи на простоту, наївна модель має принципове значення під час оцінювання складніших алгоритмів. Вона пов’язана з концепцією випадкового блукання та гіпотезою ефективного ринку, відповідно до якої поточна ціна відображає доступну інформацію, а систематичне отримання точного прогнозу лише на основі попередніх котирувань є ускладненим [38]. У роботі Ю. Фами ефективним названо ринок, ціни на якому відображають доступну інформацію; ця концепція стала одним із фундаментальних теоретичних орієнтирів аналізу фінансових часових рядів.

Наївний прогноз доцільно використовувати як обов’язковий базовий орієнтир. Новий алгоритм короткострокового прогнозування є практично обґрунтованим лише тоді, коли він демонструє стабільне зменшення похибки порівняно з прогнозом останнього відомого значення.

Одним із найпростіших способів згладжування випадкових коливань є ковзне середнє. Цей метод зменшує вплив локального шуму, однак має суттєве обмеження: усі значення в межах вибраного вікна отримують однакову вагу. Збільшення ширини вікна посилює згладжування, але водночас зменшує чутливість моделі до короткочасних змін динаміки. Для внутрішньоденних біржових даних це може спричинити запізнення прогнозу під час зміни локального тренду.

На відміну від звичайного ковзного середнього, методи експоненціального згладжування надають більшу вагу останнім спостереженням. Ваги попередніх значень зменшуються експоненціально в міру їх віддалення від поточного моменту часу [27].

Метод Голта-Вінтерса є подальшим розвитком експоненціального згладжування та додатково враховує сезонну складову. У роботі П. Вінтерса розглянуто моделі прогнозування з експоненціально спадними вагами та наведено приклади їх застосування [37].

Для біржових даних сезонність може бути пов’язана не лише з календарними періодами, але й із внутрішньоденною структурою торгової активності: початком сесії, періодами підвищеної ліквідності, виходом новин або наближенням закриття торгів. Однак на короткому одноденному фрагменті даних статистично надійне оцінювання сезонної компоненти часто є неможливим через відсутність достатньої кількості повторюваних циклів.

Авторегресійний підхід ґрунтується на припущенні, що поточне значення ряду може бути описане лінійною комбінацією його попередніх значень. Модель авторегресії порядку p, або AR(p), визначається рівнянням

$$P_t = c + \sum_{i=1}^{p} \varphi_i P_{t-i} + \varepsilon_t, \tag{1.5}$$

де c – стала; φ_i – параметри авторегресії; ε_t – випадкова похибка, яка в найпростішому випадку розглядається як білий шум.

Модель ковзного середнього похибок MA(q) описує залежність поточного значення від попередніх випадкових збурень:

$$P_t = \mu + \varepsilon_t + \sum_{j=1}^{q} \theta_j \varepsilon_{t-j}, \tag{1.6}$$

де μ – середнє значення ряду; θ_j – параметри моделі.

Поєднання авторегресійної та MA-компонент утворює модель ARMA(p, q):

$$P_t = c + \sum_{i=1}^{p} \varphi_i P_{t-i} + \varepsilon_t + \sum_{j=1}^{q} \theta_j \varepsilon_{t-j}, \tag{1.7}$$

Методологію побудови, ідентифікації та перевірки ARMA-моделей систематизовано у працях Дж. Бокса, Г. Дженкінса, Г. Райнзела та Г. Люнг [28]. Їхній підхід передбачає послідовне виконання трьох основних етапів: ідентифікації структури моделі, оцінювання параметрів і діагностичної перевірки залишків.

Порядки p та q можуть попередньо визначатися за вибірковими автокореляційною та частковою автокореляційною функціями. Після оцінювання параметрів потрібно перевірити, чи не збереглася істотна автокореляція залишків. Наявність структурованих залишків свідчить про те, що модель не повністю відтворює закономірності ряду.

Основним обмеженням ARMA-моделей є вимога стаціонарності досліджуваного процесу. Безпосереднє використання біржових цін часто не відповідає цій вимозі. Крім того, на коротких вибірках оцінки коефіцієнтів можуть бути нестійкими, особливо якщо порядок моделі є надмірно великим.

Для прогнозування нестаціонарних часових рядів застосовують інтегровані моделі авторегресії та ковзного середнього ARIMA(p, d, q). Їхньою ключовою особливістю є попереднє диференціювання ряду. Оператор першої різниці визначається як

$$\Delta P_t = P_t - P_{t-1}, \tag{1.8}$$

а різниця порядку d – як

$$\Delta^d P_t = (1-B)^d P_t, \tag{1.9}$$

де B – оператор зсуву назад: $BP_t = P_{t-1}$.

Загальний запис моделі ARIMA(p, d, q) має вигляд

$$\Phi(B)(1-B)^d P_t = c + \Theta(B)\varepsilon_t, \tag{1.10}$$

де

$$\Phi(B) = 1 - \varphi_1 B - \varphi_2 B^2 - \ldots - \varphi_p B^p,$$
$$\Theta(B) = 1 + \theta_1 B + \theta_2 B^2 + \ldots + \theta_p B^p,$$

Моделі ARIMA описують автокореляційні зв'язки у стаціонарному ряді, отриманому після диференціювання. Цей клас моделей є одним із базових інструментів статистичного прогнозування [27; 28].

За наявності сталої сезонності застосовується сезонна модель авторегресії – інтегрованого ковзного середнього SARIMA (Seasonal Autoregressive Integrated Moving Average):

$$\Phi_P(B^s)\Phi_p(B)(1-B^s)^D(1-B)^d P_t = \Theta_Q(B^s)\Theta_q(B)\varepsilon_t, \tag{1.11}$$

де s – довжина сезонного періоду; D – порядок сезонного диференціювання; P і Q – порядки сезонних авторегресійної та МА-компонент.

Перед побудовою ARIMA-моделей доцільно перевірити ряд на стаціонарність. Одним із найбільш відомих інструментів є розширений тест Дікі-Фуллера (ADF). Його регресійне рівняння може бути записане у формі

$$\Delta P_t = a + bt + \gamma P_{t-1} + \sum_{i=1}^{k} \delta_i \Delta P_{t-i} + \varepsilon_t, \tag{1.12}$$

У межах розширеного тесту Дікі-Фуллера (ADF-тесту – Augmented Dickey-Fuller test) перевіряється нульова гіпотеза про наявність одиничного кореня. Основи такого підходу закладено у праці Д. Дікі та В. Фуллера [38].

Доповнювальним інструментом є тест KPSS (Kwiatkowski, Phillips, Schmidt, Shin), запропонований Д. Квятковським, П. Філліпсом, П. Шмідтом і Й. Шином (D. Kwiatkowski, P. Phillips, P. Schmidt, Y. Shin). На відміну від ADF-тесту, його нульова гіпотеза відповідає стаціонарності ряду відносно детермінованого

тренду. [39]. Використання обох тестів дає змогу отримати більш обґрунтований висновок щодо характеру нестаціонарності.

Для вибору порядків ARIMA-моделі застосовуються інформаційні критерії. Критерій Акаіке визначається формулою

$$AIC = -2\ln L + 2k, \tag{1.13}$$

а байєсівський інформаційний критерій Шварца – формулою

$$BIC = -2\ln L + k\ln n, \tag{1.14}$$

де L – максимальне значення функції правдоподібності; k – кількість оцінюваних параметрів; n – кількість спостережень.

Критерій Акаіке спрямований на пошук компромісу між точністю опису даних і складністю моделі [40]. Критерій Шварца сильніше штрафує надмірну кількість параметрів, особливо зі збільшенням вибірки [41].

Автоматизований алгоритм Гайндмана-Хандакара (Hyndman-Khandakar) поєднує тести на одиничні корені, мінімізацію скоригованого інформаційного критерію Акаіке (AIC_C) та оцінювання параметрів методом максимальної правдоподібності [42].

Для внутрішньоденних біржових даних застосування ARIMA має низку обмежень. Структура ряду може змінюватися впродовж однієї торгової сесії, а порядок моделі, оптимальний для одного локального фрагмента, не обов’язково залишається ефективним на наступному фрагменті. На малих вибірках автоматизований вибір параметрів також може бути неефективним через недостатню кількість спостережень.

Моделі ARMA та ARIMA орієнтовані переважно на прогнозування умовного середнього значення процесу. Однак для фінансових часових рядів важливим є також моделювання волатильності. Однією з характерних властивостей біржових даних є кластеризація волатильності: періоди значних коливань мають тенденцію групуватися, як і періоди відносно стабільної динаміки [28].

Модель ARCH дає змогу описувати змінну в часі дисперсію похибок, що є важливим для фінансових застосувань [43].

Узагальненням ARCH є модель GARCH(*p, q*), запропонована Т. Боллерслевом:

На відміну від ARCH, модель GARCH ураховує не лише попередні квадрати похибок, але й попередні значення умовної дисперсії. Це дає змогу описувати триваліший вплив волатильних періодів [44].

Серед подальших модифікацій варто виокремити експоненціальну модель EGARCH (Exponential GARCH) та модель GJR-GARCH (Glosten-Jagannathan-Runkle GARCH), запропоновану Л. Глостеном, Р. Джаганнатханом і Д. Ранклом (L. Glosten, R. Jagannathan, D. Runkle) У моделі EGARCH, запропонованій Д. Нельсоном, враховано можливість асиметричного впливу позитивних і негативних новин на волатильність [45]. Модель GJR-GARCH також дає змогу описувати різний вплив додатних і від'ємних інновацій [46].

Моделі сімейства GARCH є важливими для оцінювання ризику, побудови прогнозних інтервалів і моделювання мінливості фінансових процесів. Водночас вони не розв'язують безпосередньо задачу точкового прогнозування майбутньої ціни. Тому в межах дослідження поліноміальних методів їх доцільно розглядати як засоби аналізу волатильності та як можливі компоненти гібридних алгоритмів.

Якщо доступні додаткові фактори, що можуть впливати на біржову динаміку, застосовують динамічну регресію або моделі ARIMAX (Autoregressive Integrated Moving Average with Exogenous Inputs – авторегресія – інтегроване ковзне середнє з екзогенними змінними). До екзогенних змінних можуть належати обсяги торгів, значення ринкових індексів, курси валют, процентні ставки, макроекономічні показники або технічні індикатори.

Включення додаткової інформації може підвищити точність прогнозування, але ускладнює модель і потребує перевірки доступності екзогенних факторів у момент формування прогнозу. Використання майбутніх або фактично недоступних значень факторів спричиняє витік інформації та завищену оцінку якості алгоритму. Динамічні регресійні моделі як розширення ARIMA розглянуто у праці Р. Гайндмана та Дж. Атанасопулоса [27].

У випадку одночасного аналізу кількох пов’язаних фінансових показників застосовуються векторні авторегресійні моделі VAR (Vector Autoregressive models). Такі моделі дають змогу враховувати взаємний вплив кількох змінних без необхідності заздалегідь задавати жорстку структуру причинно-наслідкових зв’язків. Важливе значення для розвитку цього підходу мала праця К. Сімса [47].

Якщо нестаціонарні ряди мають довгостроковий рівноважний зв’язок, застосовуються векторні моделі корекції похибки VECM (Vector Error Correction Model). Теоретичні засади коінтеграції та моделей корекції похибки сформульовано Р. Енглом і К. Грейнджером [48].

Застосування VAR і VECM для коротких внутрішньоденних фрагментів має обмеження. Із збільшенням кількості змінних та лагів швидко зростає кількість параметрів. За малого обсягу вибірки це може призвести до перенавчання та неточності прогнозів.

Гнучким статистичним підходом до аналізу часових рядів є моделі простору станів. Вони дають змогу окремо описувати приховану динаміку процесу та механізм спостереження. Для лінійних моделей із гауссівськими похибками рекурентне оцінювання прихованого стану виконується за допомогою фільтра Калмана. У фундаментальній праці Р. Калмана запропоновано рекурентний підхід до задач фільтрації та прогнозування дискретних процесів [49].

Моделі простору станів дають змогу адаптивно оцінювати локальний рівень, тренд, сезонність і вплив зовнішніх змінних. Їх систематичний виклад наведено у праці Дж. Дурбіна та С.-Я. Купмана [50].

Для фінансових даних цей підхід є перспективним завдяки можливості оновлення оцінок після надходження кожного нового спостереження. Однак якість результатів залежить від правильності вибраної структури моделі, припущень щодо розподілу похибок і параметрів шуму.

Біржові процеси можуть характеризуватися чергуванням різних режимів: локального зростання, зниження, відносної стабільності або підвищеної

волатильності. Для опису таких змін застосовують моделі з марковським перемиканням режимів.

У праці Дж. Гамільтона запропоновано підхід, відповідно до якого параметри авторегресійної моделі залежать від прихованого дискретного марковського процесу. Це дає змогу описувати стрибкоподібні зміни режимів динаміки [51].

Застосування моделей із перемиканням режимів є теоретично обґрунтованим для фінансових рядів, але потребує достатньої кількості спостережень для оцінювання ймовірностей переходів. У коротких внутрішньоденних вибірках окремі режими можуть бути представлені лише кількома значеннями, що знижує надійність параметричних оцінок.

Порівняння прогнозних моделей має виконуватися на даних, які не використовувалися під час оцінювання параметрів. Точність прогнозу не можна обґрунтовано визначати лише за похибками апроксимації навчальної вибірки. Для часових рядів важливо зберігати хронологічний порядок даних, оскільки випадкове перемішування спостережень може спричинити витік інформації з майбутнього [27].

Для оцінювання точності прогнозів можуть використовуватися такі показники:

$$MAE = \frac{1}{n}\sum_{t=1}^{n}\left|P_t - \hat{P}_t\right|, \tag{1.15}$$

$$MSE = \frac{1}{n}\sum_{t=1}^{n}\left(P_t - \hat{P}_t\right)^2, \tag{1.16}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{t=1}^{n}\left(P_t - \hat{P}_t\right)^2}, \tag{1.17}$$

$$MAPE = \frac{100\%}{n}\sum_{t=1}^{n}\left|\frac{P_t - \hat{P}_t}{P_t}\right|. \tag{1.18}$$

Показник *MAE* (Mean Absolute Error – середня абсолютна похибка) характеризує середню абсолютну похибку в одиницях вимірювання ціни. Показник *RMSE* (Root Mean Squared Error – середньоквадратична похибка)

сильніше реагує на великі відхилення прогнозу. Показник *MAPE* (Mean Absolute Percentage Error – середня абсолютна відсоткова похибка) спрощує порівняння результатів для активів із різними рівнями цін, але його використання є проблемним, якщо фактичні значення можуть наближатися до нуля.

Для короткострокового прогнозування доцільним є підхід walk-forward validation. На кожному кроці модель використовує лише значення, доступні до моменту t, формує прогноз $\hat{P}_{t+h|t}$, після чого до вибірки додається фактичне значення P_{t+1}. Такий підхід відповідає реальним умовам надходження біржових даних і дає змогу оцінити ефективність алгоритму до послідовних змін локальної динаміки.

Основні властивості розглянутих підходів узагальнено в додатку А.

1.3. Методи машинного навчання для прогнозування біржової динаміки

Класичні статистичні методи прогнозування часових рядів, розглянуті в підрозділі 1.2, ґрунтуються на формалізації певної структури досліджуваного процесу: трендової, авторегресійної, сезонної або стохастичної. Їх ефективність значною мірою залежить від обґрунтованості прийнятих припущень щодо стаціонарності, виду розподілу випадкових складових, порядку моделі та характеру залежностей між спостереженнями.

Для біржових часових рядів виконання таких припущень не завжди може бути гарантоване. Фінансові дані характеризуються високим рівнем шуму, нестаціонарністю, зміною волатильності, локальними трендами, нелінійними залежностями та змінами ринкових режимів. Крім того, на поведінку цін впливають численні зовнішні чинники, частина яких не може бути безпосередньо врахована в межах однієї математичної моделі. У сучасних оглядових працях підкреслюється, що саме динамічність, нерегулярність і складність фінансових даних зумовили активне застосування алгоритмів машинного навчання для аналізу та прогнозування біржових процесів [67, 71–74].

Застосування методів машинного навчання не означає відмову від врахування часової природи даних. Навпаки, часові залежності мають бути

відображені у способі формування навчальної вибірки, побудові ознак і процедурі експериментальної перевірки. Принципова відмінність полягає в тому, що функціональна залежність між вхідними ознаками та прогнозованим значенням не задається повністю апріорно, а оцінюється на основі навчальних прикладів.

Відповідно до гіпотези ефективного ринку, сформульованої Ю. Фамою, поточні ціни активів відображають доступну учасникам ринку інформацію [5]. Це істотно ускладнює отримання стабільного прогнозного ефекту. Тому результати застосування алгоритмів машинного навчання доцільно інтерпретувати не як можливість детермінованого передбачення майбутньої ціни, а як засіб виявлення обмеженої локальної передбачуваності за визначених умов, на конкретному часовому горизонті та з обов'язковою перевіркою на даних, які не використовувалися під час навчання.

У роботі під методами машинного навчання будемо розуміти сукупність алгоритмів, які дають змогу на основі накопичених спостережень оцінювати залежність між характеристиками біржового процесу та його майбутніми значеннями або напрямками зміни. До таких методів належать регуляризовані регресійні моделі, методи найближчих сусідів, метод опорних векторів, дерева рішень, випадкові ліси, бустингові алгоритми, ансамблеві моделі, а також нейромережеві архітектури. Останні доцільно розглянути окремо, оскільки вони становлять самостійний напрям розвитку інтелектуальних методів прогнозування.

Формалізуємо задачу прогнозування засобами машинного навчання таким чином – нехай задано дискретний часовий ряд біржових котирувань: $P_1, P_2, \ldots, P_N$, де P_t – значення ціни активу або іншого фінансового показника в момент часу *t*, а *N* – кількість доступних спостережень.

У загальному випадку задача прогнозування полягає у побудові відображення

$$\hat{y}_{t+h} = f\left(\mathbf{x}_t; \boldsymbol{\theta}\right), \tag{1.19}$$

де $\hat{y}_{t+h}$ – прогнозоване значення цільової змінної; h – горизонт прогнозування; $\mathbf{x}_t$ – вектор вхідних ознак, сформований на основі інформації, доступної до моменту часу t; f – прогнозна модель; $\boldsymbol{\theta}$ – параметри моделі, які визначаються під час навчання.

Навчальна вибірка має вигляд

$$\mathcal{D}_N = \left\{ \left(\mathbf{x}_t, y_{t+h} \right) \right\}_{t=p}^{N-h}, \tag{1.20}$$

де p – кількість попередніх спостережень, що використовуються для формування вхідного вектора.

Залежно від постановки задачі цільова змінна y_{t+h} може відображати:

– майбутнє значення ціни P_{t+h};

– абсолютну зміну ціни $P_{t+h} - P_t$;

– відносну або логарифмічну дохідність;

– майбутнє значення волатильності;

– напрямок зміни ціни;

– належність майбутнього стану ринку до певного класу.

Для задачі класифікації напрямку руху ціни цільову змінну можна подати у вигляді:

$$d_{t+h} = \begin{cases} 1, & P_{t+h} > P_t, \\ 0, & P_{t+h} \le P_t. \end{cases}, \tag{1.21}$$

Отже, методи машинного навчання можуть застосовуватися як для регресійного прогнозування кількісних значень, так і для класифікації майбутніх станів біржового процесу. У практичних дослідженнях вибір постановки задачі залежить від мети аналізу: оцінювання точного майбутнього значення, визначення напрямку руху, виявлення аномалій або формування торговельного рішення.

Результативність алгоритму машинного навчання залежить не лише від вибору моделі, але й від якості вхідних ознак. У дослідженнях з прогнозування фондового ринку як вхідні змінні використовують історичні ціни, обсяги торгів, дохідності, показники волатильності, технічні індикатори, фундаментальні

характеристики компаній, макроекономічні змінні та текстові дані [64, 71, 75]. У роботі Н. Руфа та співавторів зазначено, що формування інформативного набору ознак є одним із ключових етапів побудови прогнозної моделі для фондового ринку [71].

Найпростіший вектор ознак для прогнозування на основі попередніх значень цін можна подати у вигляді

$$\mathbf{x}_t = \left(P_t, P_{t-1}, \ldots, P_{t-p+1} \right)^T . \tag{1.22}$$

Для розширеного набору ознак до вектора можуть бути включені значення обсягу торгів V_t і технічних індикаторів $I_{j,t}$:

$$\mathbf{x}_t = \left(P_t, P_{t-1}, \ldots, P_{t-p+1}, V_t, I_{1,t}, \ldots, I_{m,t} \right)^T . \tag{1.23}$$

Для алгоритмів, чутливих до масштабу ознак, зокрема методу найближчих сусідів, методу опорних векторів і регуляризованої регресії, доцільно виконувати стандартизацію:

$$z_{j,t} = \frac{x_{j,t} - \mu_j}{s_j}$$

де μ_j та s_j відповідно середнє значення та стандартне відхилення j-ї ознаки.

Під час експериментального дослідження параметри масштабування потрібно обчислювати лише за навчальною частиною вибірки. Використання майбутніх спостережень для нормалізації, вибору ознак або налаштування гіперпараметрів спричиняє витік інформації та призводить до завищеної оцінки точності моделі.

Для короткострокового прогнозування внутрішньоденних рядів особливого значення набуває проблема обмеженого обсягу локальних даних. Надмірне збільшення кількості ознак за невеликої кількості спостережень може спричинити перенавчання. У зв’язку з цим доцільно використовувати компактні набори ознак, регуляризацію параметрів, адаптивний вибір довжини вікна та процедури перевірки на послідовних часових інтервалах.

У низці прикладних задач достатньо визначити не точне значення майбутньої ціни, а напрямок її зміни. Такий підхід дає змогу сформувати класифікаційну модель, виходом якої є ймовірність зростання або зниження ціни.

Одним із базових методів є логістична регресія, перевагами якої є простота реалізації, інтерпретованість коефіцієнтів і можливість оцінювання ймовірності належності спостереження до певного класу. Її доцільно застосовувати як базову модель для порівняння з більш складними класифікаторами.

Іншим простим алгоритмом є наївний баєсівський класифікатор. Він ґрунтується на припущенні про умовну незалежність ознак за заданого класу. Хоча це припущення не завжди відповідає структурі фінансових даних, алгоритм може бути корисним як обчислювально простий еталонний метод.

У роботі М. Баллінгса та співавторів здійснено порівняння класифікаторів для прогнозування напрямку руху цін акцій, зокрема логістичної регресії, методу найближчих сусідів, методу опорних векторів, нейронних мереж та ансамблевих моделей [65]. Автори показали, що результати суттєво залежать від обраного алгоритму, набору ознак і протоколу оцінювання.

Метод *k* найближчих сусідів, або *k*-NN (*k*-Nearest Neighbors), ґрунтується на припущенні, що подібним станам досліджуваного процесу можуть відповідати подібні майбутні значення. Для нового вектора ознак $\mathbf{x}_t$ визначається множина $\mathcal{N}_k(\mathbf{x}_t)$, яка містить *k* найближчих навчальних прикладів.

Для задачі класифікації клас визначається більшістю голосів найближчих сусідів. Теоретичні засади методу було сформульовано Т. Ковером і П. Гартом [56].

Перевагою *k*-NN є відсутність складної процедури оцінювання параметрів моделі. Метод може бути корисним для виявлення локальних аналогій у поведінці біржового процесу. Водночас його ефективність суттєво залежить від способу масштабування ознак, вибору метрики відстані та значення параметра *k*. Зі збільшенням розмірності простору ознак відстані між об’єктами стають менш інформативними, що обмежує застосування алгоритму без попереднього відбору ознак або зменшення розмірності.

Метод опорних векторів був запропонований К. Кортесом і В. Вапніком [57]. Початково його розроблено для задач класифікації, однак згодом метод було адаптовано до регресійного аналізу у вигляді Support Vector Regression, або SVR.

Однією з ранніх праць, присвячених застосуванню методу опорних векторів SVM (Support Vector Machines) до прогнозування фінансових часових рядів, є дослідження К.-Дж. Кіма (K.-J. Kim) [58]. Автор порівняв метод опорних векторів із нейромережевою моделлю та методом міркування за аналогіями, розглядаючи прогнозування напрямку зміни фондового індексу. Надалі SVM і SVR стали поширеними алгоритмами для аналізу фінансових даних [67, 71, 74].

До переваг SVR належать можливість моделювання нелінійних залежностей і придатність для роботи з вибірками помірного обсягу. Обмеженнями є чутливість до масштабування даних і необхідність налаштування параметрів ядра, співвідношення між складністю моделі та величиною допустимих відхилень, ширини зони нечутливості. У разі часових рядів налаштування гіперпараметрів має виконуватися виключно на навчальних часових інтервалах.

Дерева рішень утворюють ієрархічну структуру правил поділу простору ознак на області, у межах яких формується прогноз. Теоретичні та алгоритмічні засади побудови дерев класифікації та регресії CART (Classification and Regression Trees) систематизовано Л. Брейманом, Дж. Фрідманом, Р. Олшеном і Ч. Стоуном (L. Breiman, J. Friedman, R. Olshen, C. Stone) у однойменній праці [59].

Перевагою дерев є можливість відображення нелінійних залежностей і взаємодій між ознаками. Крім того, дерева не потребують обов’язкової стандартизації змінних. Водночас окреме дерево може бути нестійким: незначні зміни навчальної вибірки здатні спричинити суттєву зміну його структури.

Для підвищення стійкості Л. Брейман запропонував метод випадкового лісу, або Random Forest [60]. У регресійній задачі прогноз випадкового лісу визначається усередненням прогнозів окремих дерев. Кожне дерево навчається на випадковій підвибірці спостережень, а під час побудови вузлів

використовується випадкова підмножина ознак. Такий підхід зменшує кореляцію між деревами та дає змогу знизити дисперсію прогнозу порівняно з окремим деревом [60].

Метод Random Forest неодноразово використовувався для прогнозування напрямку руху цін та аналізу біржових даних. У дослідженні М. Баллінгса та співавторів випадковий ліс продемонстрував високі результати серед розглянутих класифікаційних моделей [65]. У роботі К. Крауса, Х. А. До та Н. Гака (K. Kraus, X. A. Do, N. Huck) випадковий ліс порівнювався з градієнтним бустингом і глибокими нейронними мережами під час аналізу акцій фондового індексу S&P 500 (Standard & Poor's 500 Index) [66]. Ці результати не означають універсальної переваги Random Forest, але підтверджують доцільність його використання як одного з базових алгоритмів порівняння.

До обмежень випадкового лісу належать ускладнена інтерпретація ансамблю порівняно з окремим деревом, схильність до зниження точності за суттєвої зміни ринкового режиму та обмежені можливості екстраполяції за межі діапазону значень, представленого в навчальній вибірці.

Бустинг передбачає послідовну побудову ансамблю слабких моделей, кожна з яких спрямована на зменшення помилок попередніх елементів ансамблю. Однією з базових робіт у цьому напрямі є дослідження Й. Фройнда та Р. Шапіро (Y. Freund, R. Schapire), у якому обґрунтовано алгоритм AdaBoost (Adaptive Boosting) [61].

У задачах прогнозування фінансових показників поширення набули алгоритми градієнтного бустингу над деревами рішень. Однією з відомих реалізацій є XGBoost (eXtreme Gradient Boosting), запропонована Т. Ченом і К. Гестріном (T. Chen, C. Guestrin) [62]. Алгоритм передбачає регуляризацію складності дерев та ефективну організацію обчислень, що забезпечує його придатність для роботи з табличними даними.

До переваг бустингових моделей належать здатність враховувати нелінійні залежності, взаємодії між ознаками та неоднорідну структуру даних. Водночас збільшення глибини дерев і кількості ітерацій може спричинити перенавчання.

Для фінансових часових рядів це особливо небезпечно, оскільки модель може адаптуватися до випадкових коливань конкретного діапазону спостережень.

Жоден окремий алгоритм не може вважатися найкращим для всіх фінансових інструментів, часових масштабів і ринкових режимів. Порівняльні дослідження демонструють, що точність прогнозування залежить від властивостей вибірки, способу подання даних, набору ознак, горизонту прогнозування та процедури оцінювання [64-68, 71, 74]. У масштабному порівняльному дослідженні О. Бустоса, А. Помарес-Кімбаї та Р. Стелліана (O. Bustos, A. Pomares-Quimbaya, R. Stellian) розглянуто декілька алгоритмів на даних багатьох фондових ринків. Автори підкреслюють відсутність загальновизнаного універсального алгоритму, який стабільно переважав би альтернативи за різних умов [74].

Одним із напрямів підвищення точності прогнозування є формування ансамблю моделей. Ансамблювання дає змогу зменшити вплив специфічних помилок окремих моделей. Проте комбінування прогнозів не гарантує автоматичного підвищення точності. Якщо базові моделі мають подібні систематичні похибки або вагові коефіцієнти визначаються з використанням майбутніх даних, ансамбль може виявитися неефективним.

Гібридні моделі передбачають поєднання алгоритмів різних класів. Наприклад, статистична модель може використовуватися для виділення трендової складової, а алгоритм машинного навчання – для прогнозування залишків. Іншим варіантом є поєднання дерев рішень, нейронних мереж і моделей градієнтного бустингу. У роботі К. Крауса та співавторів досліджено ансамбль глибоких нейронних мереж, градієнтного бустингу та випадкового лісу [66]. Сучасні огляди також визначають гібридизацію як один із перспективних напрямів розвитку методів прогнозування фінансових часових рядів [71, 73].

Для поліноміального підходу, який досліджується у дисертації, ансамблевий принцип має особливе значення. Послідовність поліноміальних прогнозних значень можна розглядати як сукупність прогнозів, отриманих за різних параметрів локальної моделі. Аналіз внутрішньої структури цієї

послідовності та адаптивний вибір інформативної підмножини прогнозів створюють передумови для формування більш точного остаточного прогнозного значення.

Більшість алгоритмів прогнозування біржової динаміки належить до методів навчання з учителем, оскільки для кожного вхідного вектора задається відповідне цільове значення. Водночас методи навчання без учителя можуть використовуватися як допоміжні інструменти інтелектуального аналізу даних.

До таких методів належать:

1) кластерний аналіз;
2) метод головних компонент;
3) виявлення аномалій;
4) зменшення розмірності;
5) сегментація часових рядів;
6) виявлення ринкових режимів.

Кластеризація може бути використана для групування схожих станів ринку, виділення періодів підвищеної або зниженої волатильності та аналізу локальної структури прогнозних значень. Метод головних компонент дає змогу зменшити кількість взаємопов’язаних ознак і сформувати компактне представлення даних. У дослідженні С. Гу, Б. Келлі та Д. Сю серед кандидатних підходів до аналізу фінансових даних розглянуто як регуляризовані регресійні методи, так і алгоритми зменшення розмірності, дерева рішень, бустинг, випадкові ліси та нейронні мережі [68].

У контексті цієї дисертації особливо перспективним є застосування кластерного аналізу до послідовності поліноміальних прогнозних значень. За наявності групи близьких прогнозів можна виділити область локальної узгодженості та використати її для адаптивного формування остаточного прогнозу. Водночас відсутність вираженого кластера може свідчити про нестійкість екстраполяції та високий рівень стохастичності локального фрагмента часового ряду.

Водночас під час порівняння методів слід уникати некритичного припущення, що складніша модель обов’язково є точнішою. У дослідженні С. Гу, Б. Келлі та Д. Сю показано, що результативність алгоритмів залежить від співвідношення між складністю моделі, обсягом вибірки та рівнем корисного сигналу у фінансових даних [18]. За умов обмеженого обсягу навчальної вибірки складні архітектури можуть виявитися нестійкими та схильними до перенавчання.

Коректне оцінювання алгоритмів машинного навчання для часових рядів потребує збереження хронологічної послідовності даних. Випадкове перемішування спостережень може призвести до використання майбутньої інформації під час навчання та сформувати завищену оцінку якості прогнозу.

У роботі В. Серкейри, Л. Торго та І. Мозетича проаналізовано різні процедури оцінювання прогнозних моделей для часових рядів. Автори наголошують, що вибір процедури залежить від властивостей ряду, а для нестаціонарних даних доцільно застосовувати методи перевірки, які відтворюють реальне послідовне надходження нових спостережень [70].

Слід враховувати, що питання застосовності перехресної перевірки до часових рядів має нюансований характер. К. Бергмайр, Р. Гайндман і Б. Ку показали, що за певних умов крос-валідація може бути коректною для авторегресійних моделей [69]. Проте для внутрішньоденних біржових рядів зі змінами локальної структури безпечнішим базовим підходом є хронологічне тестування без перемішування спостережень.

Для регресійних моделей доцільно використовувати показники MAE, MSE, RMSE і MAPE, розглянуті в підрозділі 1.2. Для класифікаційних моделей додатково можуть застосовуватися точність (accuracy), прогностична точність (precision), повнота (recall), F_1-міра (F_1-score) та площа під ROC-кривою ROC-AUC (Receiver Operating Characteristic – Area Under the Curve).

Для оцінювання правильності прогнозу напрямку зміни можна використати показник Directional Accuracy:

$$DA = \frac{1}{n}\sum_{t=1}^{n}\mathbf{1}\left[\operatorname{sign}\left(\hat{P}_{t+h} - P_t\right) = \operatorname{sign}\left(P_{t+h} - P_t\right)\right], \quad (1.24)$$

де $\mathbf{1}[\bullet]$ – індикаторна функція.

Під час оцінювання моделей потрібно чітко розмежовувати навчальну, валідаційну та тестову вибірки. Тестова вибірка має використовуватися лише для підсумкового оцінювання. Оптимізація параметрів за тестовими даними перетворює їх на частину навчального процесу та позбавляє експеримент об'єктивності.

Узагальнену характеристику розглянутих алгоритмів наведено в додатку Б.

1.4. Нейромережеві методи прогнозування часових рядів

Штучні нейронні мережі становлять окремий клас методів інтелектуального аналізу часових рядів, орієнтованих на апроксимацію складних нелінійних залежностей між попередніми та майбутніми значеннями досліджуваного процесу. На відміну від класичних статистичних моделей, у яких структура залежностей задається наперед у формалізованому вигляді, нейромережеві моделі можуть адаптивно формувати внутрішні представлення даних у процесі навчання. Це дає змогу враховувати нелінійність, взаємодію ознак, локальні зміни тенденцій і складні часові залежності.

Застосування штучних нейронних мереж для прогнозування часових рядів розпочалося задовго до поширення сучасних методів глибокого навчання. У фундаментальному огляді Дж. Чжана, Б. Патуво та М. Ху (G. Zhang, B. Patuwo, M. Hu) систематизовано результати ранніх досліджень нейромережевого прогнозування та визначено проблеми вибору структури мережі, вхідних змінних, алгоритму навчання і способу оцінювання точності [76]. У сучасних оглядах Б. Ліма і С. Зорена (B. Lim, S. Zohren), а також О. Сезера, М. Гуделека та А. Озбайоглу (O. Sezer, M. Gudelek, A. Ozbayoglu) розглянуто розвиток цього напряму від класичних багатошарових перцептронів до рекурентних, згорткових, базованих на механізмах уваги (attention-based) та гібридних архітектур [82; 83].

У задачах аналізу біржової динаміки нейромережевий підхід застосовують для прогнозування:

- абсолютних значень котирувань;
- логарифмічної або відносної дохідності;
- напряму зміни ціни;
- волатильності;
- імовірності настання заданої ринкової події;
- параметрів розподілу майбутніх значень;
- сигналів для прийняття торговельних рішень.

Нехай P_t – ціна фінансового активу в момент часу *t*.

Якщо розв'язується задача класифікації напряму руху котирувань, цільову змінну можна визначити як

$$s_{t+1} = \begin{cases} 1, & r_{t+1} > 0, \\ 0, & r_{t+1} \le 0. \end{cases}, \tag{1.25}$$

У роботі основною є задача побудови алгоритму прогнозування наступного значення ряду біржової динаміки. Її можна формалізувати як побудову відображення

$$\hat{y}_{t+h} = F_\theta\left(X_t\right), \tag{1.26}$$

де $\hat{y}_{t+h}$ – прогноз на горизонт *h*; F_θ – нейромережева модель із параметрами θ; X_θ – вектор або матриця вхідних даних, сформованих на основі доступних даних процесу.

Для одновимірного часового ряду $y_1, y_2, \ldots, y_T$ вхідний вектор із локальним вікном довжини *L* має вигляд

$$X_t = \left[y_{t-L+1}, y_{t-L+2}, \ldots, y_t\right], \tag{1.27}$$

Під час багатовимірного аналізу біржової динаміки елементами X_t можуть бути ціни відкриття, максимуму, мінімуму та закриття, обсяги торгів, значення технічних індикаторів, характеристики волатильності, зовнішні макроекономічні змінні або текстові індикатори ринкових настроїв. Водночас збільшення кількості ознак не гарантує підвищення точності: для фінансових даних особливого значення набувають запобігання перенавчанню, контроль витоку майбутньої інформації та перевірка результатів за схемою walk-forward.

Найпростішою нейромережевою моделлю прогнозування є багатошаровий перцептрон – Multilayer Perceptron (MLP). У такій моделі локальне вікно спостережень розглядають як вхідний вектор фіксованої довжини, а прогноз обчислюють у результаті послідовного проходження сигналу через один або декілька прихованих шарів.

Вихід *j*-го нейрона шару *l* визначають за формулою

$$a_j^{(l)} = \varphi\left(\sum_{i=1}^{n_{l-1}} w_{ji}^{(l)} a_i^{(l-1)} + b_j^{(l)} \right), \tag{1.28}$$

де $a_i^{(l-1)}$ – виходи нейронів попереднього шару; $w_{ji}^{(l)}$ – вагові коефіцієнти; $b_j^{(l)}$ – зміщення; $\varphi(\cdot)$ – функція активації; n_{l-1} – кількість нейронів попереднього шару.

Для задачі прогнозування числового значення котирування або дохідності як функцію втрат часто використовують середню квадратичну похибку *MSE*. З урахуванням регуляризації задача навчання нейронної мережі набуває вигляду

$$\theta^* = \arg\min_\theta \left[\frac{1}{N} \sum_{t=1}^{N} \left(y_t - \hat{y}_t \right)^2 + \lambda \Omega(\theta) \right], \tag{1.29}$$

де $\Omega(\theta)$ – регуляризаційний доданок; λ – параметр, що визначає силу регуляризації.

Перевагами MLP є відносна простота реалізації, можливість апроксимації нелінійних функцій і помірні обчислювальні витрати. Однак звичайний багатошаровий перцептрон не має внутрішнього механізму збереження стану: часовий порядок враховується лише через розташування елементів у вхідному векторі. Тому модель потребує явного вибору довжини вікна *L*, а її здатність відтворювати довгі залежності обмежена.

Для коротких внутрішньоденних вибірок ця властивість може бути не лише недоліком, але й перевагою. Порівняно з громіздкими глибокими архітектурами MLP має меншу кількість параметрів і нижчий ризик перенавчання. На основі багатошарових перцептронів побудовано також спеціалізовані моделі N-BEATS (Neural Basis Expansion Analysis for Time Series) та N-HiTS (Neural Hierarchical

Interpolation for Time Series), у яких прогноз формують шляхом послідовного уточнення компонентів сигналу. [89; 103].

Для безпосереднього врахування послідовної природи часових рядів застосовують рекурентні нейронні мережі – Recurrent Neural Networks (RNN). На відміну від MLP, рекурентна мережа використовує прихований стан h_t, який залежить не лише від поточного вхідного вектора x_t, але й від попереднього стану h_{t-1}:

$$h_t = \varphi\left(W_x x_t + W_h h_{t-1} + b_h\right), \tag{1.30}$$

$$\hat{y}_{t+1} = W_y h_t + b_y \tag{1.31}$$

У формулах W_x, W_h, W_y – матриці ваг; b_h, b_y – вектори зміщень; x_t – вхідні характеристики в момент часу *t*; h_t – прихований стан, у якому накопичується інформація про попередню динаміку процесу.

Рекурентна структура теоретично дає змогу враховувати залежності довільної тривалості. Проте практичне навчання звичайних RNN ускладнюється проблемами зникнення та вибухового зростання градієнтів. І. Бенджіо, П. Сімард і П. Фрасконі (Y. Bengio, P. Simard, P. Frasconi) показали, що зі збільшенням часової відстані між пов’язаними подіями градієнтне навчання рекурентних мереж стає дедалі складнішим [77].

У дослідженні Х. Хевамалаге, К. Бергмаєра та К. Бандари (H. Hewamalage, C. Bergmeir, K. Bandara) підкреслено, що рекурентні моделі є конкурентоспроможними інструментами прогнозування, але не можуть розглядатися як універсальна заміна статистичних методів. Їхня результативність залежить від характеристик ряду, способу попередньої обробки, сезонності, обсягу навчальних даних і коректного підбору гіперпараметрів [84].

Для рядів біржової динаміки проблема навчання RNN посилюється через нестаціонарність, шум, короткочасність локальних закономірностей і зміну ринкових режимів. Тому на практиці частіше використовують удосконалені рекурентні моделі LSTM та GRU (Gated Recurrent Unit – керований рекурентний блок). Архітектуру довгої короткострокової пам'яті LSTM (Long Short-Term

Memory) запропонували З. Хохрайтер і Дж. Шміддхубер (S. Hochreiter, J. Schmidhuber) для подолання проблеми навчання довготривалих залежностей [78]. Основною особливістю LSTM є наявність стану комірки c_t та керувальних вентилів, які визначають, яку інформацію потрібно зберігати, забувати й передавати на вихід.

У поширеному варіанті LSTM обчислення виконують за такими співвідношеннями:

$$i_t = \sigma\left(W_i x_t + U_i h_{t-1} + b_i\right), \tag{1.32}$$

$$f_t = \sigma\left(W_f x_t + U_f h_{t-1} + b_f\right) \tag{1.33}$$

$$o_t = \sigma\left(W_o x_t + U_o h_{t-1} + b_o\right) \tag{1.34}$$

$$\tilde{c}_t = \tanh\left(W_c x_t + U_c h_{t-1} + b_c\right) \tag{1.35}$$

$$c_t = f_t \odot c_{t-1} + i_t \odot \tilde{c}_t \tag{1.36}$$

$$h_t = o_t \odot \tanh(c_t) \tag{1.37}$$

де i_t – вхідний вентиль; f_t – вентиль забування; o_t – вихідний вентиль; $\tilde{c}_t$ – кандидат на оновлення стану комірки; $\sigma(\cdot)$ – сигмоїдна функція; $\odot$ – покомпонентне множення.

LSTM стала однією з найпоширеніших архітектур прогнозування фінансових часових рядів. Т. Фішер і К. Краусс (T. Fischer, C. Krauss) застосували LSTM для прогнозування напрямів зміни котирувань компонентів індексу S&P 500 та порівняли її з логістичною регресією, випадковим лісом і глибокою нейронною мережею прямого поширення (без пам’яті) [10]. Автори встановили перевагу LSTM у досліджуваній експериментальній постановці, але також зауважили зменшення надлишкової дохідності стратегії після врахування транзакційних витрат у пізніші періоди. Д. Нельсон, А. Перейра та Р. де Олівейра (D. Nelson, A. Pereira, R. de Oliveira) дослідили застосування LSTM для прогнозування напряму руху фондового ринку на основі історії цін і технічних індикаторів [86].

Практичними перевагами LSTM є здатність враховувати залежності різної тривалості та відсутність потреби вручну задавати конкретну форму функціональної залежності. Водночас використання LSTM для коротких внутрішньоденних рядів потребує обережності: велика кількість параметрів підвищує ризик перенавчання, а приховані стани мережі ускладнюють інтерпретацію причин отриманого прогнозу.

Спрощеною альтернативою LSTM є мережі з керованими рекурентними блоками GRU (Gated Recurrent Unit), запропоновані К. Чо (K. Cho) та його співавторами [79]. GRU використовує меншу кількість вентилів і не містить окремого стану комірки c_t. Це зменшує кількість параметрів і може пришвидшити навчання.

Основні співвідношення GRU мають вигляд

$$z_t = \sigma\left(W_z x_t + U_z h_{t-1} + b_z\right), \tag{1.38}$$

$$r_t = \sigma\left(W_r x_t + U_r h_{t-1} + b_r\right), \tag{1.39}$$

$$\tilde{h}_t = \tanh\left[W_h x_t + U_h\left(r_t \odot h_{t-1}\right) + b_h\right], \tag{1.40}$$

$$h_t = \left(1 - z_t\right) \odot h_{t-1} + z_t \odot \tilde{h}_t, \tag{1.41}$$

де z_t – вентиль оновлення; r_t – вентиль скидання; $\tilde{h}_t$ – кандидат на оновлення прихованого стану.

Порівняно з LSTM мережа GRU має простішу структуру. Вибір між LSTM і GRU не може бути здійснений лише на підставі загальних міркувань: ефективність конкретної архітектури необхідно перевіряти експериментально на відповідних рядах, горизонтах прогнозування та обсягах навчальних вибірок.

Для задач прогнозування біржової динаміки GRU може бути доцільною, якщо потрібно зберегти рекурентний механізм обробки даних, але зменшити параметричну складність моделі. Водночас навіть спрощена рекурентна мережа залишається моделлю типу «чорної скриньки», якщо її виходи не пов’язані з інтерпретованими проміжними характеристиками процесу.

Для моделювання часових послідовностей використовують також одновимірні згорткові нейронні мережі – Convolutional Neural Networks (CNN).

Їхня основна ідея полягає в застосуванні ковзних фільтрів для виділення локальних шаблонів у часовому ряді. На відміну від рекурентних мереж, згорткові архітектури допускають паралельне обчислення ознак для різних часових позицій.

Розширенням згорткових нейронних мереж CNN для роботи з послідовностями є часові згорткові мережі TCN (Temporal Convolutional Network). У TCN застосовують каузальні згортки, які не використовують майбутніх значень ряду, та дилатації, що збільшують рецептивне поле без пропорційного збільшення кількості шарів.

Вихід часового згорткового шару можна подати як

$$h_t^{(l)} = \varphi\left[b^{(l)} + \sum_{k=0}^{K-1} W_k^{(l)} h_{t-d_l k}^{(l-1)} \right], \tag{1.42}$$

де K – розмір ядра згортки; d_l – коефіцієнт дилатації на шарі l; $W_k^{(l)}$ – ваги згорткового фільтра; $h_{t-d_l k}^{(l-1)}$ – значення попереднього шару в доступні моменти часу.

Ш. Бай, Дж. З. Колтер і В. Колтун (S. Bai, J. Z. Kolter, V. Koltun) виконали системне порівняння згорткових і рекурентних архітектур для моделювання послідовностей та показали, що TCN можуть бути конкурентоспроможною альтернативою LSTM і GRU, зокрема завдяки довшій ефективній пам’яті та простішому паралельному навчанню [80].

Для внутрішньоденних біржових даних згорткові мережі є придатними для виділення локальних конфігурацій: коротких трендів, імпульсів, розворотів, періодів підвищеної волатильності та повторюваних мікроструктурних шаблонів. Проте якість результату залежить від масштабу часових інтервалів, способу нормалізації даних і довжини рецептивного поля.

Наступний етап розвитку нейромережевого прогнозування пов’язаний із механізмом уваги – attention mechanism. Його призначення полягає у визначенні відносної важливості окремих елементів вхідної послідовності для формування прогнозу.

Базову операцію масштабованої скалярної уваги визначають формулою

$$\text{Attention}(Q,K,V)=\text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V, \qquad (1.43)$$

де Q – матриця запитів; K – матриця ключів; V – матриця значень; d_k – розмірність векторів ключів.

Архітектуру трансформера (Transformer), що базується виключно на механізмах самоуваги, запропонували А. Васвані (A. Vaswani) та його співавтори. На відміну від класичних рекурентних моделей вона ґрунтується на механізмі уваги та не потребує послідовного оновлення прихованого стану для кожного моменту часу [81]. Початково Transformer було розроблено для обробки природної мови, однак згодом attention-based підходи набули поширення у прогнозуванні часових рядів.

Для довгих часових послідовностей класичний механізм уваги має суттєвий недолік: обчислювальна складність і витрати пам'яті квадратично залежать від довжини послідовності. У моделі Informer Х. Чжоу (H. Zhou) та його співавтори запропонували механізм розрідженої самоуваги (ProbSparse self-attention), орієнтований на зменшення обчислювальної складності обробки довгих послідовностей до порядку $O(L\log L)$ [104].

У моделі PatchTST Ю. Ні (Y. Nie) та його співавтори використали підхід пачування (поділу ряду на сегменти – patches), які розглядаються як окремі вхідні токени. Такий підхід дає змогу зберігати локальну контекстну інформацію, суттєво скорочувати розмір матриць уваги та аналізувати значно довший історичний контекст [106].

Модель iTransformer (inverted Transformer) змінює традиційну інтерпретацію вхідних токенів: механізм уваги застосовують не між часовими позиціями, а між змінними багатовимірного ряду. Автори аргументують, що така інверсія дає змогу краще враховувати міжознакові залежності та ефективніше працювати з довгими вхідними вікнами [108].

Разом із тим використання Transformer-архітектур не слід розглядати як безумовно оптимальне рішення для будь-якого часового ряду. А. Зенг, М. Чен, Л. Чжан і Ц. Сюй (A. Zeng, M. Chen, L. Zhang, Q. Xu) показали, що в низці задач

довгострокового прогнозування прості лінійні моделі можуть перевершувати складні Transformer-архітектури Автори пов’язують це з особливостями вилучення часових залежностей із впорядкованих числових послідовностей і ризиком моделювання шуму замість стійких закономірностей [105]. Цей висновок має особливе значення для коротких внутрішньоденних фінансових вибірок. Якщо доступна історія є обмеженою, а локальна структура швидко змінюється, збільшення глибини та кількості параметрів моделі не обов’язково забезпечує підвищення точності прогнозу.

Поряд із універсальними архітектурами MLP, LSTM, GRU, TCN і Transformer розроблено спеціалізовані нейромережеві моделі прогнозування, структура яких враховує особливості часових рядів.

Модель DeepAR, запропонована Д. Салінасом (D. Salinas) та його співавторами, призначена для імовірнісного прогнозування множини пов’язаних часових рядів [87].

У загальному вигляді спільний розподіл прогнозованої послідовності можна подати як

$$P\left(y_{t+1:t+H} \mid y_{1:t}, x_{1:t+H}; \theta\right) = \prod_{\tau=t+1}^{t+H} p\left(y_{\tau} \mid \eta_{\tau}\right), \tag{1.44}$$

де H – горизонт прогнозування; $x_{1:t+H}$ – доступні додаткові ознаки; η_{τ} – параметри прогнозного розподілу, сформовані рекурентною мережею.

На відміну від точкового прогнозу, імовірнісний підхід дає змогу оцінювати невизначеність. Для біржових процесів це важливо через стохастичний характер цінової динаміки. Водночас DeepAR орієнтована насамперед на навчання за множиною пов’язаних рядів, тому її застосування до одиничної короткої внутрішньоденної послідовності потребує спеціального обґрунтування.

Архітектура трансформера часового об'єднання TFT (Temporal Fusion Transformer), запропонована Б. Лімом, С. Ариком, Н. Лоффом і Т. Пфістером (B. Lim, S. Arık, N. Loeff, T. Pfister), поєднує рекурентні шари для локальної обробки, механізм інтерпретованої уваги для моделювання довготривалих залежностей,

вентильні компоненти та засоби відбору ознак [82]. TFT призначена для багатогоризонтного прогнозування за наявності статичних, історичних і наперед відомих змінних.

Для навчання моделей, які прогнозують умовні квантилі, застосовують квантильну функцію втрат:

$$L_q\left(y,\hat{y}^{(q)}\right)=\max\left[q\left(y-\hat{y}^{(q)}\right),(q-1)\left(y-\hat{y}^{(q)}\right)\right], \tag{1.45}$$

де $q\in(0,1)$ – заданий квантиль; $\hat{y}^{(q)}$ – прогноз відповідного квантиля.

TFT є прикладом архітектури, у якій висока виразність нейронної мережі поєднується з частковою інтерпретованістю. Однак така модель має значну структурну складність і потребує достатнього обсягу даних для стабільного навчання.

Модель N-BEATS – Neural Basis Expansion Analysis for Interpretable Time Series Forecasting – запропонували Б. Орєшкін, Д. Карпов, Н. Шападо та І. Бенджіо (B. Oreshkin, D. Carpov, N. Chapados і Y. Bengio) [89]. Архітектура побудована на глибоких повнозв'язних блоках із прямими та зворотними залишковими зв'язками. У кожному блоці формують дві складові: апроксимацію вхідної послідовності – backcast та частковий прогноз – forecast.

Послідовне оновлення залишку можна подати як

$$X^{(l+1)}=X^{(l)}-\hat{X}^{(l)}, \tag{1.46}$$

а остаточний прогноз – як суму прогнозів окремих блоків:

$$\hat{Y}=\sum_{l=1}^{B}\hat{Y}^{(l)}, \tag{1.47}$$

де $X^{(l)}$ – вхідний залишок для блока l; $\hat{X}^{(l)}$ – відновлена частина вхідного сигналу; $\hat{Y}^{(l)}$ – частковий прогноз; B – кількість блоків.

Інтерпретований варіант N-BEATS використовує базисні функції тренду та сезонності. Ідея виділення структурованих прогнозних компонентів є важливою для цієї дисертаційної роботи, оскільки демонструє доцільність інтегрування нейромережевих механізмів із математично обґрунтованими проміжними представленнями часового ряду.

Модель N-HiTS розвиває підхід N-BEATS шляхом застосування ієрархічної інтерполяції та багатомасштабної дискретизації вхідних даних. Це дає змогу поетапно враховувати компоненти різної частоти й масштабу, зменшуючи обчислювальні витрати під час довгогоризонтного прогнозування [103].

Для короткострокового прогнозування біржової динаміки особливий інтерес становить не стільки конкретна структура N-HiTS, скільки загальний принцип багатомасштабного аналізу: прогноз доцільно формувати не на основі єдиного жорстко заданого параметра, а шляхом адаптивного поєднання результатів моделей із різною локальною чутливістю.

У моделі TimesNet Х. Ву (H. Wu) та його співавтори запропонували концептуально новий підхід, що передбачає перетворення одновимірного часового ряду на множину двовимірних представлень відповідно до домінантних періодів, виявлених за допомогою швидкого перетворення Фур'є FFT (Fast Fourier Transform). Це дає змогу моделювати внутрішньоперіодні та міжперіодні зміни за допомогою двовимірних згорткових блоків [107].

TimesNet ілюструє тенденцію до побудови спеціалізованих представлень часових рядів перед їх обробкою нейронною мережею. Аналогічний принцип може бути реалізований для біржових рядів шляхом переходу від безпосередньої обробки сирих котирувань до аналізу структурованої послідовності локальних прогнозів.

Окремим напрямом стали попередньо навчені моделі часових рядів. У моделі Chronos А. Ансарі (A. Ansari) та його співавтори запропонували підхід, що передбачає перетворення числових значень часового ряду на дискретні токени за допомогою процедур масштабування та квантування за рівнем (квантизації), після чого застосовують класичну мовну архітектуру Transformer [109]. Модель попередньо навчено на масштабному крос-доменному корпусі даних (cross-domain dataset) із різних предметних областей, що створює теоретичні та практичні передумови для ефективного застосування прогнозування без додаткового навчання (zero-shot forecasting) на нових часових рядах.

Попередньо навчені моделі є перспективними для масштабних інформаційних систем, але їх застосування не усуває потреби в спеціалізованих методах для задач із чітко визначеними обмеженнями: малим локальним вікном, однокроковим горизонтом, високою волатильністю та необхідністю пояснити механізм формування прогнозу.

Нейромережеві методи мають декілька принципових переваг:

1) нелінійність – можливість апроксимувати складні залежності без жорсткого задання їх аналітичної форми;
2) адаптивність – автоматичне уточнення вагових коефіцієнтів на основі емпіричних даних;
3) робота з багатовимірними даними – урахування цін, обсягів, індикаторів і зовнішніх факторів;
4) виділення прихованих ознак – формування внутрішніх представлень локальної структури ряду;
5) можливість гібридизації – поєднання з класичними статистичними моделями, декомпозицією, кластеризацією та алгоритмами адаптивного вибору параметрів.

Водночас застосування нейронних мереж до біржових часових рядів має істотні обмеження:

1) ризик перенавчання в умовах малого обсягу навчальних даних;
2) нестійкість до зміни ринкових режимів;
3) складність інтерпретації прогнозів глибоких моделей;
4) залежність від способу нормалізації, вибору вікна спостережень і гіперпараметрів;
5) значні обчислювальні витрати для складних рекурентних і Transformer-архітектур;
6) можливість прихованого витоку даних, якщо часовий порядок порушено під час формування навчальної та тестової вибірок;
7) відсутність гарантії переваги складної архітектури над простішою моделлю.

Останнє обмеження є принципово важливим. Результати А. Зенга та співавторів демонструють, що навіть для задач довгострокового прогнозування прості лінійні моделі можуть виявитися сильними базовими рішеннями [105]. Отже, під час оцінювання нових нейромережевих моделей необхідно порівнювати їх не лише з іншими глибокими мережами, але й із прозорими статистичними, лінійними та локальними екстраполяційними алгоритмами.

Порівняльна характеристика основних нейромережевих архітектур наведена в додатку В.

1.5. Сучасні програмні комплекси для аналізу часових рядів біржової динаміки

Сучасний аналіз часових рядів біржової динаміки здійснюється за допомогою широкого спектра програмних засобів, які відрізняються призначенням, рівнем доступу до ринкових даних, можливостями статистичного й нейромережевого моделювання, підтримкою алгоритмічної торгівлі, засобами візуалізації та умовами поширення.

Умовно такі програмні комплекси можна поділити на п’ять основних груп:

1) професійні фінансово-аналітичні термінали;
2) торгово-аналітичні платформи технічного аналізу;
3) середовища алгоритмічної торгівлі та backtesting;
4) наукові програмні середовища для статистичного аналізу й машинного навчання;
5) системи зберігання та обробки високочастотних часових даних.

Такий поділ обумовлений тим, що не всі популярні платформи однаково придатні для наукового дослідження. Частина систем орієнтована переважно на практичний трейдинг, частина – на професійну фінансову аналітику, а найбільш придатними для відтворюваних експериментів залишаються відкриті або програмно розширювані середовища на основі Python, R, MATLAB та спеціалізованих backtesting-фреймворків.

До професійних фінансово-аналітичних терміналів належать Bloomberg Terminal, LSEG Workspace, FactSet Workstation, S&P Capital IQ Pro, Morningstar

Direct та інші інституційні системи. Їх основне призначення полягає в наданні доступу до ринкових даних, новин, корпоративної звітності, аналітики, оцінювання активів, портфельного аналізу та інструментів підтримки інвестиційних рішень.

Bloomberg Terminal позиціонується як комплексна платформа для фінансових аналітиків та трейдерів, що забезпечує доступ до ринкових даних у реальному часі, стрічок новин, аналітичного інструментарію, систем виконання торговельних угод, а також засобів передторговельного та післяторговельного аналізу (pre-trade та post-trade analysis) [90]. Платформа LSEG Workspace (London Stock Exchange Group) надає персоналізовані фінансові дані, новини та аналітичні інструменти для професійних робочих процесів, спираючись на масштабну інфраструктуру збирання й обробки даних, яка охоплює сотні тисяч користувачів на багатьох світових ринках [91]. У свою чергу, FactSet пропонує інтегровану платформу фінансових даних, аналітики та рішень на основі штучного інтелекту для інвестиційного аналізу, управління капіталом (wealth management) та автоматизації робочих процесів з великими масивами даних [92].

Основними перевагами професійних фінансових терміналів є:

- доступ до якісних ринкових, фундаментальних, новинних і макроекономічних даних;
- висока надійність джерел інформації;
- інтеграція інструментів аналізу, новин, корпоративних подій і торгових операцій;
- придатність для інституційного інвестиційного аналізу;
- наявність API або експортних механізмів для подальшої обробки даних.

З погляду наукового дослідження біржових часових рядів такі системи мають низку обмежень:

- висока вартість ліцензії або підписки;
- закритість програмного середовища;
- обмеження на масове завантаження й повторне поширення даних;
- залежність від комерційного постачальника;

- складність повної відтворюваності експериментів іншими дослідниками без аналогічної ліцензії.

Ці системи поширюються переважно за комерційною підписною моделлю або за інституційними ліцензіями. Вони найбільш доцільні для професійної фінансової аналітики, але менш зручні як основне середовище розроблення власних алгоритмів прогнозування.

До другої групи належать TradingView, MetaTrader 5, NinjaTrader, AmiBroker, MetaStock, TradeStation, Sierra Chart, MultiCharts та інші спеціалізовані системи, орієнтовані на технічний аналіз, побудову інтерактивних графіків, розроблення індикаторів, генерацію торгових сигналів та імітаційне моделювання торговельних стратегій.

Платформа TradingView є популярним хмарним вебсередовищем для візуалізації ринкової динаміки, створення користувацьких індикаторів і стратегій за допомогою вбудованої мови Pine Script, а також проведення історичного тестування (backtesting) торговельних алгоритмів із використанням розширених масивів історичних та поточних ринкових даних [93].

Платформа MetaTrader 5 орієнтована на забезпечення технічного аналізу й автоматизованої торгівлі на ринках Forex, акцій та ф’ючерсів. Вона підтримує розроблення алгоритмічних торгових роботів на базі об'єктно-орієнтованої мови MQL5, а також містить вбудований модуль тестування стратегій (Strategy Tester), який дозволяє здійснювати історичну перевірку, мультивалютне тестування та оптимізацію параметрів експертних систем (Expert Advisors) [94].

Середовище NinjaTrader фокусується переважно на ф’ючерсному трейдингу, забезпечуючи розширений графічний аналіз, симуляцію виконання угод та історичне тестування алгоритмів. Особливістю платформи є наявність інструменту відтворення ринку (Market Replay), що дозволяє симулювати торги в реальному часі на основі історичних тікових даних [95].

Платформа AmiBroker є високопродуктивним спеціалізованим середовищем для технічного аналізу та розроблення складних торгових систем. До її ключових аналітичних можливостей належать застосування високорівневої

мови формул AFL (AmiBroker Formula Language), проведення портфельного історичного тестування, багатокритеріальна оптимізація параметрів, послідовне тестування методом «ковзного вікна» (walk-forward testing), а також оцінювання ризиків та стійкості систем за допомогою статистичного моделювання за методом Монте-Карло (Monte Carlo simulation) [96].

Перевагами торгово-аналітичних платформ є:

- зручна інтерактивна візуалізація біржових рядів;
- підтримка технічних індикаторів;
- можливість швидкої побудови й перевірки торгових правил;
- наявність власних мов сценаріїв: Pine Script, MQL5, AFL тощо;
- можливість симуляції або backtesting стратегій.

До основних недоліків належать:

- орієнтація переважно на практичний трейдинг, а не на науковий аналіз;
- обмежені можливості реалізації нестандартних математичних моделей;
- складність інтеграції з власними дослідницькими pipeline;
- залежність від формату даних, брокера або платформи;
- у частини систем – закритий код і комерційна модель поширення.

Окрему групу становлять програмні комплекси, призначені для кількісного аналізу, зворотнього тестування, оптимізації та алгоритмічної торгівлі. Найбільш відомими є QuantConnect/LEAN, Backtrader, vectorbt, Zipline, bt, PyAlgoTrade.

Платформа QuantConnect є хмарним середовищем для проведення кількісних досліджень (quantitative research), історичного тестування та безпосереднього виконання алгоритмічних угод у реальному часі. Ядром платформи є програмний рушій із відкритим вихідним кодом LEAN Engine, який забезпечує інфраструктурну підтримку дослідницьких процесів, ретроспективного аналізу та торгівлі, підтримуючи розроблення алгоритмів мовами Python і C# [97].

Backtrader є відкритим Python-фреймворком для історичного тестування та алгоритмічної торгівлі, архітектура якого дозволяє досліднику зосередитися на безпосередній програмній реалізації стратегій, користувацьких індикаторів та

модулів аналізу результатів, мінімізуючи витрати на розроблення базової інфраструктури обробки даних [98].

Пакет vectorbt є спеціалізованою бібліотекою мови Python для високопродуктивного кількісного аналізу. Він інтегрується з об'єктами бібліотек pandas та NumPy і використовує механізми JIT-компіляції для прискорення обчислень (за допомогою Numba або Rust), що дає змогу здійснювати масове паралельне тестування великої кількості комбінацій стратегій та оптимізацію їхніх гіперпараметрів у векторизованій формі [99].

Узагальнюючи аналіз інструментальних засобів моделювання, можна констатувати, що перехід від закритих пропрієтарних платформ до відкритих програмних фреймворків кількісного аналізу (передусім екосистеми мови Python) забезпечує високу гнучкість програмної реалізації авторських прогнозних моделей, глибинну інтеграцію із сучасними архітектурами глибокого навчання, підтримку строгого форвардного аналізу методом «ковзного вікна» та повну відтворюваність наукових експериментів за комплексними статистичними й фінансовими метриками. Водночас застосування таких інструментів характеризується підвищеним обчислювальним порогом входження, повною інфраструктурною відповідальністю дослідника за очищення й контроль якості первинних масивів даних, а також високим ризиком виникнення прихованих системно-методичних помилок історичного тестування — зокрема, зазирання в майбутнє (look-ahead bias), помилки виживання (survivorship bias) та спрощеного моделювання ринкових транзакційних витрат, що вимагає розроблення спеціалізованих архітектурних рішень для забезпечення строгості та адекватності отриманих результатів.

Найбільш гнучкими для наукового дослідження часових рядів є універсальні середовища математичного моделювання, статистичного аналізу та машинного навчання. Python сьогодні є одним із найпоширеніших середовищ для аналізу фінансових часових рядів. Його перевага полягає у поєднанні бібліотек для обробки даних, статистики, машинного навчання, глибокого навчання, візуалізації та backtesting [100].

R традиційно є сильним середовищем для статистики, економетрики й аналізу фінансових даних.

Інструментальне середовище MATLAB широко застосовується для математичного моделювання, фінансового інжинірингу, портфельної оптимізації, економетричного прогнозування та імітаційного моделювання складних систем. Спеціалізований модуль Financial Toolbox забезпечує математичний та статистичний аналіз фінансових показників, ретроспективне тестування торговельних алгоритмів, а також оптимізацію структури інвестиційних портфелів. Модуль Econometrics Toolbox призначений для специфікації, оцінювання параметрів та прогнозування економічних часових рядів, забезпечуючи інструментарій для діагностики стаціонарності, аналізу автокореляції, тестування коінтеграції, перевірки причинності за Грейнджером та виявлення структурних зрушень у даних [101].

Під час роботи з високочастотними, внутрішньоденними та тиковими біржовими даними критичне значення має не лише вибір прогнозних моделей, а й архітектура інфраструктури збереження та швидкого доступу до інформації. Для вирішення таких задач застосовують спеціалізовані бази даних часових рядів (TSDB – Time-Series Databases), зокрема kdb+, ArcticDB, InfluxDB, ClickHouse, TimescaleDB та QuestDB.

СУБД kdb+ розроблена спеціалізовано для фінансових ринків і характеризується високою ефективністю під час збереження, обробки та низькозатримного аналізу потокових історичних високочастотних котирувань.

Високопродуктивна база даних ArcticDB орієнтована на задачі кількісного аналізу даних (quantitative data science) і здатна оперувати матричними структурами, що налічують мільярди рядків і тисячі колонок, що є критично важливим для обробки щільних масивів біржових тикових даних.

Універсальна система InfluxDB забезпечує збереження часових рядів із високою точністю часових міток (до мікро- та наносекундного рівнів), що дозволяє фіксувати точну послідовність транзакцій у системах високочастотного алгоритмічного трейдингу.

Колонкова аналітична система керування базами даних ClickHouse адаптована для виконання аналітики в реальному часі (real-time analytics), забезпечуючи високу швидкість виконання складних аналітичних запитів та масштабних агрегацій над розподіленими часовими масивами великої розмірності [101].

Проведений огляд свідчить, що сучасні програмні комплекси для аналізу часових рядів біржової динаміки суттєво відрізняються за цільовим призначенням. Наявні програмні комплекси забезпечують широкі можливості аналізу біржових часових рядів, проте більшість із них не містить спеціалізованих засобів адаптивного прогнозування на основі послідовності поліноміальних прогнозів, що обґрунтовує необхідність розроблення власного алгоритмічного та програмного забезпечення дослідження.

Узагальнену характеристику розглянутих програмних засобів наведено в додатку Г.

1.6. Критичний огляд існуючих підходів та постановка наукового завдання

Проведений у підрозділах 1.1-1.4 аналіз дає підстави розглядати задачу прогнозування біржової динаміки не лише як задачу побудови окремої прогнозної моделі, а як задачу вибору способу вилучення корисної локальної інформації з часового ряду в умовах шуму, обмеженої довжини фрагмента даних та невизначеності майбутньої поведінки ринку. Сучасні підходи до прогнозування часових рядів можна умовно поділити на статистичні, економетричні, методи машинного навчання, нейромережеві та гібридні. Кожна з цих груп має власну сферу ефективного застосування, але жодна не усуває повністю проблему прогнозування біржового ряду за малим локальним фрагментом спостережень.

Класичні статистичні підходи, зокрема моделі ARIMA/ARMA та експоненційне згладжування, є методично прозорими й добре формалізованими. Їхня перевага полягає у зрозумілій структурі, можливості аналітичної інтерпретації параметрів та наявності усталених процедур перевірки якості моделі. Разом із тим такі моделі потребують попереднього приведення ряду до форми, прийнятної для параметричної оцінки, а їхня ефективність значною мірою залежить від коректного вибору порядків моделі, способу диференціювання, довжини навчального інтервалу та припущень щодо залишків. У працях R. Hyndman і G. Athanasopoulos наголошується, що прикладне використання моделей прогнозування потребує не лише формального підбору моделі, а й обґрунтованого врахування її припущень та меж застосовності [27].

Економетричні моделі волатильності, зокрема ARCH/GARCH-підходи, мають значну цінність для аналізу умовної дисперсії та ризику, однак вони переважно орієнтовані не на безпосереднє відновлення наступного цінового значення, а на моделювання мінливості ряду. Модель GARCH, запропонована T. Bollerslev, узагальнює ARCH-процес і враховує залежність умовної дисперсії від попередніх збурень та попередніх значень дисперсії. Проте для задачі короткострокового прогнозування ціни або значення біржового індикатора це

означає, що модель волатильності часто потребує поєднання з окремою моделлю умовного середнього [7].

Методи машинного навчання, зокрема SVM, Random Forest, Gradient Boosting, MLP та інші алгоритми, здатні апроксимувати нелінійні залежності без жорсткого попереднього задання функціональної форми. У роботі V. Derbentsev, A. Matviychuk, N. Datsenko, V. Bezkorovainyi та A. Azaryan досліджено короткострокове прогнозування фінансових часових рядів за допомогою SVM, MLP, Random Forest і Stochastic Gradient Boosting Machine; автори використовували лише минулі значення цільової змінної як предиктори та застосовували one-step-ahead оцінювання [8]. Подібний напрям розвинуто також у дослідженнях V. Derbentsev, A. Matviychuk та V. Soloviev щодо прогнозування криптовалютних рядів методами BART, MLP та Random Forest [9]. Водночас такі моделі потребують ретельного налаштування гіперпараметрів, достатнього навчального набору та контролю перенавчання, особливо коли кількість доступних спостережень є малою.

Нейромережеві архітектури сучасного типу – LSTM, GRU, CNN-LSTM, attention-based моделі, Transformer-подібні підходи – забезпечують високу гнучкість для моделювання складних часових залежностей. Наприклад, N-BEATS запропоновано як глибоку архітектуру з backward/forward residual links для задачі одномірного прогнозування часових рядів, причому її інтерпретована конфігурація використовує базисні подання трендових і сезонних компонент [89]. Temporal Fusion Transformer поєднує рекурентні шари, attention-механізм та блоки відбору ознак для багатогоризонтного прогнозування з елементами інтерпретованості [88]. Informer та Autoformer орієнтовані переважно на long sequence time-series forecasting і розв’язують проблеми обчислювальної складності та довгих залежностей [104]. Однак для короткострокового прогнозу за малим локальним фрагментом біржових даних такі архітектури можуть бути надлишково складними, потребувати значних обсягів навчальних даних і не завжди забезпечувати просту інтерпретацію механізму формування конкретного прогнозу.

Окремий напрям становлять гібридні й ансамблеві методи. Результати M4-competition показали переваги комбінування статистичних і ML-компонентів: серед найкращих підходів були саме гібридні або комбінаційні рішення [112]. В українському науковому контексті гібридні та ансамблеві моделі прогнозування фінансових часових рядів досліджено, зокрема, у дисертаційній роботі К. А. Токарєвої [113], присвяченій поєднанню методів машинного навчання з класичними алгоритмами часових рядів. Такі підходи є перспективними, але їхнім недоліком є ускладнення архітектури, зростання кількості параметрів і потреба у додатковому механізмі вибору або узгодження часткових прогнозів.

Порівняння існуючих підходів наведено в таблиці 1.1

Таблиця 1.1 – Порівняння існуючих підходів до прогнозування часових рядів біржових котирувань

Група підходів	**Переваги**	**Обмеження для задачі дисертації**
ARIMA, ETS, регресійні моделі	Прозора математична структура, інтерпретованість, усталені критерії якості	Потреба у виборі порядків моделі, залежність від припущень, складність адаптації до коротких локальних фрагментів
ARCH/GARCH	Ефективні для аналізу умовної дисперсії та ризику	Не розв’язують безпосередньо задачу формування точкового прогнозу ціни
SVM, RF, GBM, MLP	Здатність апроксимувати нелінійності	Потреба у навчальних даних, гіперпараметрах, ознаках; ризик перенавчання
LSTM, GRU, Transformer-моделі	Моделювання складних часових залежностей	Висока обчислювальна складність, потреба у великих наборах даних, часткова непрозорість
Гібридні та ансамблеві моделі	Поєднання переваг різних класів методів	Ускладнення моделі, проблема вибору ваг або правил комбінування
Поліноміальна локальна екстраполяція	Робота з малим фрагментом ряду, аналітична прозорість, проста реалізація	Потреба в адаптивному виборі степеня полінома або інформативної підмножини прогнозних значень

У роботі особливе значення має останній клас підходів (табл. 1.1). Поліноміальна екстраполяція є природним інструментом локального

прогнозування, оскільки дозволяє будувати прогноз на основі останніх значень ряду без необхідності формування великої навчальної вибірки. Водночас пряме використання полінома фіксованої степені може бути недостатньо надійним: малий степінь полінома може не враховувати локальну зміну тенденції, а великий степінь може призводити до небажаних коливань прогнозних значень. Саме тому однією з ключовою ідей стає не лише побудова одного поліноміального прогнозу, а аналіз послідовності прогнозів, отриманих за різних значень параметра глибини або степеня локальної поліноміальної моделі.

Критичний аналіз наявних підходів дозволяє виділити такі невирішені аспекти:

1. Більшість класичних і ML-моделей орієнтована або на глобальне навчання, або на вибір параметрів за всією доступною історією. Для біржових даних це не завжди оптимально, оскільки локальна структура останніх спостережень може мати більшу прогностичну цінність, ніж давні фрагменти ряду.

2. У багатьох підходах параметри прогнозування задаються наперед або визначаються через перебір, але недостатньо дослідженим залишається питання адаптивного вибору параметра поліноміальної екстраполяції за поточним станом ряду.

3. Сучасні нейромережеві моделі досягають високої гнучкості, але часто втрачають аналітичну прозорість. Для дисертаційного дослідження важливо зберегти зв’язок між математичною формою прогнозу та алгоритмом його формування.

4. Гібридні й ансамблеві моделі підтверджують доцільність комбінування різних прогнозів, але не розв’язують у загальному вигляді питання: які саме часткові прогнози слід вважати інформативними для конкретного локального фрагмента біржового ряду.

5. Недостатньо формалізованим залишається використання послідовності поліноміальних прогнозних значень як самостійного об’єкта інтелектуального аналізу: її можна досліджувати з погляду кластеризації, щільності розташування

прогнозних значень, наявності екстремумів, розбіжності елементів і придатності до адаптивного усереднення.

З урахуванням виявлених обмежень існуючих підходів наукове завдання дисертаційної роботи полягає у розробленні методів та алгоритмів інтелектуального аналізу часових рядів біржової динаміки на основі поліноміального підходу, які забезпечують формування короткострокового прогнозу шляхом побудови, аналізу та адаптивного використання послідовності поліноміальних прогнозних значень.

Для розв'язання цього наукового завдання необхідно:

1. Формалізувати задачу прогнозування біржового часового ряду в умовах обмеженого обсягу локальних даних.

2. Розробити метод побудови послідовності поліноміальних прогнозних значень PPS_n для різних значень параметра глибини m та степеня локальної поліноміальної моделі.

3. Запропонувати критерії аналізу структури PPS для виявлення інформативних прогнозних значень.

4. Розробити алгоритми адаптивного вибору кількості або підмножини елементів PPS, які доцільно використовувати для формування остаточного прогнозу.

5. Дослідити можливість використання кластерного аналізу, аналізу щільних інтервалів, екстремальних значень та критеріїв розбіжності PPS для підвищення точності короткострокового прогнозування.

6. Розробити нейромережеву архітектуру, у якій поліноміальна прогнозна послідовність використовується як спеціалізований прогнозний шар або як вхід до механізму адаптивного зважування.

7. Провести експериментальну перевірку запропонованих методів на часових рядах біржової динаміки з використанням walk-forward схеми та метрик MAE, RMSE, MAPE.

Висновки до першого розділу

Часові ряди біржової динаміки є складним об’єктом інтелектуального аналізу. Їх основними властивостями є шумовий характер, негаусівські розподіли дохідностей, кластеризація волатильності, нестаціонарність, зміна ринкових режимів, залежність від часового масштабу та наявність короткочасних локальних закономірностей.

Для внутрішньоденних даних додатково необхідно враховувати мікроструктурний шум, внутрішньосесійну періодичність, нерівномірність інтенсивності торгів, нічні розриви та залежність результатів від способу агрегування спостережень.

За таких умов прогнозування на основі великої історичної вибірки не завжди є оптимальним. Актуальною є розробка адаптивних методів, які використовують обмежений локальний фрагмент часового ряду, мають контрольовану складність і дають змогу оцінювати внутрішню узгодженість прогнозів. Це обґрунтовує доцільність дослідження методів короткострокового прогнозування біржової динаміки на основі поліноміальної екстраполяції, адаптивного вибору її параметрів і структурного аналізу послідовностей поліноміальних прогнозних значень.

Класичні статистичні методи створюють фундаментальну основу прогнозування фінансових часових рядів. Їх перевагами є математична обґрунтованість, інтерпретованість параметрів і можливість формалізованого оцінювання статистичних властивостей процесу. Наївні моделі, експоненціальне згладжування та ARIMA доцільно використовувати як базові орієнтири під час експериментального порівняння алгоритмів. Моделі ARCH і GARCH є важливими для аналізу волатильності, а VAR, VECM, моделі простору станів і марковські моделі – для врахування складніших залежностей.

Водночас застосування класичних методів до коротких внутрішньоденних біржових рядів ускладнюється нестаціонарністю, випадковими коливаннями, локальними змінами тренду, нестійкістю параметрів і недостатнім обсягом даних для надійного оцінювання складних моделей. Підвищення порядку статистичної

моделі не завжди покращує прогноз, оскільки може спричинити перенавчання та значне зростання дисперсії оцінок.

Зазначені обмеження обґрунтовують доцільність розроблення адаптивних локальних методів короткострокового прогнозування. У межах дисертаційного дослідження таким напрямом є використання послідовності поліноміальних прогнозних значень, адаптивний вибір параметрів поліноміальної екстраполяції та аналіз структурних властивостей множини локальних прогнозів. Запропонований поліноміальний підхід не заперечує значення класичних статистичних моделей, а розвиває методи прогнозування для умов обмеженого обсягу даних і недетермінованості біржових процесів.

Методи машинного навчання розширюють можливості аналізу біржової динаміки порівняно з класичними статистичними моделями. Їх принциповою перевагою є здатність виявляти нелінійні залежності, враховувати взаємодії між ознаками та адаптувати параметри прогнозної моделі до структури навчальних даних. У задачах прогнозування фінансових часових рядів застосовуються регуляризовані регресійні моделі, методи класифікації, алгоритми найближчих сусідів, метод опорних векторів, дерева рішень, випадкові ліси, бустингові алгоритми, ансамблеві моделі та нейронні мережі [67, 71–74].

Водночас ефективність машинного навчання обмежується специфічними властивостями біржових процесів: нестаціонарністю, низьким співвідношенням корисного сигналу та шуму, змінами ринкових режимів, обмеженим обсягом локальних вибірок і ризиком витоку майбутньої інформації. Збільшення складності моделі саме по собі не гарантує підвищення точності. У сучасних порівняльних дослідженнях не визначено універсального алгоритму, який забезпечував би стабільну перевагу для всіх ринків, часових інтервалів і горизонтів прогнозування [68, 70, 74].

Для короткострокового прогнозування внутрішньоденних часових рядів особливого значення набувають методи, які можуть працювати з обмеженим обсягом актуальних спостережень і враховувати локальний характер біржової динаміки. Це обґрунтовує доцільність дослідження поліноміальної екстраполяції

як інструменту локального прогнозування. Важливим напрямом вдосконалення такого підходу є формування послідовності поліноміальних прогнозних значень, аналіз її структурних властивостей і адаптивний вибір інформативної підмножини прогнозів. Отримані результати можуть бути використані як самостійно, так і в складі гібридних моделей інтелектуального аналізу біржових часових рядів.

РОЗДІЛ 2. МЕТОДИЧНІ ОСНОВИ ПОЛІНОМІАЛЬНОГО ПІДХОДУ ДО АНАЛІЗУ ЧАСОВИХ РЯДІВ БІРЖОВОЇ ДИНАМІКИ

2.1. Формалізація задачі прогнозування біржового часового ряду

Прогнозування біржової динаміки належить до складних задач інтелектуального аналізу даних, оскільки фінансові часові ряди формуються під впливом значної кількості взаємопов'язаних чинників: зміни співвідношення попиту і пропозиції, очікувань учасників ринку, надходження нової інформації, макроекономічних подій, спекулятивної активності та випадкових збурень. На відміну від багатьох технічних процесів, біржові процеси не мають повністю детермінованого механізму формування значень. Їх статистичні властивості можуть змінюватися в часі, а закономірності, виявлені на одному часовому інтервалі, не завжди зберігаються на наступних інтервалах.

Теоретичне підґрунтя для розуміння обмежень прогнозованості фінансових рядів сформовано, зокрема, у роботі Ю. Фами, присвяченій гіпотезі ефективного ринку [5]. У слабкій формі цієї гіпотези припускається, що поточна ціна вже враховує інформацію, яка міститься в попередніх значеннях котирувань. Це не означає абсолютної неможливості прогнозування, однак обмежує надійність довгострокових закономірностей і підвищує вимоги до адаптивності моделей. У праці Р. Конта узагальнено емпіричні властивості фінансових рядів, зокрема важкі хвости розподілів дохідностей, кластеризацію волатильності, нелінійну залежність і неоднорідність динаміки в часі [1].

Класичні підходи до прогнозування часових рядів ґрунтуються на виявленні та формалізації статистичних залежностей між послідовними спостереженнями. До них належать авторегресійні моделі, моделі ковзного середнього, ARMA-, ARIMA- та сезонні моделі, а також моделі умовної гетероскедастичності [28; 36]. Водночас застосування таких методів до внутрішньоденних біржових даних потребує врахування локальної нестаціонарності, значного рівня шуму та обмеженої тривалості часових фрагментів, що використовуються для оперативного прогнозування.

Вагомий внесок у розвиток методів прогнозування економічних і фінансових процесів зроблено українськими науковцями. У роботі П. І. Бідюка

розглянуто системний підхід до побудови прогнозних моделей часових рядів для стаціонарних і нестаціонарних процесів, детермінованих та стохастичних трендів, гетероскедастичних і коінтегрованих процесів [114]. У праці П. І. Бідюка та А. В. Федорова запропоновано ймовірнісне прогнозування процесів ціноутворення на фондових ринках із використанням авторегресійних моделей і динамічних байєсівських мереж [115]. Автори також розглядають прогнозування на основі ковзного вікна спостережень і схеми послідовного накопичення даних. У дослідженні І. А. Шубенкової, С. К. Петрової та П. І. Бідюка розвинуто адаптивне моделювання з використанням регресійних моделей і фільтра Калмана для зменшення впливу випадкових збурень і похибок вимірювання [116].

Сучасні дослідження підтверджують, що точність прогнозу фінансового ряду не можна оцінювати ізольовано від способу формування вибірки, горизонту прогнозування та режиму перевірки моделі. У змаганні M6 Forecasting Competition окрему увагу приділено зв’язку між точністю прогнозування та якістю інвестиційних рішень. Результати змагання засвідчили складність стабільного отримання переваги на фінансових ринках навіть за умови застосування сучасних моделей і значного обсягу даних [111].

У межах цього дослідження розглядається задача однокрокового прогнозування біржового часового ряду. Її особливістю є використання локального фрагмента спостережень обмеженої довжини. Така постановка відповідає умовам оперативного аналізу внутрішньоденних котирувань, коли модель повинна адаптуватися до поточного стану процесу без використання надмірно тривалої історії, яка може містити застарілі закономірності.

Нехай у дискретні моменти часу $t = 1, 2, \ldots, T$ спостерігаються значення певного біржового показника. Біржовий часовий ряд подамо у вигляді впорядкованої послідовності:

$$\mathcal{X}_{1:T} = \{x_t\}_{t=1}^{T} = (x_1, x_2, \ldots, x_T), \tag{2.1}$$

де x_t – значення досліджуваного показника в момент часу t; T – загальна кількість спостережень.

Залежно від змісту експерименту значення x_t може відповідати ціні відкриття торговельного інтервалу $Open_t$, максимальній ціні $High_t$, мінімальній ціні Low_t, ціні закриття $Close_t$, обсягу торгів $Volume_t$ або значенню похідного фінансового індикатора. У подальшому викладі використовується узагальнене позначення x_t. В експериментальній частині роботи основним прогнозованим показником обрауj ціну закриття $Close_t$, оскільки вона характеризує підсумковий стан активу на відповідному часовому інтервалі.

Для формалізації локального характеру біржової динаміки використаємо робоче подання:

$$x_t = s_t + \varepsilon_t . \qquad (2.2)$$

де s_t –локально регулярна складова процесу; ε_t – випадкова складова, що охоплює ринковий шум, вплив неврахованих чинників і короткочасні збурення.

Подання (2.3) не означає, що регулярна та випадкова складові можуть бути точно відокремлені для кожного спостереження. Воно використовується як методологічна передумова: на короткому часовому фрагменті допускається існування локальної структури, яку можна наближено відтворити за допомогою поліноміальної моделі. Водночас зі збільшенням рівня шуму або зміною ринкового режиму стійкість такого наближення може знижуватися.

Для прогнозування наступного значення використовується не весь накопичений ряд, а локальне вікно останніх спостережень:

$$X_t = \left(x_{t-L+1}, x_{t-L+2}, \ldots, x_t\right), \qquad L \geq 2 , \qquad (2.3)$$

де X_t – локальний фрагмент часового ряду, доступний у момент часу t; L – довжина вікна спостережень.

Використання ковзного вікна має принципове значення для внутрішньоденних біржових даних. Воно дає змогу зосередити модель на поточній локальній тенденції та зменшити вплив віддалених спостережень, які могли бути сформовані в іншому ринковому режимі. Після надходження нового значення x_{t+1} вікно оновлюється шляхом вилучення найстарішого елемента та додавання нового спостереження:

$$X_{t+1} = \left(x_{t-L+2}, x_{t-L+3}, \ldots, x_t, x_{t+1}\right). \tag{2.4}$$

У дослідженнях фінансових процесів також використовується схема накопичення даних, за якої кожен наступний прогноз формується на основі всіх доступних попередніх спостережень:

$$X_t^{(\mathrm{acc})} = \left(x_1, x_2, \ldots, x_t\right). \tag{2.5}$$

Схеми прогнозування на основі ковзного вікна та послідовного накопичення даних розглянуто, зокрема, П. І. Бідюком та А. В. Федоровим [115]. Для запропонованого в дисертаційній роботі підходу основною є схема ковзного вікна (2.4), оскільки вона узгоджується з локальним характером поліноміальної екстраполяції та дає змогу досліджувати прогнозування за умов обмеженого обсягу історичних даних.

У загальному випадку задача прогнозування полягає у побудові оператора $\mathcal{F}_h$, який на основі локального вікна X_t та набору параметрів θ_t визначає прогноз на h часових кроків уперед:

$$\hat{x}_{t+h|t} = \mathcal{F}_h\left(X_t; \theta_t\right), \qquad h = 1, 2, \ldots, H. \tag{2.6}$$

де $\hat{x}_{t+h|t}$ прогноз значення x_{t+h}, сформований на основі інформації, доступної до моменту t; h – горизонт прогнозування; H – максимальний горизонт прогнозування; θ_t – параметри прогнозної моделі.

У межах дисертаційного дослідження основна увага зосереджується на однокроковому прогнозуванні, тобто на випадку $h = 1$:

$$\hat{x}_{t+1|t} = \mathcal{F}_1\left(X_t; \theta_t\right). \tag{2.7}$$

Для спрощення запису надалі поряд із повним позначенням $\hat{x}_{t+h|t}$ може використовуватися скорочене позначення x_{t+1}, якщо момент формування прогнозу однозначно визначається контекстом.

Принциповою умовою коректності прогнозування є відсутність витоку майбутньої інформації. Для формування прогнозу значення x_{t+1} допускається використання лише тих спостережень, які були доступні до моменту часу t:

$$X_t \subseteq \{x_1, x_2, \ldots, x_t\}. \tag{2.8}$$

У загальному випадку параметри прогнозного оператора можуть адаптивно змінюватися після надходження нових даних. Нехай V_t множина локальних навчальних або валідаційних позицій, доступних до моменту часу *t*. Тоді вибір параметрів моделі можна подати як задачу мінімізації емпіричного ризику:

$$\theta_t^* = \arg\min_{\theta\in\Theta} \frac{1}{|V_t|} \sum_{\tau\in V_t} \ell\left(x_{\tau+1}, \mathcal{F}_1(X_\tau;\theta)\right). \tag{2.9}$$

де Θ – множина допустимих значень параметрів; $l(\cdot)$ – функція втрат; $|V_t|$ – кількість позицій у локальній вибірці.

Похибка прогнозу визначається як різниця між фактичним і прогнозованим значеннями:

$$e_{t+1} = x_{t+1} - \hat{x}_{t+1|t}. \tag{2.10}$$

Для біржових часових рядів доцільно розглядати не лише мінімізацію середньої похибки на вибраних даних, а й чутливість алгоритму до локальних збурень. Модель, яка забезпечує високу точність на окремому часовому фрагменті, але різко змінює результат у разі додавання одного нового спостереження, має обмежену практичну цінність. Тому важливо аналізувати поведінку прогнозних значень у процесі послідовного оновлення даних.

Для коректної оцінки алгоритмів прогнозування часових рядів використовується режим послідовного тестування з ковзною точкою формування прогнозу – walk-forward validation, або rolling forecasting origin. На кожному кроці прогноз формується лише за попередніми значеннями ряду, після чого порівнюється з фактичним наступним значенням. Такий режим відтворює реальну ситуацію застосування алгоритму та запобігає використанню майбутньої інформації. Методологію оцінювання прогнозів із ковзною точкою походження детально розглянуто Р. Дж. Хайндманом і Дж. Атанасопулосом [27].

Послідовність прогнозних значень у режимі walk-forward можна подати у вигляді:

$$\hat{x}_{t+1|t} = \mathcal{F}_1\left(X_t;\theta_t\right), \qquad t = t_0, t_0+1, \ldots, T-1. \tag{2.11}$$

де t_0 – перший момент часу, для якого сформовано локальне вікно достатньої довжини; $N = T - t_0$ кількість оцінених прогнозів. Загальна схема алгоритму послідовного тестування наведена на рис. 1.

Для порівняння прогнозних алгоритмів доцільно використовувати сукупність взаємодоповнювальних метрик. Основними критеріями в експериментальній частині роботи є середня абсолютна похибка, середньоквадратична похибка, корінь із середньоквадратичної похибки та середня абсолютна відсоткова похибка.

Середня абсолютна похибка визначається за формулою:

$$MAE = \frac{1}{N}\sum_{t=t_0}^{T-1}\left|e_{t+1}\right|. \tag{2.12}$$

Середньоквадратична похибка має вигляд:

$$MSE = \frac{1}{N}\sum_{t=t_0}^{T-1}e_{t+1}^2 . \tag{2.13}$$

Корінь із середньоквадратичної похибки визначається як:

$$RMSE = \sqrt{\frac{1}{N}\sum_{t=t_0}^{T-1}e_{t+1}^2} . \tag{2.14}$$

Середня абсолютна відсоткова похибка обчислюється за формулою:

$$MAPE = \frac{100\%}{N}\sum_{t=t_0}^{T-1}\left|\frac{e_{t+1}}{x_{t+1}}\right|. \tag{2.15}$$

Метрика *MAE* характеризує середній абсолютний масштаб похибки та є стійкішою до поодиноких значних відхилень, ніж *MSE* і *RMSE*. Метрики *MSE* та *RMSE* надають більшої ваги великим похибкам, що важливо для виявлення нестійких прогнозних значень. Метрика *MAPE* забезпечує інтерпретацію похибки у відсотках і є зручною для порівняння результатів прогнозування цін різного масштабу. Водночас її використання потребує обережності для рядів, значення яких можуть наближатися до нуля.

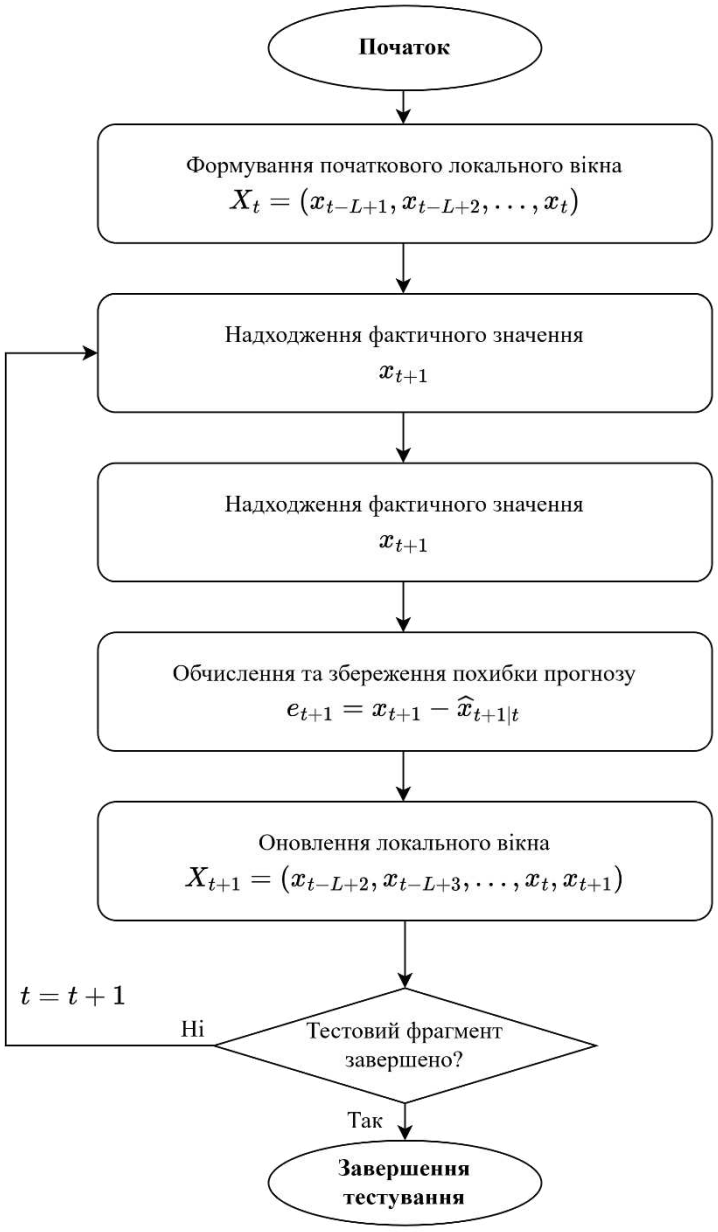


Рис. 2.1. Алгоритм послідовного тестування.

Для міжрядового порівняння додатково може застосовуватися масштабована середня абсолютна похибка *MASE*, запропонована Р. Дж. Хайндманом та А. Б. Келером [110]:

$$MASE = \frac{\frac{1}{N}\sum_{t=t_0}^{T-1}\left|e_{t+1}\right|}{\frac{1}{T_{\text{tr}}-1}\sum_{t=2}^{T_{\text{tr}}}\left|x_t - x_{t-1}\right|}. \qquad (2.16)$$

де T_{tr} – кількість спостережень у навчальному фрагменті.

Перевагою *MASE* є можливість порівняння прогнозів для часових рядів із різними масштабами. На відміну від відсоткових показників, ця метрика не

втрачає стійкості для значень, близьких до нуля. Обмеження метрик прогнозної точності та доцільність використання масштабованих похибок розглянуто в роботі [110] і в узагальнювальній праці [109].

Оскільки для біржових задач важливим є не лише абсолютне значення прогнозу, а й правильне визначення напряму зміни ціни, додатково може використовуватися показник точності прогнозування напряму руху:

$$DA = \frac{100\%}{N} \sum_{t=t_0}^{T-1} I\left[\left(x_{t+1} - x_t \right) \left(\hat{x}_{t+1|t} - x_t \right) > 0 \right]. \qquad (2.17)$$

де $I[\cdot]$ – індикаторна функція, яка набуває значення 1, якщо умова виконується, і 0 – в іншому випадку.

Показник *DA* доцільно використовувати як допоміжний критерій. Високе значення точності визначення напряму не гарантує малої абсолютної похибки прогнозу, а мінімальна абсолютна похибка не завжди забезпечує правильне визначення напряму зміни котирування. Тому обґрунтоване порівняння алгоритмів має ґрунтуватися на сукупності критеріїв.

Загальна постановка задачі (2.8) не обмежує клас прогнозних моделей. Однак для коротких локальних фрагментів біржового ряду доцільно дослідити можливість використання поліноміальної екстраполяції. Її перевагами є відсутність потреби у формуванні великої навчальної вибірки, обчислювальна простота та можливість побудови прогнозу без тривалого ітераційного налаштування параметрів.

Водночас точність поліноміальної екстраполяції істотно залежить від кількості використаних вузлів і, відповідно, від степеня інтерполяційного полінома. Надмірне збільшення степеня може спричиняти неточність прогнозу та різке посилення впливу шумових складових. Тому використання лише одного наперед заданого полінома не завжди є обґрунтованим.

У роботі Ю. В. Турбала, Г. Г. Шліхти, М. Ю. Турбала та Б. Ю. Турбала запропоновано формувати множину поліноміальних прогнозів, отриманих для різних степенів полінома, та аналізувати їхню послідовність для вибору прогнозного значення [128]. Подальший розвиток підходу пов’язаний з

усередненням інформативної частини поліноміальної екстраполяційної послідовності [129], використанням PPS для аналізу ступеня стохастичності процесу [118] та кластеризацією поліноміальних прогнозних значень [127].

Нехай на основі локального вікна X_t будується сукупність однокрокових поліноміальних прогнозів:

$$p_m = \mathcal{P}_m\left(X_t^{(m)}\right), \qquad m = 1,2,\ldots,M, \qquad M \le L, \tag{2.18}$$

де p_m – прогноз значення x_{t+1}, отриманий на основі останніх m спостережень; $X_t^{(m)}$ – підвікно з m останніх елементів локального фрагмента X_t; $\mathcal{P}_m$ – оператор поліноміальної екстраполяції; M – максимальна кількість використаних вузлів.

Індекс m визначає кількість останніх спостережень, використаних для побудови прогнозу. Відповідний інтерполяційний поліном має степінь, що не перевищує $m-1$. У програмній реалізації може застосовуватися еквівалентна нульова індексація прогнозів $P_0, P_1, \ldots, P_{M-1}$.

Упорядкований набір прогнозних значень утворює послідовність поліноміальних прогнозів – Polynomial Prediction Sequence (PPS):

$$PPS(X_t) = \left(p_1, p_2, \ldots, p_M\right). \tag{2.19}$$

Кожний елемент послідовності (2.20) відображає прогноз, отриманий для одного й того самого моменту часу $t+1$, але за різної кількості використаних вузлів. Тому PPS можна розглядати не лише як набір альтернативних прогнозів, а і як структуроване подання реакції поліноміальної моделі на зміну її параметра m.

Остаточний прогноз формується шляхом застосування оператора адаптивного агрегування:

$$\hat{x}_{t+1|t} = \mathcal{A}\left(PPS(X_t); \eta_t\right), \tag{2.20}$$

де $\mathcal{A}$ – оператор вибору або агрегування прогнозних значень; η_t – параметри адаптивної процедури.

Одним із узагальнених варіантів оператора (2.21) є зважене агрегування:

$$\hat{x}_{t+1|t} = \sum_{m=1}^{M} \alpha_{m,t} p_m + \delta_t, \qquad \alpha_{m,t} \ge 0, \qquad \sum_{m=1}^{M} \alpha_{m,t} = 1, \tag{2.21}$$

де $\alpha_{m,t}$ – ваговий коефіцієнт прогнозу p_m; δ_t – коригувальна складова.

Формула (2.22) є узагальненням кількох можливих стратегій:

1) вибір одного прогнозу – один коефіцієнт $\alpha_{m,t}$ дорівнює 1, а решта дорівнюють 0;
2) арифметичне усереднення – ваги вибраної підмножини прогнозів є однаковими;
3) адаптивне зважування – коефіцієнти визначаються на основі внутрішніх критеріїв узгодженості PPS;
4) нейромережеве агрегування – ваги прогнозів формуються навченим селектором;
5) кластерний аналіз PPS – остаточне значення визначається за центральною характеристикою інформативного кластера.

Відповідно, задачу дослідження можна сформулювати таким чином.

Нехай задано скалярний біржовий часовий ряд:

$$\mathcal{X}_{1:T} = \{x_t\}_{t=1}^{T}$$

і локальне вікно спостережень:

$$X_t = (x_{t-L+1}, x_{t-L+2}, \ldots, x_t)$$

Необхідно розробити метод короткострокового прогнозування, який для кожного моменту часу *t*:

1) формує послідовність поліноміальних прогнозів $PPS(X_t)$;
2) аналізує структурні властивості отриманої послідовності;
3) визначає інформативну підмножину прогнозних значень або адаптивні вагові коефіцієнти;
4) обчислює остаточний прогноз $\hat{x}_{t+1|t}$;
5) допускає перевірку точності в режимі walk-forward без витоку майбутньої інформації.

У формалізованому вигляді цільова задача полягає у виборі оператора $\mathcal{A}^*$, який мінімізує функціонал прогнозної похибки:

$$\mathcal{A}^* = \arg\min_{\mathcal{A}\in\mathfrak{A}} \frac{1}{N}\sum_{t=t_0}^{T-1} \ell\left(x_{t+1}, \mathcal{A}\left(PPS(X_t);\eta_t\right)\right), \tag{2.22}$$

де $\mathfrak{A}$ – множина допустимих алгоритмів вибору або агрегування прогнозів PPS.

При цьому алгоритм повинен задовольняти обмеження інформаційної доступності:

$$PPS(X_t) = PPS\left(x_{t-L+1},\ldots,x_t\right), \tag{2.23}$$

тобто під час формування прогнозу для моменту $t+1$ не можуть використовуватися значення $x_{t+1}, x_{t+2},\ldots$.

2.2. Поліноміальна екстраполяція як основа локального прогнозування

Часові ряди біржової динаміки характеризуються мінливістю статистичних властивостей, наявністю шумової складової, короткочасних трендів і локальних структурних змін. Унаслідок цього побудова єдиної глобальної апроксимувальної функції на тривалому інтервалі спостереження не завжди забезпечує достатню точність короткострокового прогнозування. Особливо виразно ця проблема проявляється під час аналізу внутрішньоденних котирувань, коли актуальність віддалених у часі даних може швидко знижуватися.

У таких умовах доцільно використовувати локальну екстраполяцію, за якої прогнозне значення визначається на основі обмеженої кількості останніх спостережень. Поліном у цьому разі не розглядається як глобальна модель біржового процесу. Він виконує роль локального інструменту відтворення регулярної компоненти ряду в околі поточного моменту часу та її продовження на один крок уперед.

Методологічні засади локального поліноміального моделювання систематизовано у праці J. Fan і I. Gijbels [119]. Автори розглядають локальні поліноміальні моделі як гнучкий засіб аналізу даних, що дає змогу враховувати зміну властивостей досліджуваного процесу в різних областях його визначення. Важливим історичним прикладом використання локальних поліномів є метод A. Savitzky і M. Golay [120], у якому поліноміальна апроксимація на рухомому вікні

застосовується для згладжування та чисельного диференціювання експериментальних даних. Метод, запропонований у дисертаційній роботі, відрізняється від локальної поліноміальної регресії: поліном будується не методом найменших квадратів, а як інтерполяційний поліном за останніми значеннями часового ряду. Проте спільною є ідея локального аналізу даних у межах рухомого вікна.

Застосування поліномів високого степеня потребує обережності. Як зазначено у працях з теорії апроксимації, зі збільшенням кількості рівновіддалених вузлів глобальна поліноміальна інтерполяція може супроводжуватися небажаними осциляціями. Цей ефект відомий як феномен Рунге. Для задачі локального прогнозування це означає, що механічне збільшення степеня полінома не гарантує зменшення похибки екстраполяції. Обґрунтованішим є формування множини альтернативних локальних прогнозів із подальшим інтелектуальним аналізом їхньої внутрішньої структури.

Окремий напрям досліджень локальної поліноміальної екстраполяції розвивається у працях українських науковців. У роботі Ю. Турбала, А. Бомби, М. Турбала та А. Дріві [124] запропоновано обчислювальну процедуру прогнозування за допомогою інтерполяційних поліномів довільного степеня на рівномірній часовій сітці. У статті Ю. Турбала, Г. Шліхти, М. Турбала та Б. Турбала [128] розглянуто алгоритм вибору оптимального степеня полінома шляхом аналізу послідовності прогнозів. Подальший розвиток цього підходу пов’язаний із формуванням послідовності всіх допустимих поліноміальних прогнозів, адаптивним визначенням її інформативної частини, усередненням прогнозних значень і кластерним аналізом [125;127-129].

У сучасних прикладних дослідженнях поліноміальна інтерполяція також використовується в комбінації з методами зваженого усереднення. Зокрема, A. Flores та співавтори [121] застосували таке поєднання для прогнозування часових рядів забруднення повітря за умов обмеженого обсягу вхідних даних. Хоча предметна область відрізняється від фінансових ринків, ця робота підтверджує

доцільність комбінування локальної поліноміальної екстраполяції з процедурами адаптивного агрегування прогнозів.

В українській науковій літературі екстраполяційні методи розглядаються також як складова статистичного аналізу часових рядів і прикладного прогнозування економічних процесів. Теоретичні основи аналізу динамічних рядів систематизовано, зокрема, у навчальному посібнику А. Єріної та О. Мазуренко [122]. О. Сенишин [123] дослідила застосування екстраполяційних методів до прогнозування економічних показників і наголосила на необхідності врахування тенденцій розвитку процесу під час вибору прогнозної моделі.

У межах дисертаційного дослідження локальна поліноміальна екстраполяція використовується як базовий механізм формування послідовності поліноміальних прогнозів – Polynomial Predictions Sequence, або PPS. На відміну від підходу, за якого визначається єдиний наперед заданий степінь полінома, запропоновано обчислювати впорядковану множину прогнозних значень для різних глибин передісторії. Це дає змогу перейти від вибору окремої моделі до аналізу структури сукупності локальних прогнозів.

Нехай задано дискретний часовий ряд біржової динаміки

$$F_n = \{f_1, f_2, \ldots, f_n\}, \tag{2.24}$$

де f_i – значення досліджуваного показника в момент часу t_i. Як досліджуваний показник можуть використовуватися ціна закриття, логарифм ціни, дохідність, значення біржового індикатора або інший числовий параметр ринкової динаміки.

Припустимо, що спостереження задано на рівномірній часовій сітці:

$$t_i = i\Delta, \qquad f_i = f(t_i), \tag{2.25}$$

де Δ – крок дискретизації часового ряду.

Необхідно визначити прогноз наступного значення $\hat{f}_{n+1}$. Для побудови локальної прогнозної моделі розглянемо останні *m* значень часового ряду:

$$W_n^{(m)} = \{f_{n-m+1}, f_{n-m+2}, \ldots, f_n\}, \qquad 1 \le m \le M \le n, \tag{2.26}$$

де *m* – кількість останніх спостережень, використаних для побудови конкретного прогнозу; *M* – максимальна кількість останніх спостережень, яку використано

під час формування послідовності поліноміальних прогнозів.; $W_n^{(m)}$ локальне рухоме вікно довжини m.

За значеннями з вікна $W_n^{(m)}$ будується інтерполяційний $P_{m-1}^{(n)}(t)$ поліном степеня не вище ніж $m-1$, який задовольняє умови

$$Q_{m-1}^{(n)}\left(t_{n-j}\right) = f_{n-j}, \qquad j = 0,1,\ldots,m-1. \tag{2.27}$$

Локальний поліноміальний прогноз на один крок уперед визначається як значення побудованого полінома в точці t_{n+1}:

$$p_m = \hat{f}_{n+1}^{(m)} = Q_{m-1}^{(n)}\left(t_{n+1}\right). \tag{2.28}$$

Таким чином, індекс m одночасно визначає:

1) кількість останніх значень ряду, використаних для побудови локальної моделі;
2) ширину локального вікна $W_n^{(m)}$;
3) максимально можливий степінь інтерполяційного полінома, що дорівнює $m-1$;
4) конкретне прогнозне значення p_m.

Усі прогнози p_m належать до одного моменту часу t_{n+1}, але формуються за різних обсягів передісторії. Саме ця властивість створює основу для подальшого порівняльного аналізу прогнозних значень.

Інтерполяційний поліном $Q_{m-1}^{(n)}\left(t_n\right)$ можна подати у формі Лагранжа:

$$Q_{m-1}^{(n)}(t) = \sum_{j=0}^{m-1} f_{n-j} L_j^{(m)}(t). \tag{2.29}$$

де $L_j^{(m)}(t)$ – базисні поліноми Лагранжа:

$$L_j^{(m)}(t) = \prod_{\substack{\ell=0 \\ \ell \neq j}}^{m-1} \frac{t - t_{n-\ell}}{t_{n-j} - t_{n-\ell}}. \tag{2.30}$$

Для визначення однокрокового прогнозу необхідно обчислити значення базисних поліномів у точці t_{n+1}. Оскільки часові відліки є рівновіддаленими, виконується співвідношення

$$L_j^{(m)}(t_{n+1}) = (-1)^j \binom{m}{j+1}. \tag{2.31}$$

Підставивши (2.32) у (2.30), одержимо спрощену формулу локального поліноміального прогнозу:

$$p_m = \hat{f}_{n+1}^{(m)} = \sum_{k=1}^{m} (-1)^{k-1} \binom{m}{k} f_{n-k+1}. \tag{2.32}$$

Формула (2.33) є ключовою для формування PPS. Вона дає змогу обчислювати прогноз без явного визначення коефіцієнтів інтерполяційного полінома та без розв'язування системи лінійних алгебраїчних рівнянь. Відповідна формула використана у працях, присвячених аналізу послідовностей поліноміальних прогнозів і кластеризації їхніх елементів [124; 125; 127-129].

Необхідно врахувати, що спрощена формула (2.32) є коректною саме для рівномірної часової сітки. Для внутрішньоденних біржових даних це означає, що PPS доцільно формувати окремо для послідовних торгових інтервалів однакової тривалості. Якщо у вихідних даних є пропущені часові відліки, їх потрібно попередньо опрацювати або застосувати загальну формулу Лагранжа (2.29) із фактичними значеннями t_i.

Важливою властивістю локальної поліноміальної екстраполяції є її безпосередній зв'язок зі скінченними різницями. Визначимо зворотну скінченну різницю першого порядку:

$$\nabla f_n = f_n - f_{n-1}. \tag{2.33}$$

Скінченна різниця *r*-го порядку визначається рекурентно:

$$\nabla^r f_n = \nabla^{r-1} f_n - \nabla^{r-1} f_{n-1}, \qquad r \geq 2. \tag{2.34}$$

Еквівалентний явний запис має вигляд

$$\nabla^r f_n = \sum_{j=0}^{r} (-1)^j \binom{r}{j} f_{n-j}. \tag{2.35}$$

На основі інтерполяційної формули Ньютона прогноз p_m можна подати як суму скінченних різниць:

$$p_m = f_n + \nabla f_n + \nabla^2 f_n + \ldots + \nabla^{m-1} f_n. \tag{2.36}$$

Звідси випливає рекурентне співвідношення:

$$p_{m+1} = p_m + \nabla^m f_n. \qquad (2.37)$$

Рекурентний запис (2.39) має важливе змістове значення: перехід від прогнозу p_m до прогнозу p_{m+1} визначається скінченною різницею m-го порядку. Отже, послідовність поліноміальних прогнозів безпосередньо відображає структуру локальних змін часового ряду різних порядків. Це дає підстави розглядати PPS не лише як набір альтернативних прогнозів, а і як окремий інформаційний об'єкт для інтелектуального аналізу. Рекурентний зв'язок між сусідніми елементами PPS і скінченними різницями наведено також у працях, присвячених кластерному аналізу PPS та оцінюванню рівня стохастичності процесів [127;118].

Для кожного допустимого значення mобчислюється прогноз p_m. У результаті утворюється впорядкована послідовність поліноміальних прогнозів, або PPS

$$P_n^{(M)} = \left(p_1, p_2, \ldots, p_M \right). \qquad (2.38)$$

З урахуванням формули (2.39) повний запис PPS має вигляд

$$P_n^{(M)} = \left(f_n, 2f_n - f_{n-1}, 3f_n - 3f_{n-1} + f_{n-2}, \ldots, \sum_{k=1}^{M} (-1)^{k-1} \binom{M}{k} f_{n-k+1} \right). \qquad (2.39)$$

Кожний елемент p_m є локальним прогнозом того самого майбутнього значення f_{n+1}, але формується на основі іншого обсягу попередніх даних та іншого степеня полінома. Таким чином, PPS – це впорядкована сукупність альтернативних локальних прогнозів, отриманих за допомогою поліноміальних моделей із різною шириною вікна та різним максимальним степенем полінома.:

$$P_n^{(M)} = p_m = \mathcal{E}_m \left(f_{n-m+1}, \ldots, f_n \right), \qquad m = 1, 2, \ldots, M. \qquad (2.40)$$

де $\mathcal{E}_m$ – оператор локальної поліноміальної екстраполяції за m останніми значеннями.

У низці практичних задач доцільно аналізувати не всю послідовність, а її робочу підпослідовність:

$$P_{n,[m_{\min},M]} = \left(p_{m_{\min}}, p_{m_{\min}+1}, \ldots, p_M \right), \qquad 1 \le m_{\min} \le M . \tag{2.41}$$

Якщо прогноз $p_1 = f_n$ використовується лише як базова наївна модель, для подальшого аналізу можна прийняти $m_{\min} = 2$.

Значення M визначає максимальну ширину локального вікна спостережень. Надто мале значення M може призвести до втрати інформативних локальних закономірностей. Надто велике значення M розширює часовий інтервал аналізу та підвищує вплив застарілих спостережень, а також може посилювати реакцію прогнозів високих порядків на шумову складову. Тому параметр M має узгоджуватися з частотою дискретизації, довжиною доступної вибірки та характером досліджуваного біржового процесу.

У працях, присвячених PPS, показано, що елементи такої послідовності можуть утворювати області концентрації, локальні екстремуми, монотонні фрагменти та віддалені значення. Тому вибір остаточного прогнозу не обов'язково має зводитися до використання одного фіксованого степеня полінома. Альтернативою є аналіз структурних властивостей PPS, адаптивний вибір її інформативної підмножини та агрегування узгоджених прогнозних значень [125; 127; 128; 129].

Послідовність $P_n^{(M)}$ можна сформувати двома способами:

1) безпосереднім обчисленням кожного прогнозу за формулою (2.32);
2) рекурентним обчисленням на основі таблиці скінченних різниць і формули (2.39).

Для програмної реалізації доцільно використовувати другий спосіб, оскільки він розкриває зв'язок між елементами PPS та скінченними різницями й не потребує багаторазового обчислення біноміальних коефіцієнтів.

Алгоритм 2.1 – Формування послідовності поліноміальних прогнозів

Вхідні дані:

$(f_{n-M+1}, \ldots, f_n)$ – останні (M) значень часового ряду; (M) – максимальна ширина локального вікна спостережень.

Вихідні дані:

$P_n^{(M)} = (p_1, p_2, \ldots, p_M)$ – послідовність поліноміальних прогнозів.

1. Сформувати допоміжний масив (d), записавши останні (M) значень часового ряду у зворотному порядку:

   $d[j] \leftarrow f_{n-j}, \;\; j = 0, 1, ..., M-1$

2. Визначити перший елемент PPS:

   $p[1] \leftarrow f_n$.

3. Для кожного значення $r = 1, 2, \ldots, M-1$ виконати такі дії:

   3.1. Для кожного значення $j = 0, 1, \ldots, M-r-1$ обчислити скінченні різниці r-го порядку: $d[j] \leftarrow d[j] - d[j+1]$.

   3.2. Після завершення внутрішнього циклу перший елемент допоміжного масиву містить скінченну різницю r-го порядку: $d[0] = \nabla^r f_n$.

   3.3. Обчислити наступний елемент PPS: $p[r+1] \leftarrow p[r] + d[0]$.

4. Повернути сформовану послідовність: $P_n^{(M)} = (p_1, p_2, \ldots, p_M)$.

У наведеному алгоритмі $d[j]$ позначає j-й елемент допоміжного масиву, який використовується для послідовного обчислення скінченних різниць. На початковому етапі масив містить останні M значень часового ряду у зворотному порядку. Після виконання зовнішнього циклу з номером r значення $d[0]$ дорівнює скінченній різниці $\nabla^r f_n$. Це дає змогу рекурентно обчислити всі елементи PPS без явної побудови інтерполяційних поліномів. Схематично алгоритм формування послідовності поліноміальних прогнозів наведений на рис. 2.2.

Для теоретичного аналізу властивостей локальної поліноміальної екстраполяції розглянемо допоміжну неперервну функцію $\varphi(t)$, значення якої у вузлах часової сітки збігаються зі значеннями регулярної складової досліджуваного процесу:

$$f_i = \varphi(t_i).$$

Нехай функція $\varphi(t)$ має неперервну похідну m-го порядку на інтервалі $[t_{n-m+1}, t_{n+1}]$.

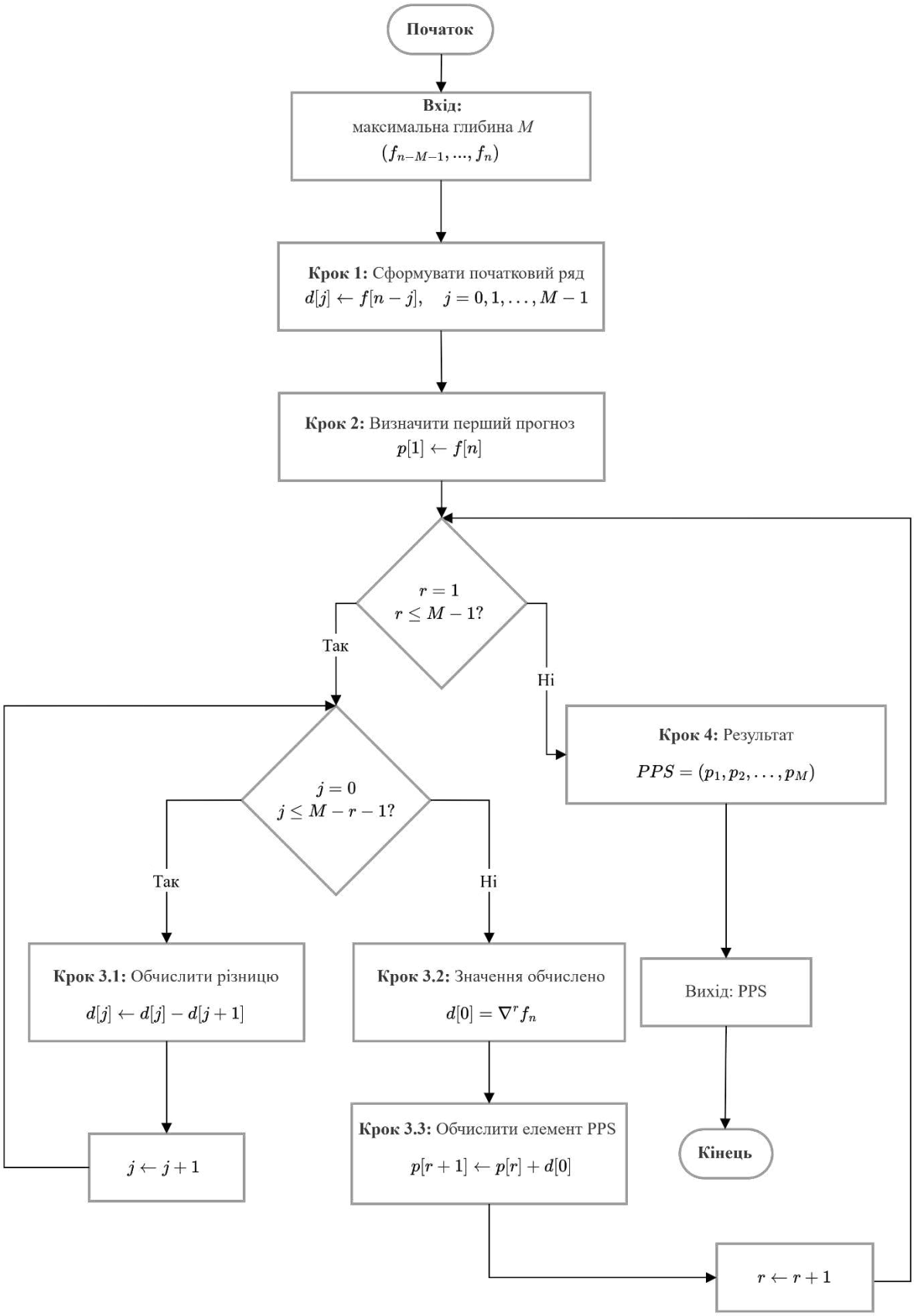


Рис. 2.2 Схема алгоритму формування послідовності поліноміальних прогнозів

За останніми m значеннями $\varphi(t_{n-m+1}), \varphi(t_{n-m+2}),\ldots,\varphi(t_n)$ побудуємо локальний інтерполяційний поліном $Q_{m-1}^{(n)}(t)$ степеня не вище ніж $m-1$. Його значення у точці t_{n+1} визначає локальний поліноміальний прогноз:

$$p_m = Q_{m-1}^{(n)}(t_{n+1}).$$

Відповідно до формули залишкового члена інтерполяції Лагранжа існує точка $\xi_m \in (t_{n-m+1}, t_{n+1})$, для якої виконується співвідношення [137]

$$\varphi(t_{n+1}) - p_m = \frac{\varphi^{(m)}(\xi_m)}{m!}\prod_{j=0}^{m-1}(t_{n+1} - t_{n-j}). \qquad (2.42)$$

Для рівномірної часової сітки з кроком Δ $t_i = i\Delta$ виконується рівність

$$t_{n+1} - t_{n-j} = (j+1)\Delta.$$

Отже,

$$\prod_{j=0}^{m-1}(t_{n+1} - t_{n-j}) = m!\Delta^m,$$

а залишковий член локальної екстраполяції набуває вигляду

$$\varphi(t_{n+1}) - p_m = \varphi^{(m)}(\xi_m)\Delta^m. \qquad (2.43)$$

Одержане співвідношення має теоретичний характер. Воно описує залишковий член локальної екстраполяції гладкої регулярної компоненти процесу, але не враховує шумову складову реального біржового часового ряду. Для спостережуваних котирувань доцільно використовувати модель

$$f_i = \varphi(t_i) + \varepsilon_i. \qquad (2.44)$$

де ε_i – випадкове збурення. У такому разі фактична похибка прогнозування визначається як властивостями локальної апроксимації, так і впливом шумової складової.

2.3 Аналіз ступеня стохастичності процесів на основі послідовності поліноміальних прогнозів

Нехай відомі значення деякого часового ряду $f_n, f_{n-1}, \ldots, f_{n-m+1}$. Якщо необхідно для кожного локального фрагмента ряду довжини m будувати однокроковий поліноміальний прогноз, то у дослідженні [128] показано, що таке прогнозне значення може бути знайдене без явного знаходження коефіцієнтів інтерполяційного полінома, а на основі прямої біноміальної формули. Це істотно спрощує обчислення для всіх значень m і дає можливість швидко формувати повну послідовність прогнозів.

Прогнозне значення, побудоване за m точками, визначається формулою

$$p_m = \sum_{k=1}^{m} (-1)^{k-1} C_m^k f_{n-k+1}. \qquad (2.45)$$

Послідовність $P = \{p_1, p_2, \ldots, p_m\}$ будемо називати послідовністю поліноміальних прогнозів (PPS). У [13] показано, що така послідовність може мати різну внутрішню структуру: тенденцію до збіжності, наявність локальних екстремумів, інтервали концентрації та віддалені значення. Саме ці структурні особливості роблять PPS придатною не лише для побудови прогнозу, а й для аналізу характеру самого процесу. А тому будемо використовувати PPS для дослідження стохастичності процесу. В роботі [148] визначено умови збіжності PPS . При цьому в ході чисельних експериментів було виявлено значну чутливість послідовності PPS від наявності стохастичних збурень вхідних даних. Ця особливість і є головною ідеєю досліджень даної роботи.

Для аналізу ступеня стохастичності будемо розглядати модельний часовий ряд, у якому детермінована компонента поєднується з випадковим збуренням заданої інтенсивності. Нехай базовий процес описується деякою детермінованою функцією $g(i)$, визначеною у вузлах рівномірної сітки. Тоді досліджувана послідовність задається співвідношенням

$$f_i = g(i) + \varepsilon \xi_i,$$

де ξ_i – випадкова величина, а параметр ε визначає рівень стохастичного збурення. У ролі базових детермінованих функцій будемо використовувати

декілька типів залежностей, які відрізняються характером локальної геометрії: монотонно зростаючі, осциляційні та змішані. Це дає змогу оцінити, наскільки втрата збіжності PPS залежить не лише від інтенсивності шуму, а й від форми детермінованої складової.

Такий підхід дозволяє варіювати ступінь випадковості процесу контрольованим способом і, відповідно, аналізувати зміни поведінки PPS в широкому діапазоні умов.

Ключовим теоретичним результатом, на який спираються дослідження в цьому розділі роботи, є встановлений у дослідженні [148] зв'язок між збіжністю PPS і поведінкою відповідних скінченних різниць. Для того щоб послідовність прогнозів можна було використовувати як індикатор, необхідно чітко розуміти математичні умови її збіжності. Теоретичний аналіз, проведений авторами, встановлює прямий зв'язок між збіжністю PPS та поведінкою скінченних різниць високих порядків. Зокрема, доведено, що послідовність $\{p_m\}$ збігається при фіксованому n та $m \to \infty$ тоді і тільки тоді, коли існує скінченна границя m-тої скінченної різниці $\Delta^m f_n$.

Це спостереження дозволило запровадити нову таксономію функцій та даних, розділивши їх на специфічні класи стабільності:

1. Біноміально-стабільні функції (клас Ξ) – клас детермінованих функцій, для яких границя скінченних різниць існує і дорівнює нулю (або константі) при будь-якому обраному кроці дискретизації h. Представниками цього класу є алгебраїчні поліноми будь-яких степенів та показникові (експоненційні) функції.

2. Умовно біноміально-стабільні функції – це клас процесів, для яких збіжність послідовності поліноміальних прогнозів не є гарантованою, а залежить від вибору величини кроку дискретизації h. Прикладом є періодичні (тригонометричні) функції. При одних значеннях кроку h (наприклад, кратних періоду) PPS може збігатися, тоді як при інших – генерувати розбіжну послідовність зі зростаючою амплітудою коливань.

3. Скінченно біноміально-стабільні дані – практичний аналог для експериментальних вибірок обмеженої довжини, де перевіряється тенденція до збіжності на скінченному наборі m точок.

При наявності випадкових збурень послідовність PPS матиме вигляд:

$$p_m = \sum_{k=1}^{m}(-1)^{k-1}C_m^k f_{n-k+1} = \sum_{k=1}^{m}(-1)^{k-1}C_m^k (g_{n-k+1} + \varepsilon\xi_{n-k+1}),$$

де $g_{n-k+1} = g\big((n-k+1)h\big)$, h – крок сітки. Виконуючи перетворення, отримаємо:

$$\sum_{k=1}^{m}(-1)^{k-1}C_m^k (g_{n-k+1} + \varepsilon\xi_{n-k+1}) = \sum_{k=1}^{m}(-1)^{k-1}C_m^k g_{n-k+1} + \sum_{k=1}^{m}(-1)^{k-1}C_m^k \varepsilon\xi_{n-k+1}.$$

Таким чином

$$\tilde{p}_m\big(n+1\big) = p_m\big(n+1\big) + \varepsilon\sum_{k=1}^{m}(-1)^{k-1}C_m^k \xi_{n-k+1}, \qquad (2.46)$$

де

$$\tilde{p}_m\big(n+1\big) = \sum_{k=1}^{m}(-1)^{k-1}C_m^k (g_{n-k+1} + \varepsilon\xi_{n-k+1}),$$

$$p_m\big(n+1\big) = \sum_{k=1}^{m}(-1)^{k-1}C_m^k g_{n-k+1}.$$

Із співвідношення (2.48) бачимо, що збіжність стохастично збуреної послідовності за умови збіжності детермінованої складової залежить від збіжності послідовності поліноміальних прогнозів для стохастичної складової. Таким чином, будемо досліджувати PPS для рівномірно розподілених, незалежних випадкових величин. Параметр ε, очевидно, впливатиме лише на модуль відповідних збурень.

Для емпіричного аналізу було розроблено програмне середовище за допомогою мови Python, яке забезпечує генерацію випадкових часових рядів із рівномірним, нормальним та показниковим розподілами і побудову відповідних послідовностей поліноміальних прогнозів. Програмний комплекс обчислює елементи PPS для різних порядків, прирости між сусідніми прогнозами, кількість змін знаку та логарифмічні характеристики зростання амплітуди. Передбачено візуалізацію вихідного ряду, значень PPS, знакової структури, сусідніх приростів і темпу зростання модуля прогнозних значень. Такий інструментарій дає змогу

досліджувати наявність інтервалів монотонності, локальних екстремумів та розбіжної поведінки PPS, яка у роботі інтерпретується як індикатор стохастично-домінованої природи процесу.

Розглянемо послідовність незалежних рівномірно розподілених випадкових величин, що мають рівномірний розподіл на інтервалі $[0,1]$, (рис.2.3).

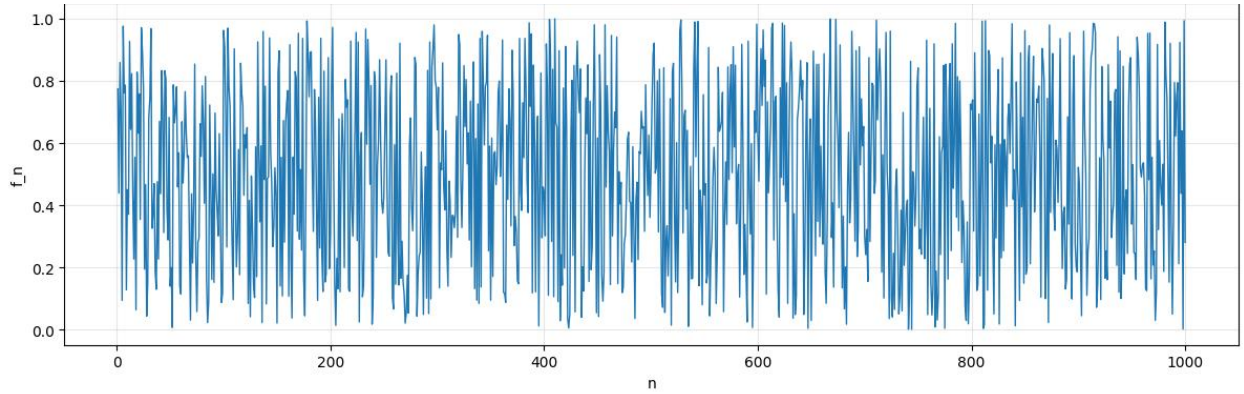
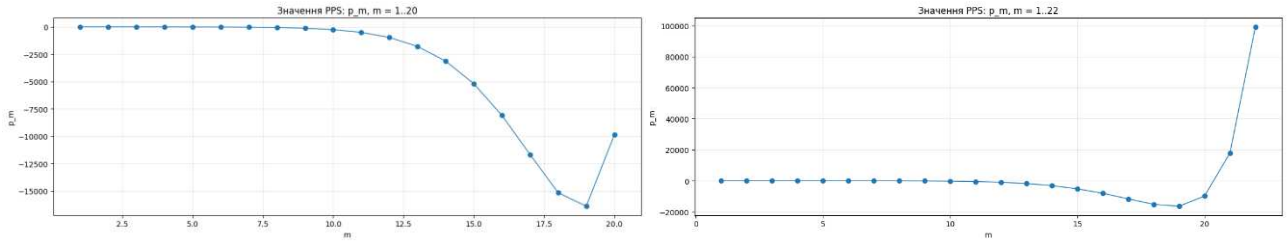
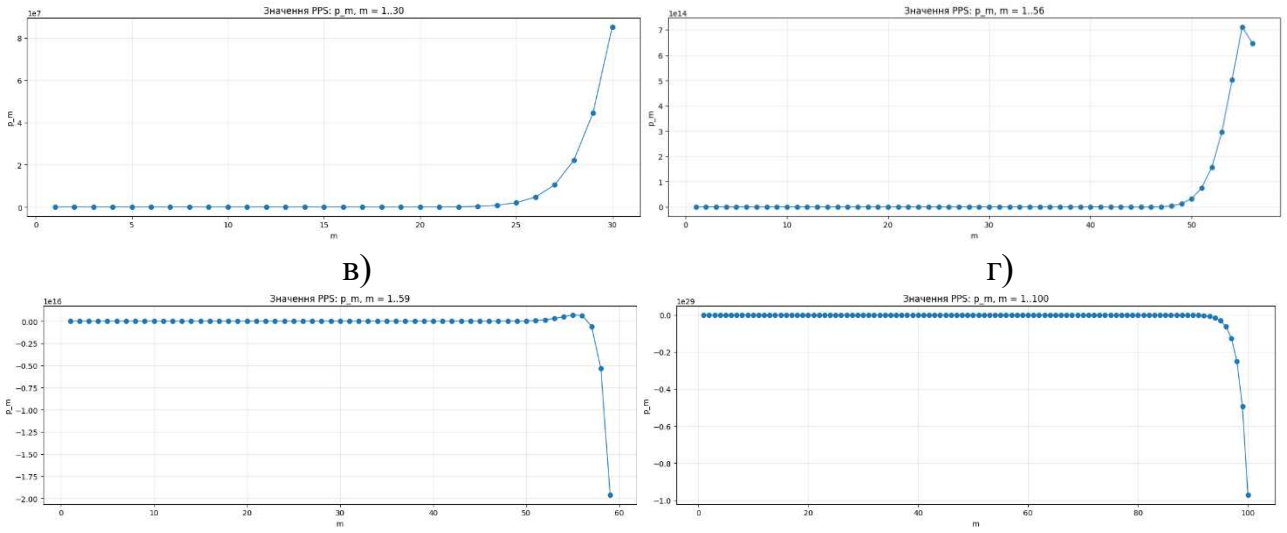


Рис. 2.3. Рівномірно розподілені випадкові величини, які мають рівномірний розподіл на інтервалі $[0,1]$, параметр $seed = 42$.

Для рівномірно розподілених випадкових даних на інтервалі $[0,1]$ послідовність поліноміальних прогнозів не має ознак збіжності у класичному розумінні, однак її поведінка не повністю хаотична у візуальному представленні. На графіках PPS можна спостерігати окремі інтервали монотонного зростання або спадання, а також локальні екстремуми, у яких напрям зміни прогнозних значень змінюється на протилежний, (рис. 2.4). Це свідчить про те, що навіть для випадкових вхідних даних біноміальна структура формули поліноміального прогнозу породжує певну внутрішню організацію послідовності p_m.



а) б)



в) г)

г) д)

Рис 2.4. Поведінка PPS для різних значень m: а) при $m = 20$; б) при $m = 22$; в) при $m = 30$; г) при $m = 56$; ґ) при $m = 59$; д) при $m = 100$.

Наведені вище чисельні результати показують, що послідовність поліноміальних прогнозів на стохастичних даних демонструє чітку детерміновану поведінку. Незважаючи на те, що вхідні дані є незалежними випадковими величинами, PPS має інтервали монотонності та локальні екстремуми. Наявність інтервалів монотонності, їх строге чергування (зростання, спадання, і т. д.) свідчить про стохастичну залежність елементів PPS. Очевидно, що легко підібрати детерміновану функцію, яка буде добре апроксимувати поведінку PPS на будь-якому інтервалі. Наприклад, це функції виду $y = ax^k \cos(x)$. Відповідні осцилюючі функції із зростаючою амплітудою розглядались в якості при дослідженні пірамідального методу прогнозування в роботі [152].

Аналогічні результати отримані для показникового та нормального розподілів (рис. 2.5-2.8).

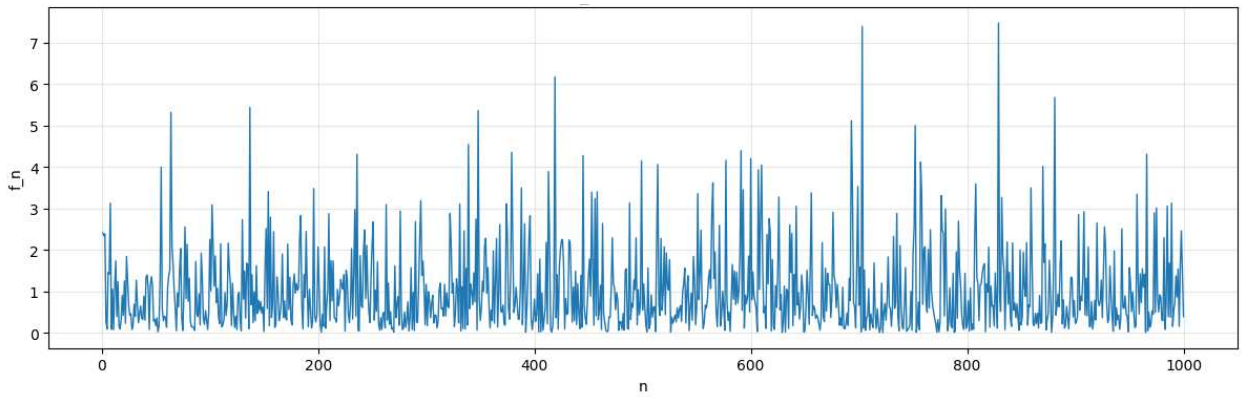
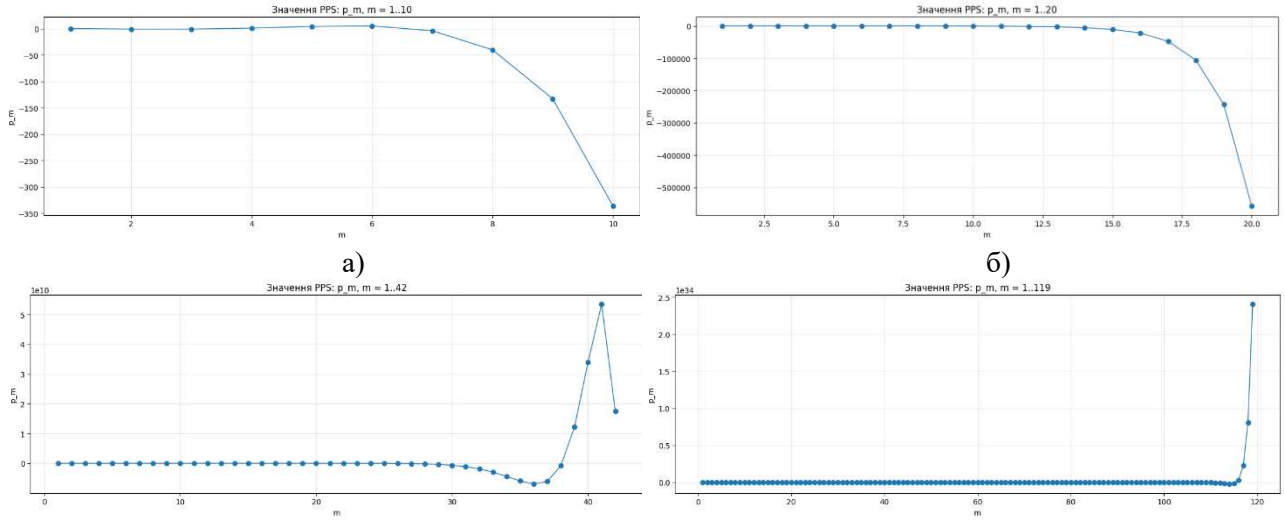


Рис. 2.5. Випадкові величини, які мають показниковий розподіл з параметром $\lambda = 1,\ seed = 42$.

На рис. 2.6 показано, як поводить себе послідовність PPS для показникового розподілу при різних значеннях m.



а) б)

в) г)

Рис 2.6. Поведінка PPS для різних значень m для показникового розподілу: а) при $m = 10$; б) при $m = 20$; в) при $m = 42$; г) при $m = 119$.

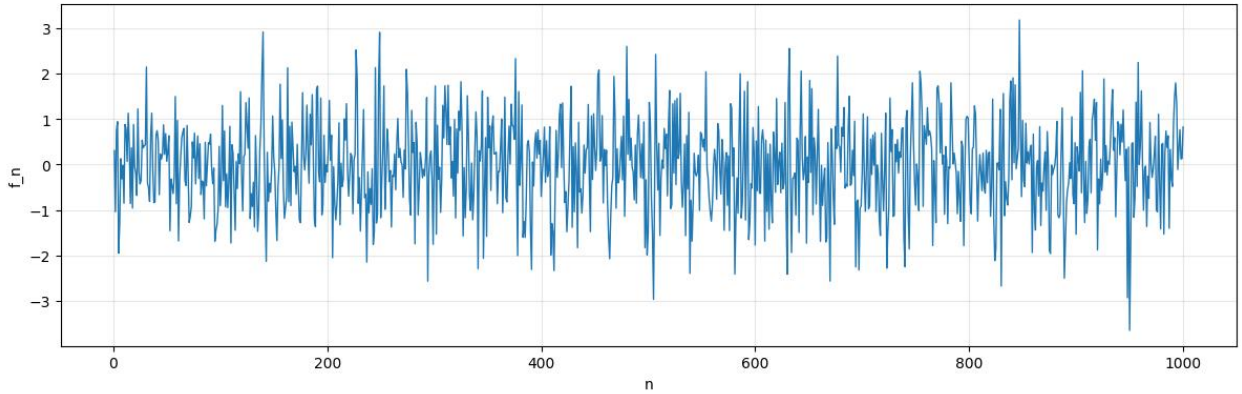
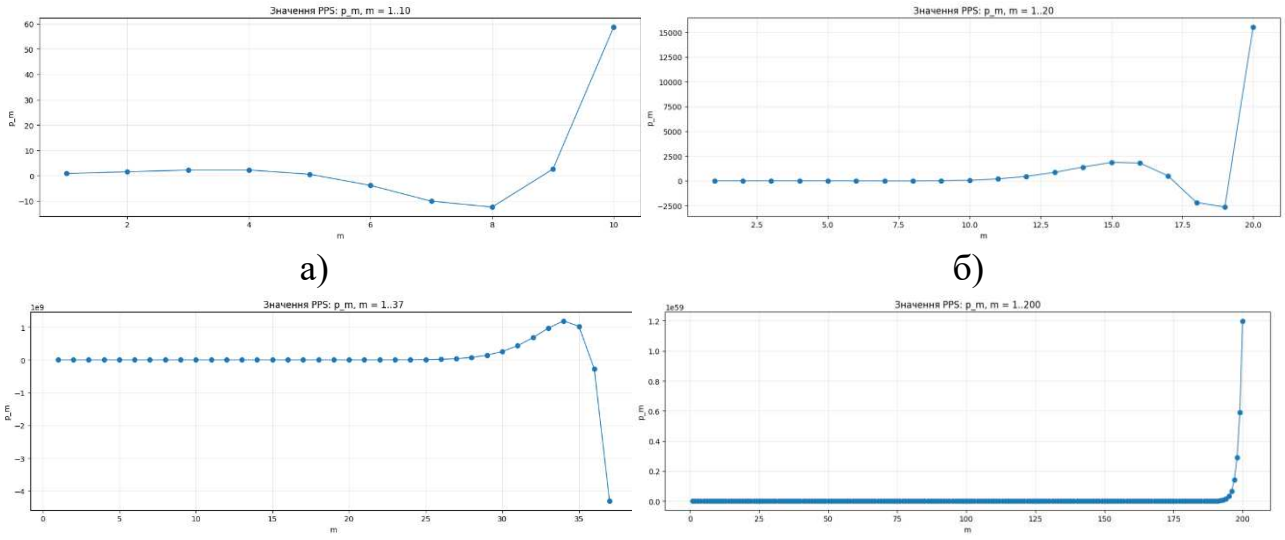


Рис. 2.7. Нормальний розподіл $N(0,1),\ seed = 42$.

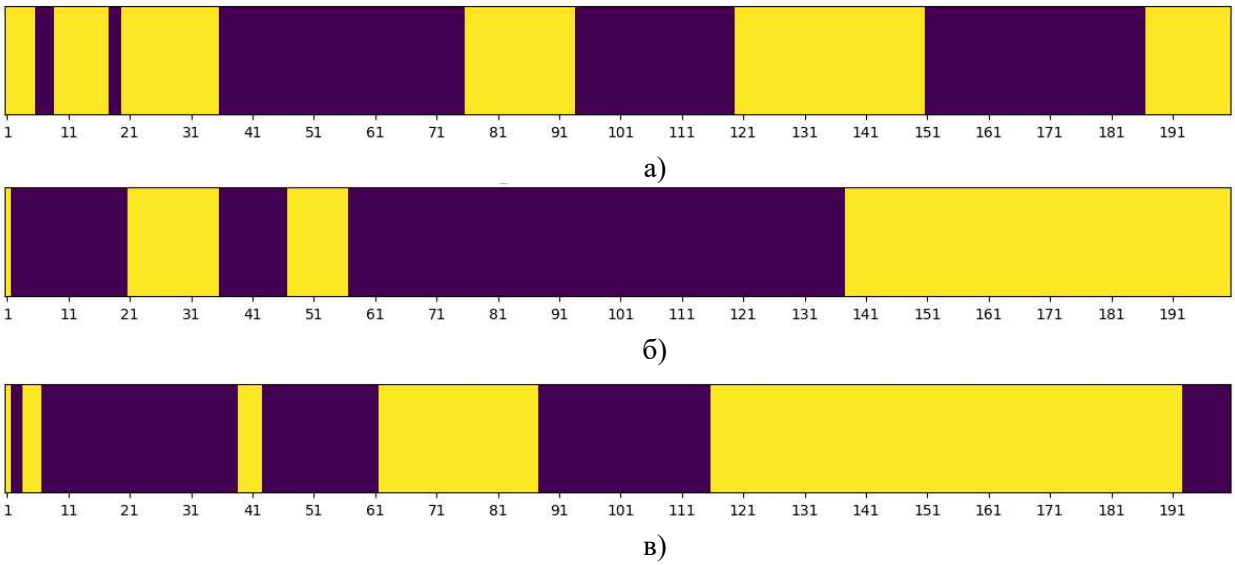


а) б)

в) г)

Рис 2.8. Поведінка PPS для різних значень m для нормального розподілу $N(0,1)$: а) при $m=10$; б) при $m=20$; в) при $m=37$; г) при $m=200$.

Графік зміни знаку p_m (Рис. 2.9) відображає послідовність знаків елементів PPS і дає змогу оцінити наявність осциляцій у поліноміальних прогнозах. Якщо знак прогнозних значень на певному інтервалі зберігається, це може вказувати на локальну узгодженість елементів PPS. Графік є допоміжним інструментом для виявлення осциляцій, інтервалів відносної стабільності та моментів зміни характеру поведінки послідовності поліноміальних прогнозів.



а)

б)

в)

Рис. 2.9. Зміна знаку послідовності PPS для різних типів розподілів (значення $m=200$, для всіх параметр $seed=42$: а) нормальний розподіл; б) рівномірний розподіл; в) показниковий розподіл.

Таким чином, чисельні експерименти показують, що PPS для стохастичних даних розбігається. Це означає, що розбігатиметься і PPS для вихідного процесу,

що складається із детермінованої та стохастичної компоненти. При цьому, за рахунок вибору різних значень параметру ε можемо отримати ефект аналога збіжності для скінченного випадку, тобто матимемо скінченно біноміально стабільні значення.

Поняття скінченно біноміально-стабільних значень вперше введено у роботі [148]. Часовий ряд будемо називати скінченно біноміально стабільним, якщо $\exists k_0 \in \mathbb{N} : \forall k > k_0 \left| p_{k+1} - p_k \right| < \left| p_k - p_{k-1} \right|$. Очевидно, що для будь-якого вхідного скінченного набору точок існує таке значення ε, при якому стохастично збурений часовий ряд буде скінченно біноміально стабільним. Однак надзвичайно висока чутливість послідовності PPS до стохастичних значень дозволяє використовувати її в якості критерію для перевірки стохастичності процесу.

Метод дослідження стохастичності процесу на основі PPS:

1. Отримуємо значення часово процесу з рівномірними інтервалами спостережень.
2. Будуємо PPS на основі співвідношення (2.47).
3. Перевіряємо біноміальну стабільність часового ряду.
4. Якщо часовий ряд є біноміально-стабільним, то вважаємо, що процес є детермінованим для об'єму вибірки n. В протилежному випадку маємо випадковість чи детермінований хаос.

Звернемо увагу на той факт, що в процесі чисельних досліджень будувались прогнозні значеннння на основі многочленів до 200 степеня, що було б чисельно неможливо за умови використання класичних підходів до побудови прогнозів на основі многочленів через обчислювальну складність задач знаходження коефіцієнтів многочленів та розв'язання систем лінійних рівнянь великої розмірності. Лише наявність прямої формули обчислення прогнозного значення (2.46) робить можливим відповідні дослідження.

Висновки до другого розділу

У підрозділі формалізовано задачу прогнозування біржового часового ряду як задачу визначення наступного значення показника x_{t+1} на основі локального

ковзного вікна X_t. Обґрунтовано доцільність однокрокового прогнозування в режимі walk-forward, який забезпечує перевірку алгоритмів без використання майбутньої інформації.

Визначено сукупність метрик оцінювання точності: *MAE, MSE, RMSE, MAPE*, а також додаткові показники *MASE* і *DA*. Сформульовано передумови застосування поліноміального підходу: на короткому часовому фрагменті допускається існування локально регулярної складової процесу, яку можна наближено відтворити методом поліноміальної екстраполяції.

На відміну від використання одного наперед заданого полінома, запропонована методологія передбачає формування послідовності поліноміальних прогнозів $PPS(X_t)$ і застосування адаптивного оператора аналізу та агрегування її елементів. Така постановка створює основу для подальшого розроблення алгоритмів вибору параметрів прогнозування, аналізу структурних властивостей PPS, оцінювання ступеня стохастичності процесів і побудови нейромережевих моделей.

У розділі розглянуто підхід до аналізу стохастичності процесів на основі поведінки послідовності поліноміальних прогнозів. Показано, що PPS може використовуватися не лише як інструмент короткострокового прогнозування, а й як засіб виявлення закономірностей у часовому ряді. Теоретичною основою такого підходу є зв’язок між збіжністю PPS і поведінкою відповідних скінченних різниць, що дозволяє розглядати збіжність або скінченну біноміальну стабільність як ознаку наявності детермінованої структури в даних.

У результаті чисельних експериментів встановлено, що для послідовностей незалежних випадкових величин із рівномірним, показниковим і нормальним розподілами PPS не демонструє збіжності у класичному розумінні. При цьому поведінка PPS не є стохастичною: на графіках спостерігаються інтервали монотонного зростання або спадання, локальні екстремуми, зміни знаку та зростання амплітуди прогнозних значень. Це свідчить про те, що біноміальна структура формули поліноміального прогнозу перетворює випадкові вхідні

флуктуації у специфічну осциляційну послідовність із внутрішніми локальними закономірностями, однак без ознак стійкої збіжності.

Отримані результати дають підстави розглядати розбіжність PPS як індикатор стохастично-домінованої поведінки процесу. Якщо для досліджуваного часового ряду послідовність поліноміальних прогнозів є скінченно біноміально стабільною, це може свідчити про наявність локальної детермінованої структури на відповідному фрагменті даних. Якщо ж PPS розбігається, має значні осциляції, часті зміни знаку та зростаючу амплітуду, то це може вказувати на наявність вираженої випадкової складової або складної динаміки, що не описується локальною поліноміальною структурою.

Запропонований підхід має прикладну цінність для задач інтелектуального аналізу коротких часових рядів, де застосування складних статистичних, фазових або нейромережевих моделей може бути обмеженим через малий обсяг вибірки, високу чутливість до параметрів або значні обчислювальні витрати. Використання прямої біноміальної формули для обчислення поліноміальних прогнозів дає змогу досліджувати PPS для високих порядків без явного знаходження коефіцієнтів інтерполяційних поліномів, що істотно спрощує проведення чисельних експериментів.

Перспективи подальших досліджень полягають у формалізації кількісних показників розбіжності PPS, зокрема через аналіз приростів між сусідніми елементами, частоти зміни знаку, швидкості зростання амплітуди та довжини інтервалів монотонності. Окремим напрямом подальшої роботи є порівняння PPS-критерію зі стандартними методами аналізу складних процесів, зокрема найбільшим показником Ляпунова, рекурентним аналізом і дисперсією Аллана, а також перевірка запропонованого підходу на реальних фінансових, технічних та природничих часових рядах.

РОЗДІЛ 3. ІНТЕЛЕКТУАЛЬНІ МЕТОДИ АДАПТИВНОГО ПРОГНОЗУВАННЯ ЧАСОВИХ РЯДІВ БІРЖОВОЇ ДИНАМІКИ НА ОСНОВІ ПОСЛІДОВНОСТІ ПОЛІНОМІАЛЬНИХ ПРОГНОЗІВ

3.1. Загальна схема адаптивного поліноміального прогнозування

Класичні підходи до прогнозування часових рядів здебільшого передбачають попередній вибір структури моделі, оцінювання її параметрів та подальше використання отриманої моделі для прогнозування [27; 28]. У фінансових часових рядах така процедура ускладнюється тим, що локальна поведінка котирувань може швидко змінюватися, а обсяг інформативного фрагмента даних часто є обмеженим. Саме тому в сучасних підходах до прогнозування поширеними є ідеї адаптивного вибору моделі, локального перенавчання, комбінування прогнозів та використання різних варіантів агрегування прогнозних значень [29; 30]. У працях з комбінування прогнозів показано, що поєднання кількох прогнозів може зменшувати ризик невдалого вибору єдиної моделі, особливо за умов невизначеності структури процесу [6; 7].

Поліноміальна складова такого підходу спирається на загальні положення теорії інтерполяції, екстраполяції та скінченних різниць. У працях К. Брезінського та М. Редіво-Зальї узагальнено розвиток методів екстраполяції, раціональної апроксимації та пов’язаних чисельних процедур [31]. Окремий напрям становлять поліноміальні та індуктивні моделі, пов’язані з ідеями О. Г. Івахненка щодо побудови моделей оптимальної складності на основі обмеженого обсягу експериментальних даних [33]. У його роботі поліноміальні предиктивні моделі розглядаються як засіб опису складних динамічних систем, а МГУА подано як підхід до побудови поліноміального опису системи за малою кількістю спостережень.

У контексті цієї дисертаційної роботи загальна схема адаптивного поліноміального прогнозування розглядається як базова методична конструкція, на основі якої конкретизуються окремі способи вибору інформативних елементів PPS.

Нехай задано дискретний часовий ряд біржової динаміки

$$F_N = \{f_1, f_2, \ldots, f_N\},$$

де f_i – значення досліджуваного біржового показника в момент часу *i*, наприклад ціна закриття *Close*, ціна останньої угоди, значення індексу або іншого фінансового індикатора.

Для побудови короткострокового прогнозу в момент часу *n* використовується локальний фрагмент останніх *M* значень:

$$W_n^{(M)} = \{f_{n-M+1}, f_{n-M+2}, \ldots, f_n\}.$$

Параметр *M* визначає максимальну глибину локального фрагмента, на якому будуються поліноміальні прогнозні значення. У межах одного фрагмента формується не один прогноз, а набір прогнозів, кожний з яких відповідає різній кількості використаних останніх точок, тобто різній степені локального полінома.

Якщо для побудови *m*-го прогнозу використовуються останні *m* значень ряду, то відповідний інтерполяційний поліном має степінь $m-1$. Значення цього полінома в наступному часовому вузлі $n+1$ розглядається як *m*-й поліноміальний прогноз.

Для рівновіддалених часових вузлів поліноміальний прогноз на один крок уперед можна подати через скінченні різниці. Позначимо через $\nabla^r f_n$ різницю *r*-го порядку, обчислену в кінцевій точці локального фрагмента:

$$\nabla^r f_n = \sum_{k=0}^{r} (-1)^k C_r^k f_{n-k}, \quad r = 0,1,\ldots,m-1. \tag{3.1}$$

Тоді прогноз, побудований за поліномом степеня $m-1$, можна записати у вигляді:

$$p_m = \sum_{r=0}^{m-1} \nabla^r f_n. \tag{3.2}$$

Еквівалентна форма цієї формули через біноміальні коефіцієнти має вигляд:

$$p_m = \sum_{k=1}^{m} (-1)^{k-1} C_m^k f_{n-k+1}, \quad m = 1,2,\ldots,M. \tag{3.3}$$

Послідовність поліноміальних прогнозних значень, сформовану в момент часу *n*, позначимо як $PPS_n^{(M)} = (p_1, p_2, \ldots, p_M)$.

Тут $PPS_n^{(M)}$ – не часовий ряд у звичайному розумінні, а впорядкована послідовність альтернативних прогнозів, побудованих для одного й того самого моменту $n+1$. Індекс m у цій послідовності не є часовим індексом, а визначає кількість останніх значень, використаних для побудови прогнозу, або, що еквівалентно, степінь відповідного локального полінома плюс одиниця.

У роботах, присвячених удосконаленню поліноміальних прогнозів, показано, що безпосередній вибір полінома високого степеня не гарантує зменшення помилки екстраполяції, оскільки навіть точна інтерполяція на локальному фрагменті може давати значну похибку за його межами [128]. Тому доцільним є аналіз не одного прогнозу, а всієї послідовності *PPS*, зокрема її монотонності, збіжної тенденції, локальних концентрацій та віддалених значень [127; 128].

Загальна ідея адаптивного прогнозування полягає в тому, що остаточний прогноз не задається наперед фіксованим номером m, а визначається на основі поточної структури послідовності $PPS_n^{(M)}$. Для цього вводиться оператор вибору інформативної підмножини:

$$I_n = D\left(PPS_n^{(M)}, \theta_n\right), \tag{3.4}$$

де $I_n \subseteq \{1,2,\ldots,M\}$ – множина індексів прогнозів, які в поточному локальному фрагменті вважаються інформативними; $D(\cdot)$ – правило вибору; θ_n – набір параметрів адаптації в момент часу n.

Параметр θ_n може включати максимальну глибину M, мінімальну кількість елементів для агрегування, поріг відхилення, параметри пошуку інтервалу концентрації, параметри кластеризації або інші налаштування методу.

Остаточний прогноз формується як агрегована характеристика елементів *PPS*, індекси яких належать до I_n:

$$\hat{f}_{n+1} = A\left(\{p_m \mid m \in I_n\}\right), \tag{3.5}$$

де $A(\cdot)$ – оператор агрегування.

Одним з варіантів з таких операторів є середнє арифметичне вибраних прогнозних значень:

$$\hat{f}_{n+1} = \frac{1}{|I_n|}\sum_{m\in I_n} p_m . \tag{3.6}$$

Більш загальний варіант передбачає використання вагового агрегування:

$$\hat{f}_{n+1} = \sum_{m\in I_n} w_m p_m, \quad w_m \geq 0, \quad \sum_{m\in I_n} w_m = 1. \tag{3.7}$$

Для врахування зміни локальної структури біржового ряду параметри методу можуть уточнюватися на ковзному контрольному фрагменті. Нехай Θ – множина допустимих параметрів, а L – довжина контрольного інтервалу. Тоді локальний критерій якості можна записати так:

$$Q_n(\theta) = \frac{1}{L}\sum_{i=n-L}^{n-1} \ell\left(f_{i+1}, \hat{f}_{i+1}(\theta)\right), \quad \theta \in \Theta. \tag{3.8}$$

Тоді адаптивний вибір параметрів може бути поданий як задача мінімізації деякого локального критерію:

$$\theta_n^* = \arg\min_{\theta\in\Theta} Q_n(\theta). \tag{3.9}$$

У межах адаптивної схеми функція втрат $l(\cdot)$ використовується для локального оцінювання відхилення прогнозного значення $\hat{f}_{i+1}$, сформованого на основі PPS, від фактичного значення f_{i+1}. На відміну від підсумкових метрик якості прогнозування, функція втрат характеризує помилку на одному прогнозному кроці та може застосовуватися для вибору поточних параметрів адаптації. Зокрема, як функцію втрат можна використати абсолютну або квадратичну похибку:

$$\ell_1(f_{i+1}, \hat{f}_{i+1}) = |f_{i+1} - \hat{f}_{i+1}|,$$
$$\ell_2(f_{i+1}, \hat{f}_{i+1}) = (f_{i+1} - \hat{f}_{i+1})^2.$$

Таким чином, адаптація реалізується не лише через зміну параметрів, а й через зміну способу інтерпретації поточної PPS. Це є суттєвою відмінністю від підходів, у яких модель фіксується після одноразового навчання або оцінювання параметрів.

Загальна схема адаптивного поліноміального прогнозування на основі PPS може бути подана у вигляді такого алгоритму.

Алгоритм 3.1. Загальна схема адаптивного поліноміального прогнозування.

Вхід: часовий ряд $F_N = \{f_1, f_2, \ldots, f_N\}$; максимальна глибина M; множина допустимих параметрів Θ; правило вибору D; оператор агрегування A.

Вихід: прогнозне значення $\hat{f}_{n+1}$.

1. Вибрати поточний локальний фрагмент

$$W_n^{(M)} = \{f_{n-M+1}, f_{n-M+2}, \ldots, f_n\}.$$

2. Для кожного $m = 1, 2, \ldots, M$ побудувати поліноміальний прогноз

$$p_m = \sum_{k=1}^{m} (-1)^{k-1} C_m^k f_{n-k+1}.$$

3. Сформувати послідовність поліноміальних прогнозів

$$PPS_n^{(M)} = (p_1, p_2, \ldots, p_M).$$

4. Визначити поточні параметри адаптації

$$\theta_n^* = \arg\min_{\theta \in \Theta} Q_n(\theta)$$

або задати їх відповідно до обраної процедури аналізу PPS.

5. Вибрати інформативну підмножину індексів

$$I_n = D\left(PPS_n^{(M)}, \theta_n\right).$$

6. Сформувати остаточний прогноз

$$\hat{f}_{n+1} = A\left(\{p_m \mid m \in I_n\}\right).$$

7. Після надходження фактичного значення f_{n+1} обчислити похибку прогнозування та перейти до наступного моменту часу.

У вигляді блок-схеми алгоритм наведений на рис. 3.1.

Адаптивність у запропонованій схемі проявляється на трьох рівнях.

Перший рівень – локалізація даних. Для кожного моменту прогнозування використовується не вся історія ряду, а поточний локальний фрагмент $W_n^{(M)}$. Це дає змогу враховувати короткочасні зміни біржової динаміки без побудови громіздкої глобальної моделі.

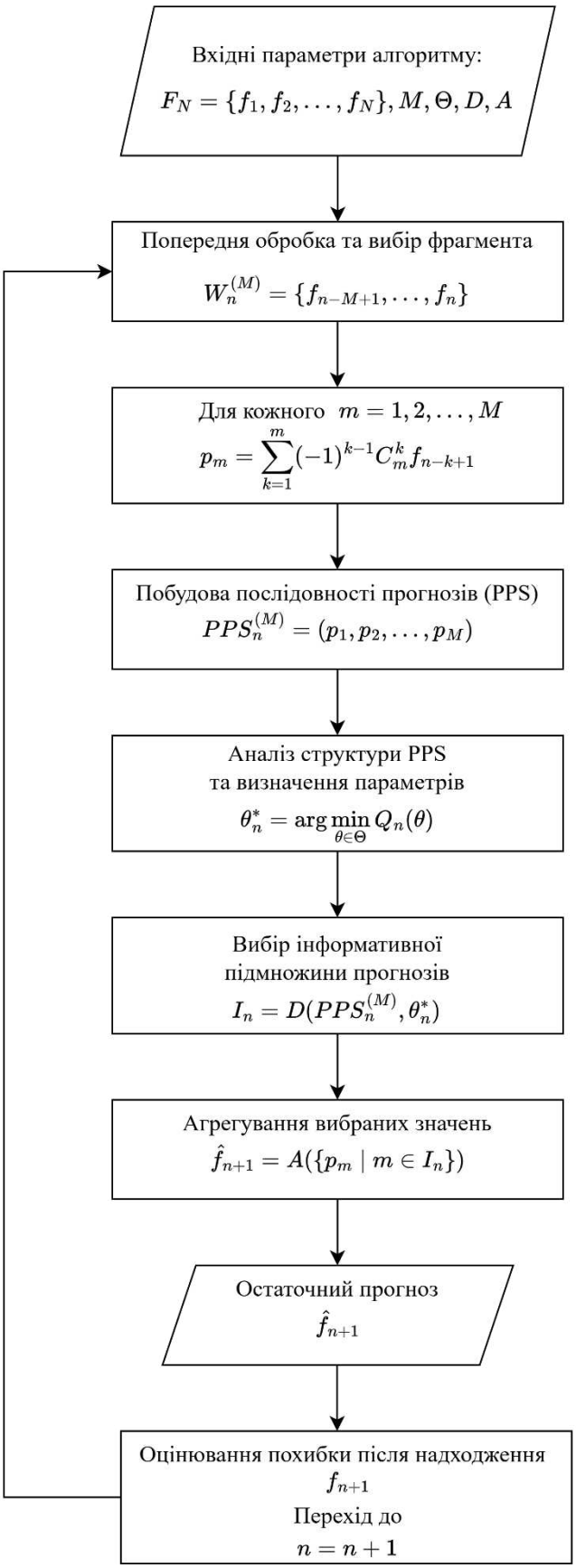


Рис. 3.1. Блок-схема алгоритму адаптивного поліноміального прогнозування на основі PPS

Другий рівень – варіативність поліноміальних прогнозів. Замість фіксації одного значення степеня полінома формується послідовність $PPS_n^{(M)}$, яка містить кілька альтернативних прогнозних значень для одного й того самого моменту часу. Тому задача прогнозування зводиться не лише до екстраполяції, а й до аналізу структури отриманої прогнозної послідовності.

Третій рівень – адаптивний вибір або агрегування елементів PPS. Остаточний прогноз визначається на основі правила *D*, яке може враховувати форму PPS, локальні відхилення, концентрацію прогнозних значень або кластерну структуру. Саме цей рівень є основою для подальших методів, що наведені в наступних підрозділах.

У роботах українських науковців також простежується близька ідея адаптивного вибору параметрів або моделей для прогнозування фінансових і економічних процесів. Зокрема, у працях В. Дербенцева, А. Матвійчука, Н. Даценко, В. Безкоровайного та А. Азаряна досліджено застосування методів машинного навчання до короткострокового прогнозування фінансових часових рядів [34]. У роботах Ю. Зайченка та Г. Зайченко розглянуто нечіткий метод групового урахування аргументів для задач фінансового прогнозування [35]. Це підтверджує актуальність поєднання інтелектуального аналізу даних, адаптивного вибору параметрів і поліноміальних моделей у задачах біржової динаміки.

3.2. Алгоритм прогнозування з адаптивним вибором порядку поліноміальної екстраполяції

У класичних підходах до екстраполяції вибір степеня полінома або кількості опорних точок істотно впливає на похибку прогнозу. Відомо, що екстраполяція за межі інтервалу спостережень є значно чутливішою до шуму та локальних збурень, ніж інтерполяція [31; 130]. Для біржових рядів ця проблема ускладнюється наявністю волатильності, локальних розривів динаміки та короткочасних змін режиму [131; 132]. Тому доцільно не фіксувати наперед один степінь полінома, а формувати множину локальних поліноміальних прогнозів і адаптивно визначати, яку частину цієї множини використовувати.

У цьому алгоритмі параметр *m* розглядається як кількість поліноміальних прогнозів, що включаються до усереднення. Водночас максимальний степінь полінома, який бере участь у такому усередненні, дорівнює $m-1$. Отже, вибір *m* фактично задає глибину локальної поліноміальної екстраполяції.

Нехай задано біржовий часовий ряд $F_N = \{f_1, f_2, \ldots, f_N\}$

де f_i – значення досліджуваного показника у момент часу *i*, наприклад ціна закриття *Close*, *n* – номер останнього відомого спостереження.

Для поточного моменту *n* розглянемо локальне вікно передісторії:

$$W_n^{(M)} = \{f_{n-M+1}, f_{n-M+2}, \ldots, f_n\}, \tag{3.10}$$

Для кожного $m = 1,2,\ldots,M$ будується поліноміальний прогноз p_m, який відповідає екстраполяції на один крок уперед за інтерполяційним поліномом степеня $m-1$, побудованим за останніми *m* точками ряду:

$$p_m = f_{n+1}^{(m)} = P_{m-1}((n+1)\Delta) = \sum_{k=1}^{m} (-1)^{k-1} C_m^k f_{n-k+1}, \quad m = 1,2,\ldots,M\ . \tag{3.11}$$

Отримана послідовність $PPS_n^{(M)} = (p_1, p_2, \ldots, p_M)$ є множиною альтернативних локальних прогнозів, побудованих із різною глибиною використання значень часового ряду.

Остаточний прогноз у методі усередненої поліноміальної екстраполяції формується як середнє перших *m* елементів PPS:

$$\bar{p}_m = \frac{1}{m} \sum_{i=1}^{m} p_i\ . \tag{3.12}$$

Основна задача полягає у виборі *m*. Однак після введення усереднення виникає нова задача: необхідно визначити, скільки саме елементів PPS потрібно включити до середнього. Якщо вибрати занадто мале *m*, прогноз може недостатньо враховувати локальну структуру ряду. Якщо ж вибрати занадто велике *m*, до усереднення можуть потрапити прогнози, побудовані за високими степенями поліномів, що підвищує чутливість до випадкових коливань.

Природним способом вибору параметра *m* є мінімізація похибки прогнозу:

$$e_m = f_{n+1} - \bar{p}_m\ . \tag{3.13}$$

Безпосередньо мінімізувати фактичну похибку (3.13) неможливо, оскільки в момент прогнозування значення f_{n+1} ще невідоме. Саме ця обставина зумовлює здійснити перехід від фактичної похибки до її внутрішньої оцінки, яка може бути обчислена лише за вже відомими значеннями ряду.

Отже, ключова ідея полягає в тому, щоб замінити невідоме майбутнє значення f_{n+1} його допоміжною оцінкою. Для цього використовується інформація про локальні прирости ряду:

$$f_n - f_{n-1}, \quad f_{n-1} - f_{n-2}, \quad \ldots, \quad f_{n-m+2} - f_{n-m+1}.$$

На їх основі вводиться додаткова оцінка майбутнього значення:

$$\tilde{p}_m = f_n + \frac{1}{m-1}\sum_{k=1}^{m-1}(-1)^{k-1}C_m^{k+1}\left(f_{n-k+1} - f_{n-k}\right), \quad m \geq 2. \tag{3.14}$$

Ця оцінка інтерпретується як додатковий прогноз f_{n+1}, побудований не безпосередньо за значеннями ряду, а за локальною динамікою його приростів. Такий підхід узгоджується з ідеєю використання різницевих перетворень у часових рядах, однак, на відміну від ARIMA-підходу, тут одночасно враховуються аналоги різниць різних порядків без окремої процедури ідентифікації порядку диференціювання [128; 133].

Далі на основі $\tilde{p}_m$ формується оцінка прогнозу на наступний крок:

$$q_m = \hat{f}_{n+2}^{(m)} = m\tilde{p}_m + \sum_{k=2}^{m}(-1)^{k-1}C_m^k f_{n-k+2}. \tag{3.15}$$

Величина q_m показує, яким був би наступний прогноз, якби замість невідомого f_{n+1} у локальній схемі використовувалась його оцінка $\tilde{p}_m$. Тому розбіжність між q_m і $\tilde{p}_m$ може бути використана як внутрішній критерій узгодженості поліноміального продовження.

Для кожного $m = 2, 3, \ldots, M$ вводиться нормована величина

$$\Delta_m = \frac{1}{m}\left|q_m - p_m\right|. \tag{3.16}$$

Параметр *m* вибирається за правилом мінімізації:

$$m^* = \arg\min_{2\le m\le M} \Delta_m . \tag{3.17}$$

Вибраний параметр m^* задає кількість елементів PPS, що використовуються для формування остаточного прогнозу:

$$\hat{f}_{n+1} = \overline{p}_{m^*} = \frac{1}{m^*}\sum_{i=1}^{m^*} p_i . \tag{3.18}$$

Максимальний степінь полінома, який фактично враховується у прогнозній процедурі, дорівнює $d^* = m^* - 1$. Отже, алгоритм не потребує попереднього фіксування степеня полінома. Він обирає його адаптивно для кожного локального вікна на основі внутрішньої оцінки розбіжності між двома способами продовження ряду.

Алгоритм 3.2. Адаптивне прогнозування на основі усередненої послідовності поліноміальних екстраполяцій

Вхід: локальне вікно $W_n^{(M)} = \{f_{n-M+1}, f_{n-M+2}, \ldots, f_n\}$; максимальна глибина M.

Вихід: прогноз $\hat{f}_{n+1}$; адаптивно вибраний параметр m^*; послідовність значень Δ_m.

1. Для кожного $m = 1, 2, \ldots, M$ обчислити поліноміальний прогноз:

$$p_m = \sum_{k=1}^{m} (-1)^{k-1} C_m^k f_{n-k+1} .$$

2. Для кожного $m = 2, 3, \ldots, M$ обчислити додаткову оцінку:

$$\tilde{p}_m = f_n + \frac{1}{m-1}\sum_{k=1}^{m-1} (-1)^{k-1} C_m^{k+1} \left(f_{n-k+1} - f_{n-k}\right).$$

3. Для кожного $m = 2, 3, \ldots, M$ обчислити прогноз на наступний крок:

$$q_m = \hat{f}_{n+2}^{(m)} = m\tilde{p}_m + \sum_{k=2}^{m} (-1)^{k-1} C_m^k f_{n-k+2} .$$

4. Обчислити внутрішній критерій розбіжності:

$$\Delta_m = \frac{1}{m}\left|q_m - p_m\right|.$$

5. Визначити оптимальне значення:

$$m^* = \arg\min_{2\le m\le M} \Delta_m .$$

6. Сформувати остаточний прогноз:

$$\hat{f}_{n+1} = \overline{p}_{m^*} = \frac{1}{m^*}\sum_{i=1}^{m^*} p_i \,.$$

У разі однакових або близьких значень Δ_m доцільно обирати менше значення m, оскільки воно відповідає меншому максимальному степеню полінома і зменшує ризик надмірного реагування на локальні флуктуації.

Критерій Δ_m не є фактичною похибкою прогнозу, оскільки фактичне значення f_{n+1} у момент прогнозування невідоме. Його слід розглядати як внутрішню оцінку узгодженості локального поліноміального продовження. Якщо значення Δ_m є малим, це означає, що прогноз, побудований безпосередньо за значеннями ряду, та прогноз, побудований із використанням оцінки на основі приростів, дають близькі результати.

У дослідженні варто зазначити, що застосування алгоритму для рядів із вираженою стохастичною компонентою є доцільним за умови відносно малих значень оцінки похибки. Малі значення такої оцінки можуть свідчити про наявність локальної детермінованої складової, яка робить короткострокове прогнозування обґрунтованим.

Для практичного використання можна додатково ввести відносний показник:

$$R_{m^*} = \frac{\Delta_{m^*}}{|\hat{f}_{n+1}| + \eta} \cdot 100\%\,, \tag{3.19}$$

де $\eta > 0$ – мале число, що запобігає діленню на нуль.

Цей показник може використовуватись як допоміжний індикатор надійності локального прогнозу. Якщо R_{m^*} суттєво зростає, то відповідне локальне вікно може містити різкий перехід, аномальне значення або зміну короткочасного режиму динаміки.

3.3. Метод прогнозування на основі кластеризації поліноміальних прогнозних значень

На відміну від підходів, у яких підсумковий прогноз визначається одним поліномом фіксованого порядку або усередненням усіх побудованих прогнозів, у цій роботі використовується ідея виділення найбільш щільної області (кластера) у множині альтернативних поліноміальних прогнозних значень (PPS). Загальна схема методу прогнозування на основі кластеризації значень послідовності поліноміальних прогнозів показана на рис. 3.2.

Нехай задано часовий ряд $f_i = f(i\Delta), i = n, n-1, \ldots, n-N+1,$ де Δ – крок дискретизації, N – кількість останніх доступних спостережень, які використовуються для прогнозування. Потрібно побудувати однокроковий прогноз значення $f_{n+1} = f((n+1)\Delta)$. Для цього для кожного $m = 1, 2, \ldots, k$ будується поліноміальний прогноз p_m, який відповідає використанню останніх m значень ряду. Значення p_m визначається формулою

$$p_m = f_{n+1}^m = \sum_{k=1}^{m} (-1)^{k-1} \binom{m}{k} f_{n-k+1}., \tag{3.20}$$

де $\binom{m}{k} = C_m^k$ – біноміальні коефіцієнти.

Ця формула відповідає швидкому способу обчислення прогнозу на основі інтерполяційного многочлена степеня $m-1$ і використовується як у пропонованому підході, так і в методі усереднення поліноміальної екстраполяційної послідовності [129].

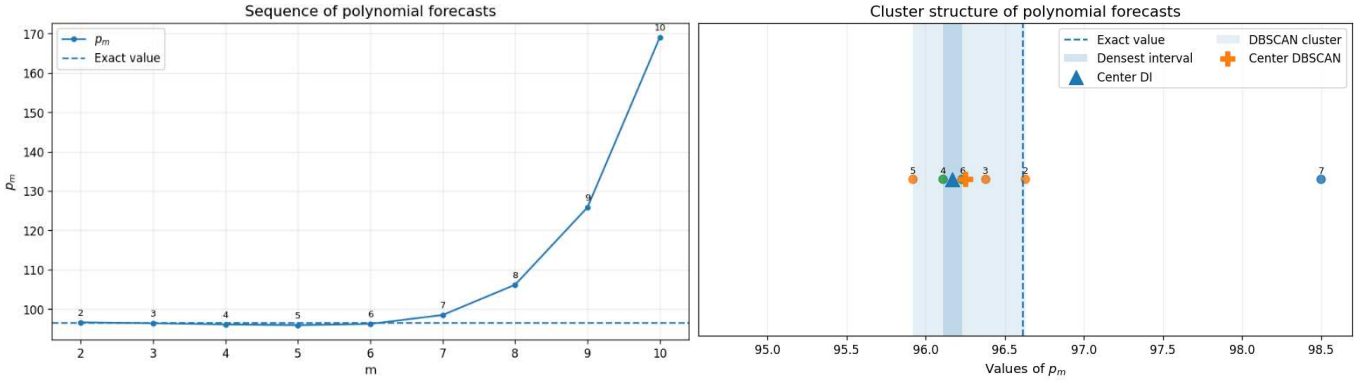


Рис. 3.2. Значення послідовності поліноміальних прогнозів та її кластерна структура.

Отже, для одного й того самого моменту прогнозування формується послідовність PPS: $P=\{p_1,p_2,\ldots,p_m\}$. Особливістю такого підходу є те, що кожне значення p_m є альтернативною оцінкою майбутнього значення f_{n+1}, отриманою на основі різної глибини передісторії. Саме тому інформативним є не лише окреме значення p_m, а й структура всієї послідовності P: наявність згущень, розкиду, локальних груп і віддалених значень. На Рис. 3. 1 показаний приклад такої PPS для 9 точок, обчислених на основі значень біржових котирувань NFLX та їх розміщення на числовій осі:

На відміну від підходів, у яких фінальний прогноз ототожнюється з одним поліномом фіксованого порядку або з усередненням частини послідовності P, у цій роботі використовується кластерний підхід, за якого остаточне значення визначається на основі області найбільшої концентрації елементів множини P. Для аналізу внутрішньої структури послідовності p_m розглянемо два сусідні її елементи:

$$p_m=\sum_{k=1}^{m}(-1)^{k-1}\binom{m}{k}f_{n-k+1},$$

$$p_{m+1}=\sum_{k=1}^{m+1}(-1)^{k-1}\binom{m+1}{k}f_{n-k+1}.$$

Тоді

$$p_{m+1}-p_m=\sum_{k=1}^{m}(-1)^{k-1}\left(\binom{m+1}{k}-\binom{m}{k}\right)f_{n-k+1}+(-1)^m f_{n-m}.$$

З використанням тотожності Паскаля

$$\binom{m+1}{k}-\binom{m}{k}=\binom{m}{k-1}$$

одержуємо

$$p_{m+1}-p_m=\sum_{k=1}^{m}(-1)^{k-1}\binom{m}{k-1}f_{n-k+1}+(-1)^m f_{n-m}.,\qquad(3.21)$$

Використаємо заміну $j=k-1$:

$$p_{m+1} - p_m = \sum_{j=0}^{m} (-1)^j \binom{m}{j} f_{n-j}.., \tag{3.22}$$

Для $m = 1$:

$$p_2 - p_1 = f_n - f_{n-1} = \nabla f_n,$$

для $m = 2$:

$$p_3 - p_2 = f_n - 2f_{n-1} + f_{n-2} = \nabla^2 f_n,$$

для $m = 3$:

$$p_4 - p_3 = f_n - 3f_{n-1} + 3f_{n-2} - f_{n-3} = \nabla^3 f_n,$$

і так далі. Таким чином вираз (3) є m-тою зворотною скінченною різницею в точці f_n, тобто

$$p_{m+1} - p_m = \nabla^m f_n.$$

Отримали рекурентну формулу для послідовності поліноміальних прогнозів:

$$p_{m+1} = p_m + \nabla^m f_n.., \tag{3.23}$$

Таким чином, послідовність поліноміальних прогнозів має внутрішню дискретну структуру. Співвідношення (3.23) показує, що кожен наступний член послідовності поліноміальних прогнозів відрізняється від попереднього на відповідну зворотну скінченну різницю. Тому сама послідовність P є інформативним об'єктом дослідження, а не лише набором ізольованих прогнозних значень.

Для порівняльного аналізу використовується метод, у якому остаточне прогнозне значення будується як середнє арифметичне певної кількості перших елементів послідовності поліноміальних прогнозів . У відповідній роботі запропоновано шукати прогноз у вигляді усереднення підпослідовності $\{p_1, p_2, \ldots, p_m\}$, а оптимальне значення m визначати шляхом мінімізації внутрішнього відхилення між прогнозом у точці $n+1$ та оцінкою, побудованою для точки $n+2$ [129].

Згідно з цією постановкою, усереднений прогноз має вигляд

$$\hat{f}_{n+1}^{(avg,m)} = \frac{1}{m}\sum_{i=1}^{m} f_{n+1}^{(i)}.$$

У наведеній роботі також подано еквівалентний аналітичний запис цього усереднення через значення часового ряду, а для оцінки якості вибору m використано внутрішню характеристику

$$\Delta_m = \frac{1}{m}\left|\hat{f}_{n+2}^{(m)} - f_{n+1}^{(m)}\right|.$$

Після цього оптимальне значення m визначається як

$$m^* = \arg\min_{m\in\{1,2,\ldots,n\}} \Delta_m.$$

Тоді підсумковий прогноз методу усереднення задається як

$$\hat{f}_{n+1}^{(avg)} = \hat{f}_{n+1}^{(avg,m^*)}.$$

Саме цей підхід далі використовується як базовий метод порівняння для оцінки ефективності кластерного формування прогнозу. Його ключова ідея полягає у виборі початкового фрагмента послідовності поліноміальних прогнозів. У запропонованому в дослідженні підході вибір виконується не за індексною позицією елементів, а за їх просторовим розташуванням на числовій осі. Виділяються області найбільшої концентрації значень p_m та будується підсумковий прогноз на основі центральної характеристики цієї області.

Нехай C підмножина значень P, яка відповідає знайденому кластеру або найщільнішому інтервалу. Тоді фінальний прогноз визначається як

$$f_{n+1} = \Phi(C),$$

де $\Phi(C)$ – правило обчислення центру кластера. У програмній реалізації передбачено три варіанти:

середнє арифметичне:

$$\Phi(C) = \frac{1}{|C|}\sum_{x\in C} x;$$

медіана:

$$\Phi(C) = \text{median}(C);$$

середина інтервалу:

$$\Phi(C) = \frac{L_C + R_C}{2},$$

де L_C та R_C – ліва і права межі знайденого кластера або інтервалу.

В дослідженні реалізовані два кластерні підходи – метод найщільнішого інтервалу (DI – Densest Interval) та DBSCAN (Density-Based Spatial Clustering of Applications with Noise).

Перший кластерний підхід (DI) ґрунтується на пошуку найщільнішого інтервалу серед значень послідовності P. Для цього елементи множини впорядковуються за зростанням $v_1 \le v_2 \le \ldots \le v_s$.

Тоді розглядаються всі можливі неперервні підпослідовності $[v_l, v_s]$, $1 \le l < r \le s$, які містять не менше, ніж q точок. Для кожного такого інтервалу обчислюється його ширина $w_{l,r} = v_r - v_l$ та класичний показник щільності

$$\rho_{l,r} = \frac{r - l + 1}{\max(v_r - v_l, \varepsilon_0)},$$

де $\varepsilon_0 > 0$ – мале число для уникнення ділення на нуль.

Оптимальним вважається інтервал $(l^*, r^*) = \arg\max_{l,r} \rho_{l,r}$.

Після цього формується кластер $C_{DI} = \{v_{l^*}, v_{l^*+1}, \ldots, v_{r^*}\}$, а прогнозне значення визначається як $\hat{f}_{n+1}^{(DI)} = \Phi(C_{DI})$.

Другий кластерний підхід базується на алгоритмі DBSCAN. Множина P розглядається як одновимірна вибірка $P = \{p_1, p_2, \ldots, p_m\}$.

Алгоритм використовує два параметри: ε – радіус локального сусідства; $MinPts$ – мінімальна кількість точок у ε-околі.

Точки, що мають достатню кількість сусідів у межах радіуса ε, утворюють щільні області, які трактуються як кластери. Точки, що не належать жодній такій області, розглядаються як шум.

У результаті формується множина кластерів $C_{DB} = \{C_1, C_2, \ldots, C_K\}$.

Як підсумковий обирається кластер з найбільшою кількістю елементів:

$$C_{DB}^{*} = \arg\max_{C_j \in C_{DB}} \left| C_j \right|.$$

Тоді фінальний прогноз визначається як

$$\hat{f}_{n+1}^{(DB)} = \Phi(C_{DB}^{*}).$$

Для дослідження властивостей запропонованого підходу використовуються три типи даних. У контрольованих експериментах аналізуються значення функцій

$$y = \cos(hx),$$
$$y = \cos(h \ln x),$$
$$y = e^{hx},$$

де h – параметр функції.

Для вивчення поведінки алгоритмів на шумових послідовностях використовуються випадкові дані, згенеровані з нормального розподілу $y_i \sim \mathcal{N}(0, \sigma^2)$, де σ визначає рівень шуму.

Для оцінки практичної придатності методу використовуються реальні фінансові часові ряди, які завантажуються з Excel-файлів. В якості експериментальних даних розглянуті інтраденні котирування акцій Netflix (NFLX) протягом одного операційного дня (08.12.2025) з кроком дискретизації $\Delta = 30$ хвилин (поля: Timestamp, Close). Також ефективність методу була перевірена на щоденних котируваннях акцій Apple (AAPL) протягом 2019 -2025 років. Із вибраного стовпця (Close) формується послідовність значень, де останнє спостереження використовується як контрольне, а попередні – як історія для побудови прогнозу.

Для кожного з підходів (DI, DBSCAN) обчислюється абсолютна похибка

$$\Delta_{abs} = \left| f_{n+1} - \hat{f}_{n+1} \right|$$

та відносна похибка

$$\Delta_{rel} = \frac{\left| f_{n+1} - \hat{f}_{n+1} \right|}{\left| f_{n+1} \right|} \cdot 100\%.$$

Крім числових характеристик, у програмній реалізації застосовуються засоби візуального аналізу, зокрема графік вихідного часового ряду з точним і прогнозними значеннями, графік значень p_m з нумерацією точок, графік кластерного аналізу на числовій осі з виділенням меж кластерів та гістограма розподілу значень p_m. Окремо зафіксовано, які саме значення p_m увійшли до знайдених кластерів. Це дозволяє аналізувати не лише точність підсумкового прогнозу, а й внутрішню структуру множини альтернативних поліноміальних прогнозів.

3.4. Метод прогнозування на основі аналізу екстремумів PPS

У попередніх підрозділах послідовність поліноміальних прогнозних значень PPS розглядалася як результат формування набору локальних прогнозів, що відповідають різним степеням поліноміальної екстраполяції. Далі пропонується інший спосіб використання цієї послідовності: не лише як множини альтернативних прогнозів, а як впорядкованого ряду значень, у якому зміна прогнозу при переході від одного степеня полінома до іншого містить додаткову інформацію про локальну зміну біржового часового ряду.

Ідея аналізу екстремумів узгоджується з підходами, у яких локальні максимуми та мінімуми використовуються для виявлення точок зміни динаміки фінансових часових рядів. Зокрема, у працях R. El-Yaniv та A. Faynburd локальні екстремальні точки розглядаються як turning points, що позначають початок нових короткочасних коливань фінансової послідовності [134]. У роботі F. Yazdani, M. Khashei та S. R. Hejazi задача виявлення turning points подана як окремий етап побудови ефективної торгової або прогнозної стратегії [135].

Методологічно запропонований підхід також пов’язаний із локальним поліноміальним моделюванням, що використовується для опису коротких фрагментів складних часових залежностей [119]. У контексті української наукової школи важливими є праці О. Г. Івахненка та В. Г. Лапи, у яких розвинуто ідеї кібернетичного прогнозування та поліноміального опису складних систем [32 ;33]. Зокрема, у поліноміальній теорії складних систем Івахненка прогнозні поліноми розглядаються як засіб дослідження динамічних об’єктів, а GMDH-

підхід орієнтований на побудову моделей за обмеженим обсягом експериментальних даних [33].

Сучасні українські дослідження також розглядають інтелектуальні методи прогнозування у фінансовій сфері. Зокрема, у роботах Ю. Зайченка, Г. Заїченко та О. Кузьменка досліджено LSTM-мережі та гібридні нейромережі на основі GMDH для коротко- та середньострокового прогнозування фінансових показників [136]. Це підтверджує актуальність поєднання локальних поліноміальних моделей із адаптивними процедурами вибору прогнозного значення.

У запропонованому методі локальний екстремум послідовності поліноміальних прогнозних значень (PPS) не використовується як остаточне прогнозне значення. Його доцільно розглядати як певний орієнтир, який дозволяє виділити інформативну частину PPS. Такий підхід зменшує залежність результату від окремого поліноміального прогнозу та дає змогу використовувати групу близьких прогнозних значень, що відображають локальну динаміку фінансового часового ряду.

Нехай для моменту часу n сформовано послідовність поліноміальних прогнозних значень $PPS_n^{(M)} = \left(p_{n,1}, p_{n,2}, \ldots, p_{n,M}\right)$, де $p_{n,m}$ – прогноз значення f_{n+1}, отриманий на основі поліноміальної екстраполяції з використанням попередніх значень та полінома відповідного степеня.

На відміну від методу фіксованого вибору параметра m, у цьому підході аналізується характер зміни значень $p_{n,m}$ при збільшенні m. Якщо послідовність PPS має локальний екстремум, то це може свідчити про момент, у якому додавання наступних різницевих складових починає істотно змінювати прогнозне значення. Тому локальний екстремум PPS розглядається як індикатор інформативного значення параметра m.

Для аналізу зміни PPS введемо різницю між сусідніми поліноміальними прогнозами:

$$g_{n,m} = p_{n,m+1} - p_{n,m}, \quad m = 1,2,\ldots,M-1. \tag{3.24}$$

Значення $g_{n,m}$ характеризує зміну прогнозу при переході від *m*-го до $m+1$ -го елемента PPS. Якщо знак цієї різниці змінюється, то у відповідній точці послідовності виникає локальний максимум або локальний мінімум.

Локальний екстремум у точці *m* визначається умовою

$$\left(p_{n,m} - p_{n,m-1}\right)\left(p_{n,m+1} - p_{n,m}\right) < 0, \quad m = 2,3,\ldots,M-1. \tag{3.25}$$

Якщо виконується умова

$$p_{n,m-1} < p_{n,m} > p_{n,m+1}, \tag{3.26}$$

то $p_{n,m}$ є локальним максимумом PPS.

Якщо виконується умова

$$p_{n,m-1} > p_{n,m} < p_{n,m+1}, \tag{3.27}$$

то $p_{n,m}$ є локальним мінімумом PPS.

Таким чином, задача прогнозування зводиться до пошуку такого значення *m*, для якого відповідний елемент PPS є локальним екстремумом і може бути використаний як адаптивно вибране прогнозне значення.

У реальних біржових даних невеликі коливання PPS можуть бути спричинені не зміною локальної структури ряду, а шумовими компонентами або чисельними ефектами високих різниць. Тому доцільно використовувати не лише факт зміни знака, а й величину локального відхилення.

Для цього введемо показник виразності екстремуму:

$$A_{n,m} = \min\left(\left|p_{n,m} - p_{n,m-1}\right|, \left|p_{n,m} - p_{n,m+1}\right|\right). \tag{3.28}$$

Показник $A_{n,m}$ характеризує мінімальну амплітуду відхилення екстремального елемента від двох сусідніх прогнозів. Для відсікання малозначущих коливань введемо порогове значення (показник локальної мінливості)

$$\varepsilon_n = \alpha \cdot \text{median}_{1 \le j \le M-1}\left|p_{n,j+1} - p_{n,j}\right|. \tag{3.29}$$

де α – емпіричний коефіцієнт чутливості, який може задаватися в межах $0 < \alpha < 1$.

Тоді множину значущих екстремумів PPS можна записати у вигляді

$$E_n = \left\{ m : 2 \le m \le M-1, \left(p_{n,m} - p_{n,m-1}\right)\left(p_{n,m+1} - p_{n,m}\right) < 0, A_{n,m} \ge \varepsilon_n \right\}. \quad (3.30)$$

Якщо множина E_n не є порожньою, то адаптивний параметр m_n^* визначається як перший значущий екстремум:

$$m_n^* = \min E_n . \quad (3.31)$$

Вибір першого значущого екстремуму пояснюється тим, що початкові елементи PPS відповідають нижчим степеням поліноміальної екстраполяції, тоді як подальше збільшення *m* може призводити до надмірного врахування високих різниць. Отже, перший виражений екстремум розглядається як компроміс між недостатнім урахуванням локальної динаміки та надмірною чутливістю до випадкових коливань.

Допустимий діапазон після екстремуму визначимо як $D_n^{ext} = \left[p_{n,m^*} - \delta_n, p_{n,m^*} + \delta_n \right]$. Параметр допустимого відхилення задається через локальну мінливість PPS: $\delta_n = \beta V_n^{PPS}$, де $\beta > 0$ – параметр ширини допустимого діапазону. Після екстремуму до інформативної підмножини включаються лише ті значення ($p_{n,j}$), для яких виконується умова

$$\left| p_{n,j} - p_{n,m^*} \right| \le \delta_n, \quad j = m^*+1, m^*+2, \ldots, M . \quad (3.32)$$

Перегляд значень, які знаходяться після екстремуму, здійснюється послідовно. Якщо для деякого індексу *j* умова (3.32) перестає виконуватись, усі наступні елементи PPS до інформативної підмножини не включаються.

Позначимо через k_n останній індекс PPS, який входить до інформативної підмножини. Тоді,

$$k_n = \max \left\{ k : m^* \le k \le M, \left| p_{n,j} - p_{n,m^*} \right| \le \delta_n, j = m^*+1, \ldots, k \right\}. \quad (3.33)$$

Отже, інформативна підмножина PPS має вигляд

$$PPS_n^{inf} = \left(p_{n,1}, p_{n,2}, \ldots, p_{n,k_n} \right).$$

Остаточне прогнозне значення визначається як середнє арифметичне елементів інформативної підмножини:

$$\hat{f}_{n+1}^{ext} = \frac{1}{k_n} \sum_{j=1}^{k_n} p_{n,j} \,. \tag{3.34}$$

Якщо $p_{n,1} = f_n$ розглядається не як самостійний поліноміальний прогноз, а як початковий елемент PPS, то усереднення доцільно виконувати з другого елемента послідовності.

Якщо значущих екстремумів не виявлено, тобто $E_n = \varnothing$, то PPS розглядається як монотонна або майже монотонна послідовність, а прогноз формується за базовим правилом:

$$\hat{f}_{n+1}^{ext} = p_{n,2} \,. \tag{3.34}$$

У (3.34) $p_{n,2}$ відповідає лінійній екстраполяції, яка використовується як резервний прогноз у випадку, коли PPS не містить достатньо вираженої внутрішньої зміни.

Алгоритм 3.3. Метод прогнозування на основі аналізу екстремумів PPS

Вхід**:** локальне вікно даних часового ряду $f_{n-M+1}, \ldots, f_n$; максимальна глибина *M*; коефіцієнт чутливості до локальних коливань PPS α, параметр ширини допустимого діапазону після екстремуму β.

Вихід: прогнозне значення $\hat{f}_{n+1}^{ext}$; адаптивно вибраний параметр m_n^*, останній індекс інформативної підмножини PPS (кількість елементів для усереднення) k_n.

1. Сформувати послідовність поліноміальних прогнозів:

$$PPS_n^{(M)} = \left(p_{n,1}, p_{n,2}, \ldots, p_{n,M} \right).$$

2. Для кожного $m = 1,2,\ldots,M-1$ обчислити різниці між сусідніми елементами PPS:

$$g_{n,m} = p_{n,m+1} - p_{n,m} \,.$$

3. Обчислити поріг значущості ε_n:

$$\varepsilon_n = \alpha \cdot \text{median}_{1 \le j \le M-1} \left| p_{n,j+1} - p_{n,j} \right|$$

4. Для кожного $m = 2,3,\ldots,M-1$ перевірити умову локального екстремуму:

$$g_{n,m-1}g_{n,m} < 0$$

5. Для кожного знайденого екстремуму обчислити показник виразності $A_{n,m}$.
6. Сформувати множину значущих екстремумів E_n.
7. Якщо $E_n \neq \varnothing$, вибрати $m_n^* = \min E_n$ і сформувати прогноз

$$\hat{f}_{n+1}^{ext} = p_{n,m_n^*}$$

8. Якщо $E_n = \varnothing$, використати резервне правило:

$$\hat{f}_{n+1}^{ext} = p_{n,2}$$

Висновки до третього розділу

У підрозділі сформовано загальну схему адаптивного поліноміального прогнозування часових рядів біржової динаміки на основі послідовності поліноміальних прогнозів. На відміну від підходів, у яких прогноз будується за одним наперед заданим поліномом, запропонована схема передбачає формування множини локальних прогнозних значень PPS, подальший аналіз її структури та вибір інформативної частини для побудови остаточного прогнозу.

Запропонована формалізація включає: вибір локального фрагмента часового ряду, побудову поліноміальних прогнозів різного степеня, формування PPS, адаптивний вибір підмножини прогнозів, агрегування та оцінювання похибки. Така схема є методичною основою для подальшого викладення конкретних алгоритмів адаптивного прогнозування, зокрема методів на основі критерію розбіжності, екстремумів PPS, інтервалів концентрації та кластерного аналізу.

У підрозділі формалізовано алгоритм прогнозування біржового часового ряду з адаптивним вибором параметра *m*, який визначає кількість елементів PPS, що використовуються для усереднення. На відміну від підходів із наперед заданим степенем полінома, запропонована схема вибирає локальну глибину екстраполяції за критерієм Δ_m, побудованим на основі розбіжності між поліноміальним прогнозом та додатковим прогнозом, отриманим через скінченні різниці.

Остаточний прогноз формується як середнє значення перших m^* елементів PPS. Така схема поєднує переваги локальної поліноміальної екстраполяції та усереднення прогнозів, зменшуючи вплив окремих високостепеневих поліномів на результат. Алгоритм є обчислювально простим, інтерпретованим і придатним для роботи з короткими фрагментами внутрішньоденних біржових рядів.

РОЗДІЛ 4. РОЗРОБЛЕННЯ АРХІТЕКТУРИ НЕЙРОННОЇ МЕРЕЖІ ДЛЯ ПРОГНОЗУВАННЯ БІРЖОВИХ ЧАСОВИХ РЯДІВ НА ОСНОВІ ПОЛІНОМІАЛЬНОГО ПІДХОДУ

4.1. Обґрунтування необхідності нейромережевого розширення поліноміального підходу

Нехай задано локальний фрагмент часового ряду: $X = (x_1, x_2, \ldots, x_n)$. Для нього формується послідовність поліноміальних прогнозів: $PPS = (P_0, P_1, \ldots, P_{n-1})$, де P_m – прогноз наступного значення, отриманий на основі поліноміальної екстраполяції порядку m. На основі PPS можна сформувати послідовність усереднених прогнозів:

$$l_m = \frac{1}{m}\sum_{i=0}^{m-1} P_i, \qquad m = 1, \ldots, n.$$

Кожне значення l_m є окремим кандидатом на підсумковий прогноз. Відповідно, задача прогнозування зводиться до адаптивного вибору такого індексу m, для якого прогноз l_m найкраще відповідає локальній конфігурації ряду.

Алгоритмічні способи вибору параметра m, зокрема використання внутрішніх критеріїв узгодженості прогнозів або кластеризації значень PPS, дозволяють враховувати структуру поліноміальної послідовності. Водночас такі підходи ґрунтуються на наперед заданих правилах і можуть бути чутливими до вибору параметрів. Це обмежує їхню здатність адаптуватися до різних локальних режимів біржової динаміки.

З огляду на це виникає необхідність розроблення PPS-орієнтованої нейромережевої архітектури, у якій послідовність поліноміальних прогнозів інтегрується безпосередньо в структуру моделі. На відміну від традиційної нейронної мережі, яка навчається формувати прогноз лише на основі початкових значень часового ряду, PPS-орієнтована архітектура використовує проміжний прогнозний шар, компоненти якого мають явний математичний зміст. Кожен нейрон цього шару відповідає поліноміальному прогнозу певного порядку.

Такий підхід дозволяє поєднати переваги аналітичних і навчальних методів. Поліноміальний шар формує інтерпретовану множину прогнозних

ознак, а селекторний шар виконує адаптивний вибір кандидата залежно від локальної структури PPS. У результаті підсумкове прогнозне значення не формується довільно, а обирається з множини математично обґрунтованих альтернатив.

У межах дослідження наукова проблема формулюється так: необхідно розробити PPS-орієнтовану нейромережеву архітектуру для однокрокового прогнозування фінансових часових рядів, яка забезпечує адаптивний вибір підсумкового прогнозу на основі аналізу послідовності поліноміальних прогнозних значень, враховує дисбаланс можливих класів прогнозів і обмежує вплив аномальних значень, що виникають унаслідок використання поліноміальної екстраполяції підвищених порядків.

Для розв’язання сформульованої проблеми необхідно:

1) сформувати спеціалізований шар, нейрони якого обчислюють поліноміальні прогнози різних порядків;
2) сформувати спеціалізований шар нейронів для побудуви усереднення елементів PPS;
3) розробити селекторний шар для адаптивного вибору підсумкового прогнозного значення;
4) в процесі навчання визначити вагові коефіцієнти селекторного шару з урахуванням дисбалансу класів і необхідності регуляризації;
5) експериментально порівняти запропоновану архітектуру з алгоритмічними способами формування прогнозу на основі PPS.

Отже, розроблення PPS-Net спрямоване на створення інтерпретованої гібридної архітектури, яка поєднує поліноміальну екстраполяцію, структурний аналіз послідовності прогнозів та адаптивний вибір підсумкового значення. Це дозволяє розглядати PPS не лише як допоміжну множину прогнозів, а як повноцінне внутрішнє представлення локальної динаміки фінансового часового ряду.

4.2. Архітектура нейронної мережі з автоматичним вибором параметра

Виходячи з розглянутих вище завдань заппропонуємо таку архітектуру нейронної мережі (рис. 4.1):

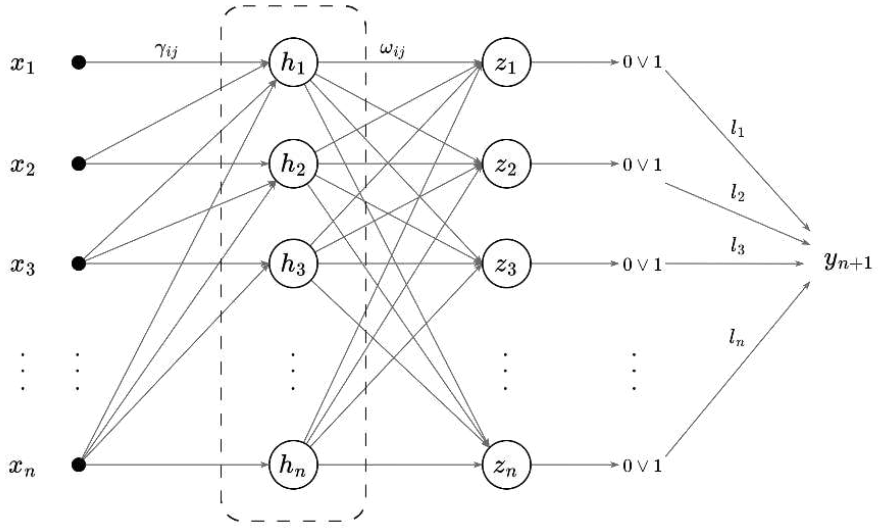


Рис. 4.1. Структура нейронної мережі PPS-Net.

Нейронна мережа складається з трьох основних шарів нейронів та адитивного суматора:

$$X \to H \to Z \to y_{n+1}$$

На вхід подаються дані часового ряду $X = (x_1, x_2, \ldots, x_n)$. Перший шар формує поліноміальні прогнозні значення: $h_i = P_{n-1}(X), i = 1,2,\ldots,n.$ Кожен поліноміальний прогноз обчислюється за формулою (4.1):

$$P_m(X) = \sum_{k=1}^{m+1} (-1)^{k-1} C_{m+1}^k x_{n-k+1}., \qquad (4.1)$$

Вагові коефіцієнти першого шару вважаються вже навченими і визначаються, як $\gamma_{ij} = (-1)^j C_i^j, i = \overline{1,n},\ j = \overline{1,n}.$ Активаційна функція першого шару H лінійна та має вигляд $f(x) = x$. Таким чином шар H формує не один прогноз, а множину альтернативних прогнозних значень, що відповідають різним степеням полінома. Ця множина розглядається як послідовність поліноміальних прогнозів PPS (Polynomial Prediction Set).

Після формування значень h_i для кожного m обчислюється величина $l_m = \frac{1}{m} \sum_{i=1}^{m} h_i, m = 1,2,\ldots,n$, тобто формується послідовність усереднених

поліноміальних прогнозів $L=(l_1,l_2,\ldots,l_n)$. Саме значення l_m і є кандидатами на підсумковий прогноз. Нейронна мережа повинна навчитися вибирати, яке саме значення l_m є найбільш доцільним для поточного вхідного фрагмента.

Третій рівень Z є навчальним шаром. Кожен нейрон шару H зв'язаний із кожним нейроном шару Z. На кожному зв'язку між нейроном h_j та нейроном z_i знаходиться невідомий ваговий коефіцієнт ω_{ij}. Активаційна функція цього шару має вигляд (4.2):

$$z_i=\begin{cases}1, & якщо \sum_{i=1}^{n}\omega_{ij}h_i>\delta\\ 0, & якщо \sum_{i=1}^{n}\omega_{ij}h_i\le\delta\end{cases}, \tag{4.2}$$

В запропонованій архітектурі коефіцієнт зміщення не використовується, тому шар Z задається матрицею ваг $\Omega=(\omega_{ij})_{n\times n}$. Шар Z виконує функцію вибору одного з варіантів l_m. Його вихід має вигляд one-hot (унітарного) вектора: $Z=(z_1,z_2,\ldots,z_n)$, де тільки один нейрон має значення 1, а всі інші 0.

Будемо називати описану вище нейронну мережу PPS-Net (Polynomial Prediction Sequence Network).

Для навчання мережі та пошуку відповідних коефіцієнтів ω_{ij} запропоновано декілька підходів.

4.3. Алгоритми навчання та використання запропонованої нейромережевої моделі

Softmax-підхід до навчання нейронної мережі

Для кожного навчального вікна $X=(x_1,x_2,\ldots,x_n)$ обчислюється шар $H=\left(h_1,h_2,\ldots,h_n\right)$, де $h_i=P_{n-i}(X)$. Потім обчислюються значення $l_m=\frac{1}{m}\sum_{i=1}^{m}h_i,\quad m=1,2,\ldots,n.$ Оскільки для навчального прикладу відоме фактичне наступне значення x_{n+1}, то оптимальне значення l_m визначається, як:

$$m^*=\arg\min_m\left|l_m-x_{n+1}\right|.$$

Знайдене значення m^* є правильною відповіддю для шару Z. Таким чином мережа навчається вибирати той нейрон z_m, який відповідає найкращому усередненому поліноміальному прогнозу. Для кожного нейрона шару Z обчислюється сигнал (логіт):

$$s_i = \sum_{i=1}^{n} \omega_{ij} h_i.$$

Після цього до значень s_i застосовується функція softmax:

$$p_i = \frac{\exp(s_i)}{\sum_{q=1}^{n} \exp(s_q)}.$$

Завдяки нормалізації, всі отримані значення p_i знаходяться в діапазоні від 0 до 1, а їхня загальна сума дорівнює 1. Це дозволяє математично коректно трактувати їх як ймовірності вибору відповідного варіанта l_i. Навчання полягає в тому, щоб підібрати такі ω_{ij}, щоб для правильного індексу m^* значення p_{m^*} було максимальним.

Оскільки шар Z виконує вибір одного з n можливих усереднених поліноміальних прогнозів l_m, навчання мережі зводиться до задачі класифікації. Для кожного навчального вікна правильним класом вважається індекс m^*, для якого значення l_m має мінімальне відхилення від фактичного наступного значення ряду. Тому, функція втрат (функція перехресної ентропії) для одного навчального прикладу має вигляд:

$$\mathcal{L} = -\ln p_i,$$

а для всіх навчальних прикладів:

$$\mathcal{L} = -\frac{1}{R} \sum_{r=1}^{R} \sum_{i=1}^{n} y_i^{(r)} \ln p_i^{(r)},$$

де R – кількість навчальних вікон, а

$$y_i = \begin{cases} 1, & i = m^*, \\ 0, & i \neq m^*. \end{cases}$$

Коефіцієнти ω_{ij} оновлюються методом градієнтного спуску або його сучасною модифікацією, наприклад Adam:

$$\omega_{ij} \leftarrow \omega_{ij} - \eta \frac{\partial \mathcal{L}}{\partial \omega_{ij}},$$

де η – швидкість навчання.

Спеціалізований метод навчання нейронної мережі

Для визначення параметрів запропонованої нейромережевої архітектури використано навчання з учителем. Метод навчання ґрунтується на інтерпретації виходу третього шару як результату вибору оптимального варіанта усереднення послідовності поліноміальних прогнозів.

Відповідно, цільовий вихід третього шару задається у вигляді one-hot-вектора: нейрон, що відповідає індексу m, набуває значення 1, а виходи решти нейронів дорівнюють 0. Вагові коефіцієнти зв'язків між другим і третім шарами визначаються таким чином, щоб для кожного навчального вікна максимальний сигнал формувався саме на виході нейрона, який відповідає оптимальному усередненому прогнозу.

Нехай маємо навчальні пари для кожного навчального вікна:

$$H^i \to Y^i, \quad \text{де } H^i = \begin{pmatrix} P^i_{n-1}(X) \\ P^i_{n-2}(X) \\ \vdots \\ P^i_0(X) \end{pmatrix}, \quad Y^i = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \tag{4.3}$$

Кожному вікну відповідає система рівнянь:

$$\begin{cases} P^i_0\omega_{0n-1} + P^i_1\omega_{1n-1} + \ldots + P^i_{n-1}\omega_{n-1n-1} = 0 \\ P^i_0\omega_{0n-2} + P^i_1\omega_{1n-2} + \ldots + P^i_{n-2}\omega_{n-2n-2} = 0 \\ \vdots \\ P^i_0\omega_{0n-i} + P^i_1\omega_{1n-i} + \ldots + P^i_{n-i}\omega_{n-in-i} = \delta_{emp} \\ \vdots \\ P^i_0\omega_{00} = 0 \end{cases}, \tag{4.4}$$

де δ_{emp} – деяке числове значення, яке обирається емпіричним шляхом, наприклад, $\delta_{emp} = 0{,}1$.

Матриці коефіцієнтів при невідомих ω_{ij} мають вигляд:

$$\begin{pmatrix} 0 & 0 & \dots & 0 & 0 & \dots & 0 & P_0^i & P_1^i & \dots & P_{n-1}^i \\ 0 & \dots & 0 & P_0^i & P_1^i & \dots & P_{n-2}^i & 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ P_0^i & 0 & \dots & 0 & 0 & \dots & \dots & \dots & \dots & \dots & 0 \end{pmatrix}, \qquad (4.5)$$

Розв'язуючи вказану СЛАР одним із існуючих методів, знаходяться коефіцієнти ω_{ij}. Тоді, вихід кожного нейрона шару Z має вигляд:

$$Out_{n-1} = \begin{cases} 1, & \sum_{i=0}^{n-1} P_i \omega_{i,n-1} > \delta_{cr} \\ 0, & \sum_{i=0}^{n-1} P_i \omega_{i,n-1} \le \delta_{cr} \end{cases},$$

$$Out_{n-2} = \begin{cases} 1, & \sum_{i=0}^{n-2} P_i \omega_{i,n-2} > \delta_{cr} \\ 0, & \sum_{i=0}^{n-2} P_i \omega_{i,n-2} \le \delta_{cr} \end{cases}, \qquad (4.6)$$

$$\vdots$$

$$Out_0 = \begin{cases} 1, & P_0 \omega_{0,0} > \delta_{cr} \\ 0, & P_0 \omega_{0,0} \le \delta_{cr} \end{cases},$$

де δ_{cr} – деяке числове значення, яке може бути визн ачено, наприклад, як $\delta_{cr} = \dfrac{\delta_{emp}}{2}$ для того, щоб однозначно виконувались нерівності в умовах виходів (4.6).

У випадку, коли рівнянь під час навчання мережі більше, ніж невідомих коефіцієнтів, використано підхід, при якому застосовано метод найменших квадратів. В матричному вигляді:

$$\sum_{r=1}^{R} (\Omega H^r - Y^r)^2 \to \min.$$

Експериментальна перевірка запропонованої архітектури виконувалася мовою програмування Python. Для завантаження біржових даних використано бібліотеку yfinance, для чисельних обчислень – NumPy та pandas, для реалізації методу DBSCAN – scikit-learn, для побудови графіків – matplotlib.

Як експериментальний часовий ряд використано внутрішньоденні біржові котирування акцій Netflix за тікером NFLX за останні 60 днів із 30-хвилинним інтервалом спостережень. Для прогнозування використовувався параметр Close. Дані опрацьовувалися у хронологічному порядку без випадкового перемішування, оскільки задача належить до класу задач прогнозування часових рядів (рис. 4.2).

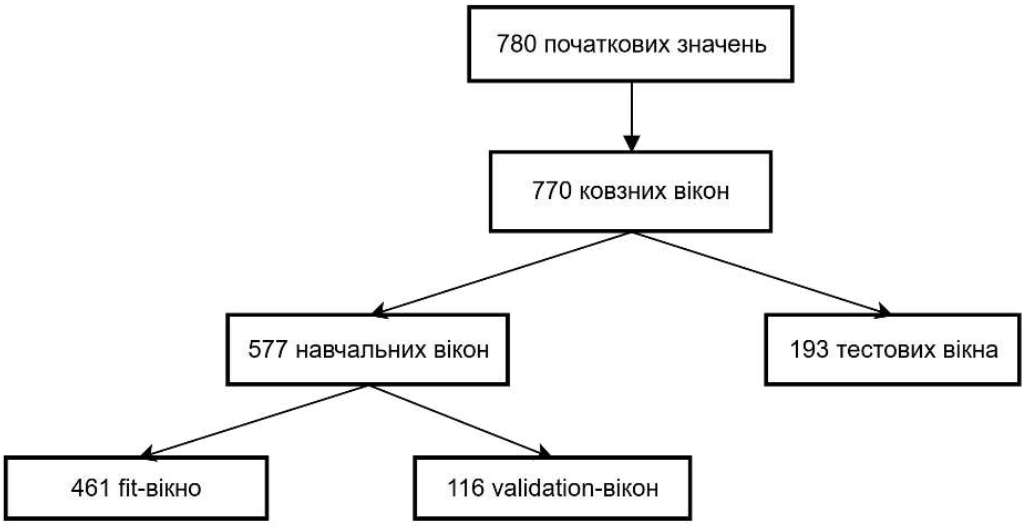


Рис. 4.2. Загальна схема поділу часового ряду та формування вікон.

Початковий набір даних містив 780 внутрішньоденних значень ціни Close акцій NFLX. Для кожного прогнозу формувалося ковзне вікно з 10 послідовних значень та одне цільове значення, що відповідало наступному часовому кроку. У результаті було сформовано 770 прогнозних вікон. Дані поділялися хронологічно: перші 75 % вікон використовувалися як навчальна вибірка, а останні 25 % – як тестова. Внаслідок цілочисельного округлення було отримано 577 навчальних і 193 тестові вікна. Для вибору параметра регуляризації λ (4.8) навчальна частина додатково поділялася на 461 fit-вікно і 116 validation-вікон (рис. 4.2).

Для кожного навчального прикладу формувалося ковзне вікно з 10 послідовних значень $X=(x_1,x_2,\ldots,x_{10})$, де x_{10} – останнє відоме значення у поточному вікні. Цільовим значенням було наступне значення ряду x_{11}. На основі кожного вікна формувалася послідовність поліноміальних прогнозів порядків від нульового до дев'ятого: $PPS=(P_0,P_1,\ldots,P_9)$. Поліноміальний шар архітектури мав вигляд:

$$H=(h_1,h_2,\ldots,h_{10})=(P_{9,}P_8,\ldots,P_0).$$

Кожне значення P_m обчислювалося за формулою (4.1). На основі послідовності $PPS=(P_0,P_1,\ldots,P_9)$ була сформована множина усереднених поліноміальних прогнозів $l_m=\frac{1}{m}\sum_{i=1}^{m-1}P_i, m=1,2,\ldots,10$. Проведені експерименти показали, що використання абсолютних значень $H=(P_{9,}P_8,\ldots,P_0)$ може призводити до виродження селектора до вибору $l_1=P_0$, тобто до наївного прогнозу. Тому в остаточній реалізації для селекторного шару використовувалися не абсолютні значення H, а структурні ознаки PPS:

$$F_i=\frac{h_i-P_0}{|P_0|-\varepsilon},\quad i=1,2,\ldots,10, \tag{4.7}$$

де ε – деяке мале додатне число, яке використовується для уникнення ділення на нуль. Такий перехід дозволяє аналізувати не рівень ціни, а форму послідовності поліноміальних прогнозів відносно базового прогнозу P_0.

Для кожного навчального вікна відоме фактичне наступне значення x_{11}. Тому правильний клас визначався як індекс:

$$m^*=\arg\min_m|l_m-x_{11}|.$$

У фінальній реалізації вагові коефіцієнти ω_{ij} між шаром структурних ознак F та селекторним шаром Z визначалися не ітераційним навчанням нейронної мережі, а за допомогою зваженого регуляризованого методу найменших квадратів.

Для навчальної вибірки формувалася матриця структурних ознак F_{train} і матриця one-hot векторів правильних класів Y. Матриця ваг W визначалася, як розв'язок задачі:

$$W^* = \arg\min_W \left(\left\| D^{1/2} \left(F_{train} W - Y \right) \right\|^2 + \lambda \left\| W \right\|^2 \right), \tag{4.8}$$

де D – діагональна матриця ваг навчальних прикладів, що використовується для компенсації дисбалансу класів, а λ – параметр регуляризації. Елементи матриці D обчислюються наступним чином:

$$d_c = \frac{R}{M \cdot n_c},$$

де R – кількість навчальних прикладів, M – кількість класів $l_1,\ldots,l_{10}$, n_c – кількість прикладів класу c. В нашому випадку $R = 577$, $M = 10$. Таким чином матриця D формується на основі частот появи класів l_m навчальній вибірці. Це дозволяє зменшити вплив домінантних класів, зокрема $l_1 = P_0$, і підвищити внесок рідкісних класів під час знаходження матриці ваг W.

Параметр регуляризації λ обирався експериментально на валідаційній частині навчальної вибірки. Для цього розглядалася скінченна множина кандидатів $\lambda \in \left\{10^{-8}, 10^{-6}, 10^{-4}, 10^{-3}, 10^{-2}, 10^{-1}, 1, 10, 100\right\}$. Для кожного значення λ знаходилася матриця ваг W селекторного шару, після чого оцінювалася якість прогнозування на валідаційній вибірці. Оптимальним вважалося таке значення λ, для якого досягалося мінімальне значення MAPE (Mean Absolute Percentage Error) на валідаційних даних.

Після знаходження матриці W для кожного тестового прикладу обчислювалися сигнали селекторного шару $S = FW$. Вибраним вважався той кандидат l_m, для якого відповідний сигнал був максимальним:

$$m = \arg\max_i S_i.$$

Підсумковий прогноз визначався як $x_{11} = l_m$.

Оскільки навіть після використання структурних ознак окремі кандидати l_m можуть давати недопустимі прогнозні значення, до прогнозу PPS-Net було

додано фільтр стабільності. Він перевіряє, наскільки вибране значення l_m є віддаленим від базового прогнозу P_0.

Критерій фільтрації задавався у вигляді:

$$|l_m - P_0| > \tau |P_1 - P_0|, \tag{4.9}$$

де τ – параметр допустимого відхилення. У програмній реалізації використовувалося емпіричне значення $\tau = 5$. Отже, правило формування підсумкового прогнозу має вигляд:

$$x_{11} = \begin{cases} l_m, & |l_m - P_0| \le \tau |P_1 - P_0| \\ P_0, & |l_m - P_0| > \tau |P_1 - P_0| \end{cases}.$$

Запропонований механізм дозволяє обмежити вплив аномальних значень, які можуть виникати внаслідок посилення локальних коливань при використанні поліноміальної екстраполяції підвищених порядків. Параметр τ доцільно визначати за валідаційною вибіркою або шляхом аналізу чутливості.

Висновки до четвертого розділу

У розділі запропоновано та експериментально перевірено нейромережеву архітектуру прогнозування фінансових часових рядів на основі послідовності поліноміальних прогнозів.

Матриця ваг селекторного шару визначалася за допомогою зваженого регуляризованого методу найменших квадратів. Використання ваг навчальних прикладів дозволило врахувати дисбаланс класів l_m, а регуляризація – обмежити надмірне зростання коефіцієнтів моделі. Додатково було застосовано фільтр локальної допустимості прогнозу, який замінював вибране значення l_m на базовий прогноз P_0, якщо відхилення виходило за межі локально допустимого діапазону.

Експериментальну перевірку проведено на внутрішньоденних біржових даних акцій Netflix за тікером NFLX. Використано значення ціни Close за останні 60 торгових днів із 30-хвилинним інтервалом. Після формування ковзних вікон із 10 послідовних значень було отримано 577 навчальних і 193 тестові приклади.

Для кожного вікна виконувався однокроковий прогноз наступного значення часового ряду.

Аналіз розподілу еталонних класів у навчальній вибірці показав, що для прогнозування внутрішньоденних фінансових даних доцільно використовувати поліноміальні прогнози невисоких порядків. Водночас наявність інших класів обґрунтовує необхідність адаптивного вибору параметра m, а не використання одного фіксованого прогнозного правила.

Подальші дослідження доцільно спрямувати на перевірку методу для різних тікерів, часових інтервалів і ринкових режимів; проведення аналізу чутливості до параметрів λ і τ; порівняння з класичними статистичними, машинними та нейромережевими методами прогнозування; а також на розроблення процедури автоматичного визначення допустимого локального діапазону прогнозу.

РОЗДІЛ 5. ЕКСПЕРИМЕНТАЛЬНЕ ДОСЛІДЖЕННЯ ЕФЕКТИВНОСТІ РОЗРОБЛЕНИХ МЕТОДІВ

5.1. Опис експериментальних даних

Експериментальне дослідження ефективності розроблених методів прогнозування часових рядів біржової динаміки виконувалося на основі реальних біржових котирувань, отриманих з відкритих фінансових джерел. Використання саме ринкових даних, а не лише модельних або синтетичних послідовностей, є принциповим для перевірки прикладної придатності запропонованих алгоритмів, оскільки фінансові часові ряди характеризуються нерівномірністю локальних змін, наявністю шумових складових, короткочасних імпульсів, зміною волатильності та обмеженою інформативністю коротких фрагментів спостережень. Подібні властивості фінансових рядів у класичному вигляді систематизовано у працях Р. Конта, де описано емпіричні властивості дохідностей активів, зокрема важкі хвости розподілів, кластеризацію волатильності та нелінійні залежності [1]. У роботах Е. Ло та А. Мак-Кінлея також показано, що припущення про випадкове блукання цін не завжди узгоджується з емпіричними даними, що обґрунтовує доцільність пошуку локальних закономірностей у фінансових рядах [157].

У сучасних дослідженнях прогнозування фінансових часових рядів значна увага приділяється вибору експериментальних даних, оскільки результат порівняння методів істотно залежить від частоти спостережень, довжини навчальної вибірки, горизонту прогнозування та способу формування тестових фрагментів. У систематичному огляді О. Б. Сезера, М. У. Гюделека та А. М. Озбайоглу зазначено, що фінансові часові ряди є одним із провідних об’єктів застосування методів машинного та глибокого навчання, а тип даних і схема експерименту безпосередньо впливають на коректність оцінювання моделей [75]. Огляд Ч. Чжана, Н. Сжаріф та Р. Ібрагім додатково підкреслює, що в задачах прогнозування цін фінансових активів важливими є не лише архітектура моделі, а й структура вхідної вибірки, спосіб подання локальної динаміки та розмежування навчальних і тестових даних [158].

У межах української наукової школи задачі моделювання та прогнозування складних процесів пов’язані, зокрема, з працями О. Г. Івахненка, який розвинув індуктивний підхід до побудови моделей складних систем і метод групового урахування аргументів. Для цього напряму характерне використання експериментальних даних як основи для автоматизованого формування та відбору моделей оптимальної складності, у тому числі поліноміального типу [33]. Запропонований поліноміальний підхід також спирається на локальне представлення даних і адаптивний вибір параметрів прогнозування.

Окремий внесок українських дослідників у проблематику аналізу фінансових процесів пов’язаний із працями П. І. Бідюка, І. О. Калініної, О. П. Гожого та Н. В. Кузнєцової, у яких розглядаються методи аналізу, моделювання та прогнозування фінансових процесів, зокрема питання волатильності, імовірнісно-статистичного моделювання та обробки неповної інформації [159; 160]. Такі підходи підтверджують необхідність деталізованого опису експериментальної вибірки, оскільки фінансові дані не є нейтральним набором числових значень, а відображають стохастично-детермінований процес із мінливою локальною структурою.

Як основний експериментальний набір у роботі використано внутрішньоденні біржові котирування акцій Netflix, Inc. з тікером NFLX, що котируються на біржі NASDAQ. Історичні дані щодо інструмента NFLX доступні через відкриті фінансові сервіси, зокрема Yahoo Finance та сторінку інвесторів Netflix [161; 162]. Для автоматизованого завантаження даних у середовищі Python використовувалась бібліотека yfinance, яка надає доступ до ринкових даних Yahoo Finance API та підтримує завантаження історичних котирувань із різними часовими інтервалами [163]. Згідно з документацією yfinance, допустимими внутрішньоденними інтервалами є, зокрема, $1m$, $2m$, $5m$, $15m$, $30m$, $60m$, $90m$, $1h$, а внутрішньоденні дані мають обмеження щодо глибини завантаження до останніх 60 днів [163].

У роботі використано часовий інтервал 30 хвилин, що забезпечує компроміс між надмірною деталізацією високочастотних тикових даних і надто

грубим поданням денних котирувань. Такий таймфрейм дозволяє досліджувати короткострокову біржову динаміку в межах торговельної сесії, але водночас зменшує вплив випадкових мікроколивань, характерних для даних вищої частоти. В дослідженнях ринкової мікроструктури зазначається, що високочастотні дані відображають особливості сучасної електронної торгівлі, зокрема вплив швидкісного трейдингу та зміну характеру торговельного процесу [164]. Тому використання агрегованих 30-хвилинних барів є доцільним для перевірки методів короткострокового прогнозування без переходу до спеціалізованих моделей тикової мікроструктури.

Експериментальний набір містить стандартні поля біржових OHLCV-даних:

$$Open_t, High_t, Low_t, Close_t, Volume_t$$

де $Open_t$ – ціна відкриття інтервалу, $High_t$ – максимальна ціна в межах інтервалу, Low_t – мінімальна ціна, $Close_t$ – ціна закриття інтервалу, $Volume_t$ – обсяг торгів за відповідний інтервал часу.

В запропонованому дослідженні в якості цільової змінної використано саме ціну закриття $Close_t$, оскільки вона є узагальненим результатом торговельної активності протягом відповідного 30-хвилинного інтервалу та найчастіше застосовується у задачах порівняння методів прогнозування біржової динаміки. Таким чином, початковий багатовимірний запис OHLCV було зведено до одновимірного часового ряду цін закриття:

$$F = \{f_t\}_{t=1}^{N}, \quad f_t = Close_t, \tag{5.1}$$

де N – кількість спостережень у часовому ряді після попередньої обробки даних.

Вибір ціни закриття як основної змінної зумовлений тим, що запропоновані у роботі методи орієнтовані на короткострокове прогнозування наступного значення біржового часового ряду на основі локального ретроспективного фрагмента спостережень. У цьому випадку використання одного цінового показника дозволяє уникнути змішування декількох різнорідних ознак та безпосередньо оцінити властивості поліноміального прогнозування.

В більшості досліджень, пов’язаних з аналізом та прогнозуванням біржових котирувань, для аналізу локальних змін розглядають ряд логарифмічних приростів: $r_t = \ln \frac{f_t}{f_{t-1}}, \quad t = 2,3,\ldots,N.$

Проте в основному експерименті прогнозуванню підлягає не r_t, а безпосередньо значення f_{t+1}. Такий вибір відповідає постановці задачі дисертаційного дослідження, де результатом роботи алгоритму є точкова оцінка наступного значення біржового ряду:

$$\hat{f}_{t+1} = A(F_t^{(L)}), \tag{5.2}$$

де $A(\cdot)$ – алгоритм прогнозування, а $F_t^{(L)}$ – локальний фрагмент спостережень довжини L.

Для забезпечення однакових умов порівняння всіх методів початковий часовий ряд було перетворено на послідовність ковзних фрагментів. Кожен такий фрагмент містить L останніх відомих значень ряду, а цільовим значенням є наступне значення після цього фрагмента.

Формально локальний ретроспективний фрагмент задається так:

$$F_t^{(L)} = (f_{t-L+1}, f_{t-L+2}, \ldots, f_t). \tag{5.3}$$

Цільове значення для цього фрагмента: $y_t = f_{t+1}$. У проведених експериментах для нейромережевого порівняння використовувалася довжина вхідного фрагмента $L = 15$. Тобто кожен навчальний або тестовий приклад формувався за схемою:

$$(f_{t-14}, f_{t-13}, \ldots, f_t) \to f_{t+1}. \tag{5.4}$$

Така схема відповідає однокроковому прогнозуванню, при якому модель не використовує майбутніх значень ряду і кожного разу формує прогноз лише на основі доступного на поточний момент локального фрагмента.

Для методів, пов’язаних із формуванням послідовності поліноміальних прогнозних значень, із вхідного фрагмента виділялася частина останніх спостережень, на основі яких будувалася PPS. У реалізованому

експериментальному протоколі для побудови послідовності поліноміальних прогнозів використовувалися $M = 10$ останніх значень локального фрагмента:

$$F_t^{(M)} = (f_{t-M+1}, f_{t-M+2}, \ldots, f_t), \quad M = 10. \tag{5.5}$$

Тут M визначає максимальну кількість останніх спостережень, що залучаються до формування PPS, а відповідний номер елемента послідовності поліноміальних прогнозів пов'язаний зі степенем локальної поліноміальної екстраполяції.

Оскільки часові ряди мають природний порядок спостережень, у роботі не застосовувалося випадкове перемішування прикладів. Поділ виконано зі збереженням хронологічної послідовності: перша частина сформованих фрагментів використовувалась для налаштування моделей, а остання – для перевірки якості прогнозування. Такий підхід відповідає загальноприйнятим рекомендаціям щодо оцінювання прогнозних моделей на часових рядах. Л. Ташман підкреслював доцільність позавибіркового тестування та схем із рухомим початком прогнозування для підвищення надійності експериментального оцінювання [165]. К. Бергмейр і Х. Бенітес також розглядали особливості застосування крос-валідації до часових рядів та обґрунтовували необхідність спеціальних схем перевірки, що враховують часову залежність даних [166].

Загальна кількість сформованих ковзних прикладів у робочому наборі становила $N_w = 770$. З них для навчальної частини використано $N_{train} = 577$, а для тестової частини $N_{test} = 193$.

У відсотковому співвідношенні такий поділ наближено відповідає схемі $75\%/25\%$, що дозволяє отримати достатню кількість прикладів для навчання моделі та водночас зберегти окрему тестову частину для незалежного оцінювання результатів. Узагальнена характеристика експериментального набору подана в табл. 5.1.

Таблиця 5.1. Основні характеристики експериментального набору даних

Характеристика	**Значення**
Фінансовий інструмент	Netflix, Inc.
Біржовий тікер	NFLX
Ринок	NASDAQ
Джерело даних	Yahoo Finance / yfinance
Тип даних	внутрішньоденні OHLCV-дані
Часовий інтервал	30 хвилин
Цільова змінна	$Close_t$
Тип задачі	однокрокове прогнозування
Довжина вхідного фрагмента	$L = 15$
Кількість значень для формування PPS	$M = 10$
Кількість ковзних прикладів	$N_w = 770$
Навчальна частина	$N_{train} = 577$
Тестова частина	$N_{test} = 193$

Попередня обробка експериментальних даних передбачала виконання таких етапів:

1) завантаження внутрішньоденних котирувань NFLX з інтервалом 30 хвилин;
2) відбір поля $Close_t$ як основної змінної прогнозування;
3) перевірку наявності пропущених значень;
4) упорядкування записів за часовою міткою;
5) формування ковзних локальних фрагментів;
6) хронологічний поділ на навчальну та тестову частини;
7) збереження однакового набору тестових прикладів для всіх порівнюваних методів.

У випадку наявності пропущених значень відповідні записи вилучалися або не включалися до формування ковзних фрагментів. Для збереження коректності однокрокового прогнозування не використовувалося заповнення

пропусків майбутніми значеннями або інтерполяція, яка могла б створити приховане інформаційне зміщення.

Контроль відсутності пропущених значень у цільовому ряді можна формально подати так:

$$\left|\{t : f_t = \varnothing, t = 1,2,\ldots,N\}\right| = 0. \tag{5.6}$$

Ця умова означає, що після попередньої обробки всі значення f_t, які використовуються для побудови навчальних і тестових фрагментів, є визначеними.

Сформований набір даних використовується для перевірки трьох груп результатів дисертаційного дослідження:

1) методів формування послідовності поліноміальних прогнозних значень;
2) алгоритмів адаптивного вибору інформативної частини PPS;
3) нейромережевої архітектури, що використовує PPS як спеціалізоване представлення локальної біржової динаміки.

Особливість обраного набору полягає в тому, що він містить достатню кількість коротких локальних фрагментів для багаторазового walk-forward оцінювання, але водночас не є надмірно великим. Це відповідає одній із ключових умов дисертаційного дослідження – перевірці методів у ситуації обмеженого обсягу доступних даних. Саме в такому режимі доцільно оцінювати поліноміальні методи, оскільки вони не потребують тривалого навчання великої кількості параметрів і можуть формувати прогноз на основі короткого ретроспективного фрагмента.

На відміну від задач довгострокового прогнозування, у цьому експерименті основна увага приділяється локальній зміні ціни $f_t \to f_{t+1}$, тобто прогнозуванню наступного 30-хвилинного значення після поточного моменту часу. Такий формат дозволяє оцінити не загальну здатність моделі описувати тривалий тренд, а її придатність до короткострокового реагування на зміну локальної структури біржового ряду.

Отже, експериментальний набір даних сформовано таким чином, щоб забезпечити порівнюваність результатів усіх досліджуваних методів, відтворюваність процедури оцінювання та відповідність постановці задачі короткострокового прогнозування часових рядів біржової динаміки на основі поліноміального підходу.

5.2. Проведення обчислювальних експериментів з використанням кластерного аналізу

Експериментальні дослідження було проведено з метою оцінки ефективності запропонованого методу прогнозування на основі кластеризації значень послідовності поліноміальних прогнозів та його порівняння з методом прогнозування на основі усереднення поліноміальної екстраполяційної послідовності [129]. В усіх експериментах використовувалась єдина схема побудови послідовності поліноміальних прогнозів $P = \{p_1, p_2, \ldots, p_m\}$ (1) на заданому діапазоні значень m. Після чого для цієї множини застосовувалися два підходи кластерного аналізу: метод найщільнішого інтервалу (DI) та метод DBSCAN.

Дослідження проведено для трьох класів даних:

- детерміновані функції;
- стохастичні (випадкові) послідовності;
- реальні фінансові часові ряди.

Для кожного типу даних виконано побудову послідовності прогнозів p_m при змінних параметрах $m \in [m_{min}, m_{max}]$, після чого досліджено їхню структуру, щільність розподілу та поведінку.

У якості тестових функцій використано:

$y = \cos(hx), \quad y = \cos(h \ln x), \quad y = e^{hx}$.

На Рис. 5. 1 показано фрагмент графіка функції $y = \cos(hx)$ та графік послідовності поліноміальних прогнозів із такими заданими параметрами: $x_0 = 1$ (початкове значення), $N = 30$ (кількість історичних точок), $h = 0,3$ (параметр), $\Delta = 1$ (крок), $m_{\min} = 2, m_{\max} = 20$ (мінімальне та максимальне значення кількості

точок для послідовності PPS), $\varepsilon = 0{,}5$ (радіус локального сусідства для DBSCAN).

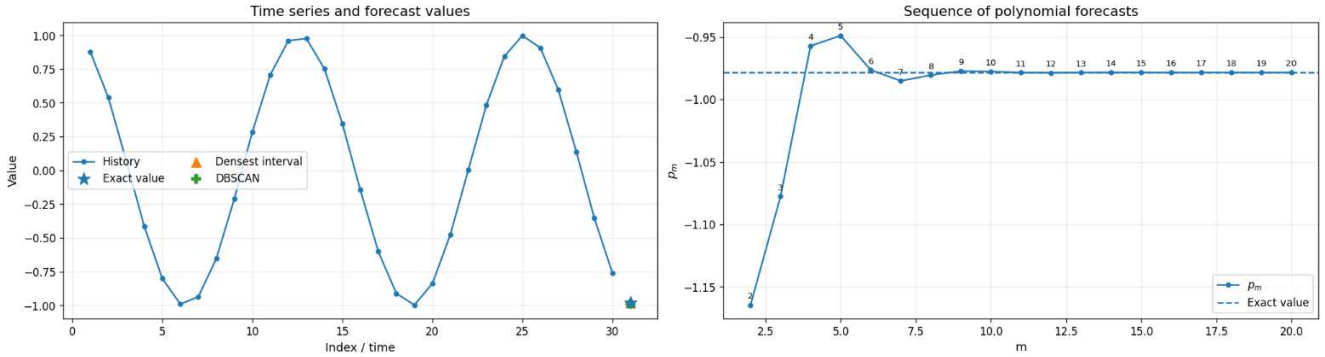


Рис. 5.1. Графік функції $y = \cos(hx)$ та її послідовність поліноміальних прогнозів (PPS).

Для детерміновнаних функцій, як видно з Рис. 5.1, послідовність поліноміальних прогнозів p_m демонструє структуровану та передбачувану поведінку, а саме:

- спостерігаються локальні екстремуми;
- відсутній випадковий шум;
- значення концентруються в обмеженому інтервалі.

Це пояснюється тим, що вхідні дані мають регулярну структуру, а поліноміальна екстраполяція фактично відновлює локальну гладкість функції.

Розподіл значень PPS на числовій осі, структуру кластера, гістограму та оцінку щільності KDE (Kernel Density Estimation) показано на Рис. 5.2. Для кращої візуалізації розміщення точок, було використано параметр вертикального зміщення (jitter).

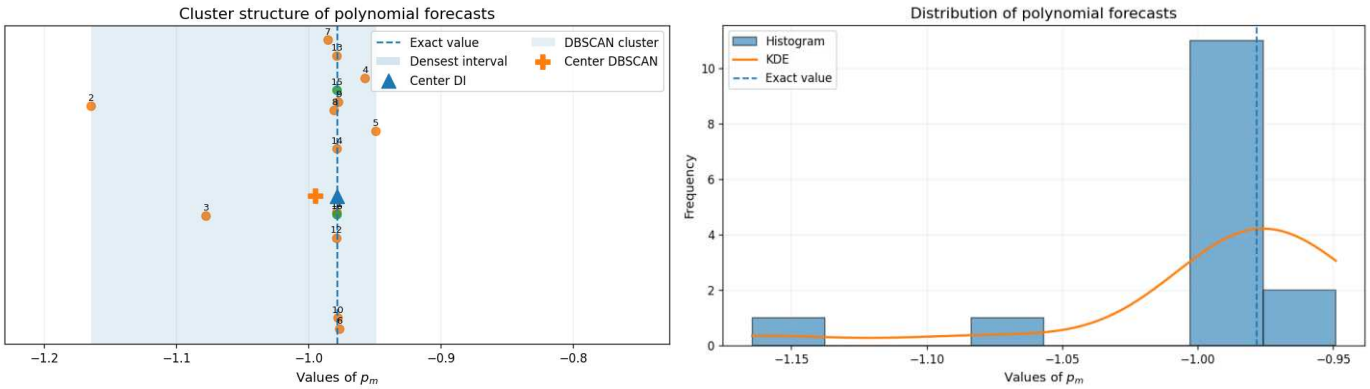


Рис. 5.2. Структура кластера та розподіл PPS для $y = \cos(hx)$.

Результати кластерного аналізу методами DI та DBSCAN наведено в Додатку 5.

Стохастичні дані для проведення експериментів формувалися за допомогою генератора псевдовипадкових чисел як вибірка з нормального (гаусового) розподілу з нульовим математичним сподіванням та заданим стандартним відхиленням σ, яке визначало рівень шуму в ряді. Генерація значень описується співвідношенням $y_i \sim N(0,\sigma^2)$

Для забезпечення відтворюваності результатів використовувався фіксований початковий стан генератора (seed). У кожному експерименті генерувалася послідовність із $N+1$ значень, де перші N елементів використовувалися як історія $\{f_0, f_1, \ldots, f_{N-1}\}$, а останнє значення f_N розглядалося як контрольне (істинне) значення для оцінювання точності прогнозу. Вхідні дані формувалися без додаткових припущень щодо тренду чи сезонності, що дозволило дослідити поведінку послідовності поліноміальних прогнозів p_m в умовах відсутності детермінованої структури та оцінити здатність методу кластеризації виявляти внутрішні закономірності навіть у випадкових даних.

На Рис. 5.3 показано графік стохастичних даних та графік послідовності поліноміальних прогнозів із такими заданими параметрами: $N = 30$ (кількість історичних точок), $Seed = 42$ (параметр відтворюваності псевдовипадкового розподілу), $\sigma = 0{,}5$ (шум), $m_{\min} = 2, m_{\max} = 20$ (мінімальне та максимальне значення кількості точок для послідовності PPS), $\varepsilon = 0{,}5$ (радіус локального сусідства для DBSCAN).

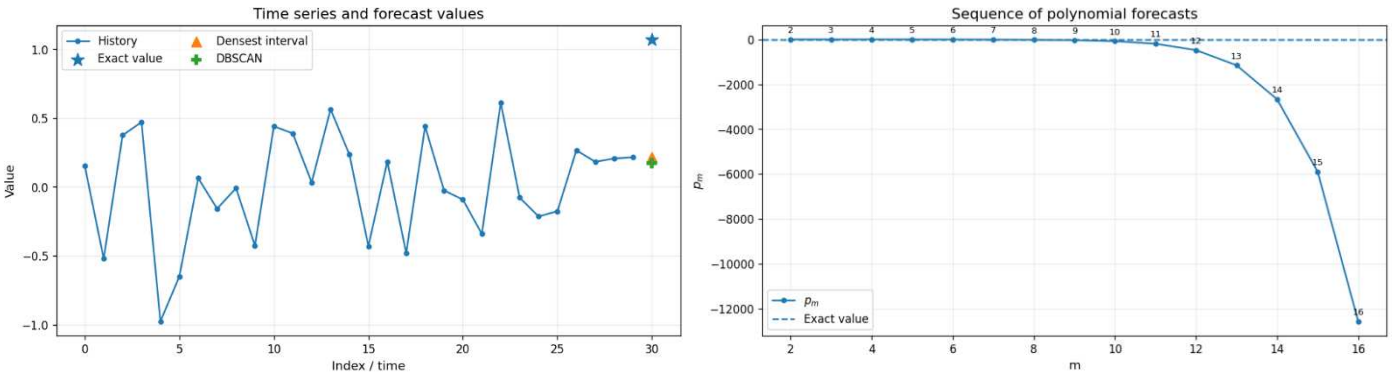


Рис. 5.3. Графік стохастичних даних та його послідовність поліноміальних прогнозів (PPS).

Незважаючи на випадкову природу вхідного ряду, послідовність p_m демонструє ознаки детермінованої поведінки. А саме:

- відсутня хаотична структура;
- зберігаються інтервали монотонності;
- виникають локальні екстремуми;
- спостерігається кластеризація значень.

Розподіл значень PPS на числовій осі, структуру кластера, гістограму та оцінку щільності KDE (Kernel Density Estimation) показано на Рис. 5.4. Для кращої візуалізації розміщення точок, було використано параметр вертикального зміщення (jitter).

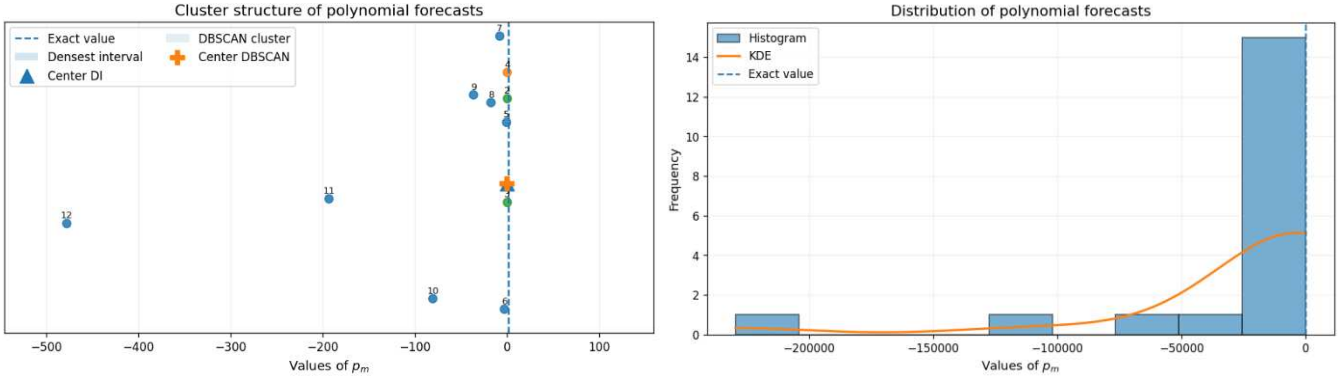


Рис. 5.4. Кластерна структура та розподіл PPS для стохастичних даних.

Результати кластерного аналізу послідовності поліноміальних прогнозів для стохастичних даних методами DI та DBSCAN наведено в додатку 5.

Для експериментів на реальних фінансових даних використано внутрішньоденний набір котирувань акцій компанії Netflix (NFLX), поданий у форматі Excel-файлу з 30-хвилинним кроком дискретизації. Як основний досліджуваний показник обрано параметр Close, тобто ціну закриття відповідного 30-хвилинного інтервалу.

На Рис. 5.5 показано графік біржових котирувань тикера NFLX та графік послідовності поліноміальних прогнозів із такими заданими параметрами: $N = 31$ (кількість історичних точок), $m_{\min} = 2, m_{\max} = 12$ (мінімальне та

максимальне значення кількості точок для послідовності PPS), $\varepsilon = 0{,}5$ (радіус локального сусідства для DBSCAN).

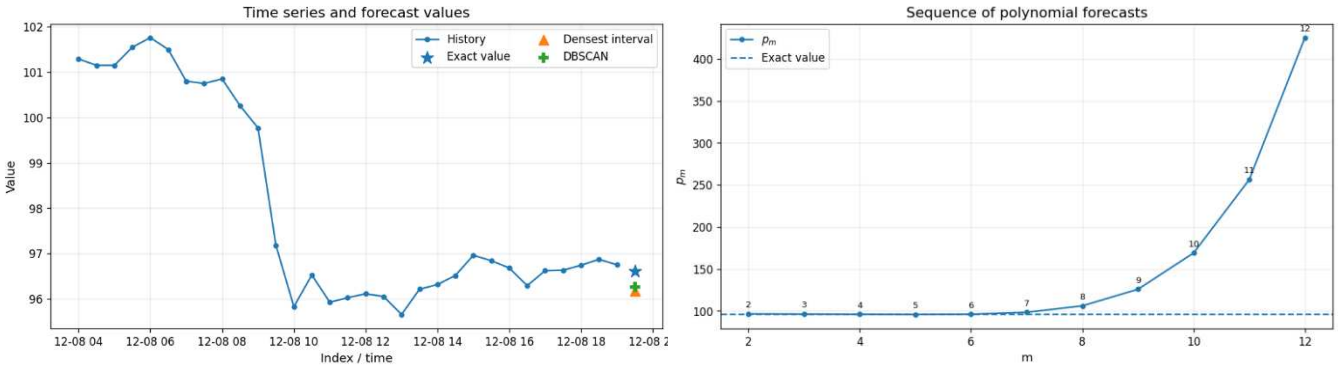


Рис. 5.5 Графік даних про котирування (NFLX, 30 хв) та його послідовність поліноміальних прогнозів (PPS).

Для біржових котирувань послідовність PPS має свої особливості, а саме:

- має комбіновану природу: частково гладку, частково шумову;
- чітко виділяються домінантні кластери;
- DBSCAN стабільно виділяє найбільш репрезентативну групу значень (Рис. 5.6).

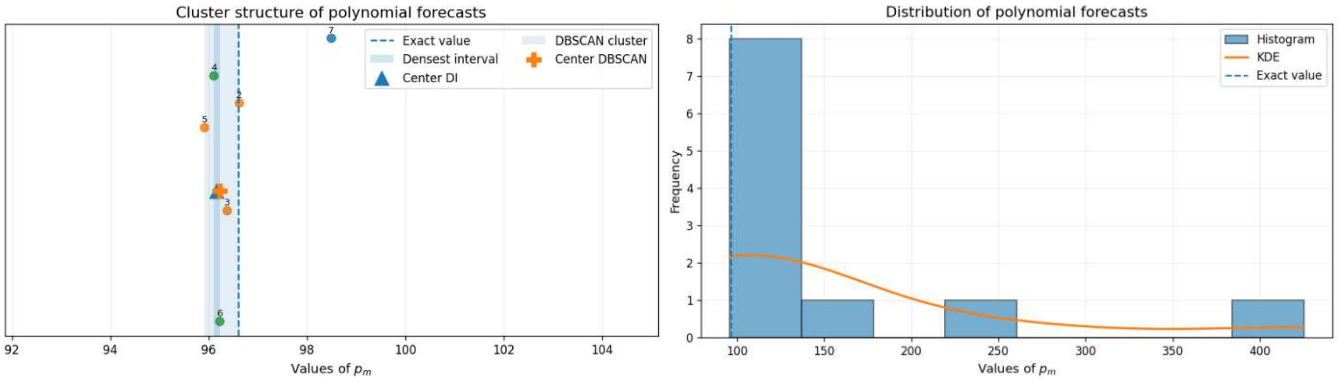


Рис. 5.6. Кластерна структура та розподіл PPS для даних про акції (тікер NFLX, 30 хв)

Результати кластерного аналізу послідовності поліноміальних прогнозів для для даних про акції методами DI та DBSCAN наведено в Додатку 5.

Отримані результати показують, що незалежно від типу вхідних даних послідовність поліноміальних прогнозів p_m має внутрішню структуру, яка проявляється у вигляді локальних екстремумів та щільних областей значень. Це обґрунтовує доцільність застосування методів кластеризації для визначення фінального прогнозу. Найбільш стабільні результати отримано при використанні

алгоритму DBSCAN, який дозволяє виділяти домінантні групи значень навіть у присутності шуму.

5.3. Дослідження ефективності запропонованої нейромережевої архітектури

Для порівняння з запропонованою PPS-Net архітектурою було реалізовано два альтернативні методи. Перший – з адаптивним вибором порядку поліноміальної екстраполяції (Δ-метод). У ньому вибір прогнозу здійснюється на основі внутрішнього критерію узгодженості поліноміальних прогнозів. Другий – метод прогнозування на основі кластеризації поліноміальних прогнозних значень (DBSCAN PPS). У цьому підході до множини значень $\{P_0, P_1, \ldots, P_9\}$ застосовувалася кластеризація DBSCAN. Підсумковий прогноз визначався як центральна характеристика знайденого кластера. Якщо кластер не виявлявся, використовувався резервний прогноз P_0.

Результати прогнозування по значеннях параметра Close за останній день експериментальної вибірки подані на рис. 5.7.

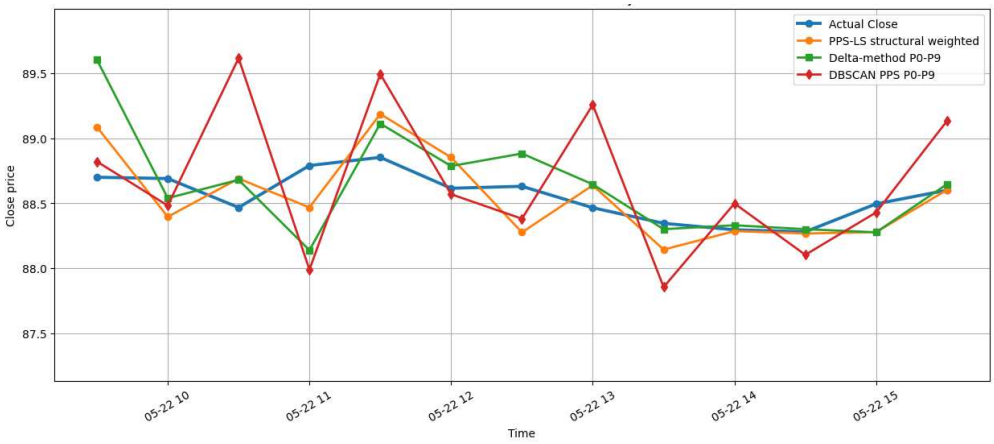


Рис. 5.7. Порівняння результатів прогнозування нейронної мережі PPS-Net із альтернативними методами

Для аналізу поведінки селекторного шару було досліджено розподіл класів l_m у навчальній і тестовій вибірках. Для кожного навчального вікна правильним класом вважався індекс кандидата l_m, який мав найменше відхилення від фактичного наступного значення часового ряду $m^* = \arg\min_m |l_m - x_{11}|$. Найчастіше еталонним класом був $l_1 = P_0$, частка якого становила 38,30%. Класи l_1, l_2, l_3 разом

охоплювали 67,07% навчальних прикладів, а класи $l_1,\ldots,l_6$ – 91,33%. Кандидати $l_7,\ldots,l_{10}$, до складу яких входять поліноміальні прогнози вищих порядків, використовувалися значно рідше (табл.5.2).

Таблиця 5.2 – Розподіл еталонних l_m класів у навчальній вибірці.

Індекс	**Клас**	**Кількість вікон**	**Частка, %**
0	l_1	221	38,30
1	l_2	87	15,08
2	l_3	79	13,69
3	l_4	47	8,15
4	l_5	61	10,57
5	l_6	32	5,55
6	l_7	25	4,33
7	l_8	10	1,73
8	l_9	11	1,91
9	l_{10}	4	0,69

Для 193 тестових вікон селектор PPS-Net переважно обирав класи l_1,l_2,l_3,l_4. Їхня сумарна частка становила 97.92 %. Класи $l_7,\ldots,l_{10}$ не були використані для формування підсумкового прогнозу. Такий розподіл узгоджується з особливостями короткострокового прогнозування внутрішньоденних фінансових даних: на малих часових інтервалах найбільш інформативними переважно є кандидати, сформовані на основі поліноміальних прогнозів невисоких порядків.

Після вибору кандидата застосовувався фільтр локальної допустимості. Якщо вибране значення l_m надмірно відхилялося від базового прогнозу P_0, воно замінювалося на P_0. Отже, розподіл класів на тестовій вибірці характеризує остаточну поведінку алгоритму з урахуванням постобробки прогнозних значень (табл. 5.3).

Таблиця 5.3 – Розподіл класів l_m, використаних для формування підсумкового прогнозу PPS-Net після фільтрації.

Індекс	Клас	Кількість вікон	Відсоток, %
0	l_1	63	32,64
1	l_2	41	21,24
2	l_3	42	21,76
3	l_4	43	22,28
4	l_5	3	1,55
5	l_6	1	0,52

Якість прогнозування оцінювалася за класичними метриками MAE, RMSE, MAPE. Порівняння ефективності результатів прогнозування за допомогою запропонованої нейронної мережі PPS-Net з альтернативними методами прогнозування наведені в таблицях 5.4 та 5.5.

Таблиця 5.4. – Результати на повній тестовій вибірці

Метод	MAE	RMSE	MAPE, %
PPS-Net	0,3643	0,5120	0,4120
Δ-метод	0,4469	0,7402	0,5054
DBSCAN PPS	0,6812	1,3672	0,7688

Таблиця 5.5 – Результати на останньому тестовому дні

Метод	MAE	RMSE	MAPE, %
PPS-Net	0,2133	0,2486	0,2406
Δ-метод	0,2423	0,3462	0,2733
DBSCAN PPS	0,4217	0,5358	0,4762

На рис. 5.8 наведено динаміку абсолютної похибки однокрокового прогнозування для останнього торгового дня тестової вибірки. Для кожного моменту часу абсолютна похибка визначалася як $e_t = |x_t - \hat{x}_t|$, де x_t – фактичне значення ціни закриття, а $\hat{x}_t$ – прогнозне значення відповідного методу.

Графічний аналіз абсолютних похибок підтверджує, що запропонований підхід формує більш рівномірний профіль похибки та не допускає різких прогнозних викидів. Отримані результати показують, що запропонована архітектура нейромережі PPS-Net на основі поліноміального підходу дозволяє підвищити точність однокрокового прогнозування внутрішньоденних фінансових часових рядів. Водночас наведені висновки стосуються дослідженої вибірки NFLX і не повинні автоматично узагальнюватися на всі біржові активи.

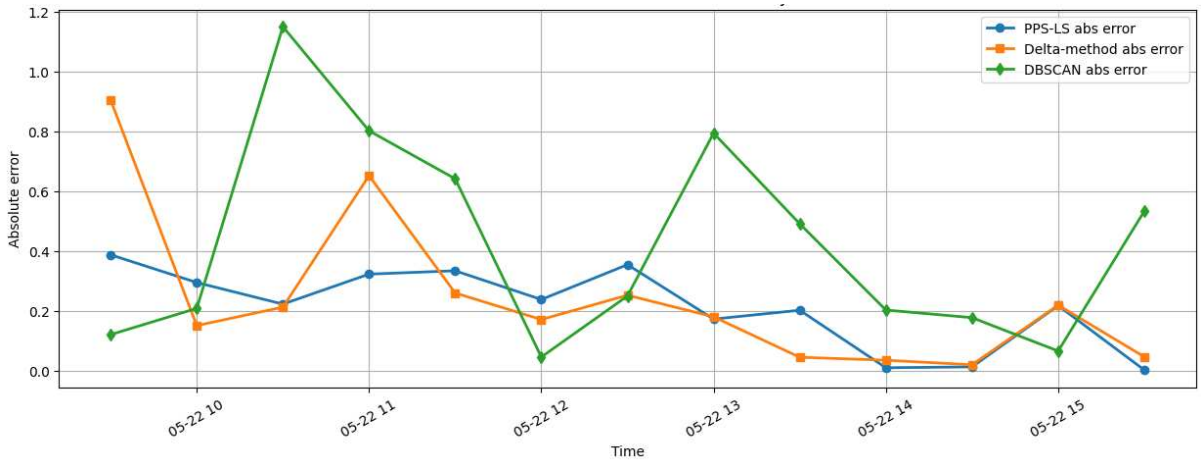


Рис. 5.8. Динаміка абсолютної похибки однокрокового прогнозування ціни Close акцій NFLX для останнього тестового дня

Аналіз розподілу еталонних класів у навчальній вибірці показав, що для прогнозування внутрішньоденних фінансових даних доцільно використовувати поліноміальні прогнози невисоких порядків. Водночас наявність інших класів обґрунтовує необхідність адаптивного вибору параметра m, а не використання одного фіксованого прогнозного правила.

На тестовій вибірці запропонований метод PPS-Net забезпечив найменші значення всіх використаних метрик похибки (табл. 5.4, табл. 5.5). Порівняно з Δ-методом запропонований підхід забезпечив зменшення усередненої вдносної похибки (MAPE) на 18,47%, а з DBSCAN – 46,40%.

На останньому тестовому дні метод PPS-Net зменшив MAPE – на 11,95% порівняно з Δ-методом, відповідно для DBSCAN зменшення MAPE становить 49,46%.

5.4. Особливості програмної реалізації та межі застосування

Використання розроблених у дисертаційній роботі методів доцільно розглядати не лише як застосування окремої прогнозної формули, а як побудову цілісного програмного модуля інтелектуального аналізу часових рядів біржової динаміки. Такий модуль має охоплювати етапи завантаження та попередньої обробки даних, формування локального вікна передісторії, побудови послідовності поліноміальних прогнозних значень PPS, адаптивного вибору інформативної частини PPS, формування остаточного прогнозу, оцінювання

похибки та моніторингу якості прогнозування в режимі послідовного надходження нових біржових значень.

Необхідність саме такого підходу узгоджується із сучасною практикою аналізу часових рядів, де для коректного оцінювання прогнозних моделей рекомендується зберігати часовий порядок спостережень і використовувати процедури rolling forecasting origin або walk-forward validation.

З погляду комп'ютерних наук розроблені методи доцільно реалізовувати як модуль прогнозування, що працює з одновимірним або багатовимірним потоком біржових даних. У роботі базовим варіантом є одновимірний часовий ряд значень ціни закриття $F = \{f_1, f_2, \ldots, f_N\}$, де f_t – значення біржового показника, наприклад Close, у момент часу t, N кількість доступних спостережень. На кожному кроці прогнозування формується локальне вікно з попередніх значень $X_t = (f_{t-L+1}, f_{t-L+2}, \ldots, f_t)$, де L – довжина вхідного вікна, що визначає кількість попередніх значень, які використовуються для побудови прогнозу. На основі локального вікна будується послідовність поліноміальних прогнозних значень $PPS_t = (p_{t,1}, p_{t,2}, \ldots, p_{t,M})$, де M – максимальна глибина побудови поліноміальних прогнозів, а $p_{t,m}$ – прогнозне значення, отримане за поліноміальною екстраполяцією порядку, що відповідає використанню m останніх точок локального вікна.

Якість практичного використання розроблених методів суттєво залежить від коректності підготовки вхідних даних. Для експериментальних і навчальних цілей можуть використовуватися відкриті джерела біржових даних, зокрема бібліотека yfinance, однак її документація прямо зазначає, що інструмент призначений для дослідницького й освітнього використання та не є офіційно афілійованим із Yahoo Finance . Тому для промислового або комерційного використання бажано застосовувати ліцензовані джерела ринкових даних із контрольованою якістю, затримкою та правилами доступу.

Перед застосуванням PPS-методів рекомендується виконати такі дії:

1) перевести часові позначки до єдиного часового поясу;

2) відфільтрувати неторгові інтервали, пропуски та дублікати;
3) привести дані до сталої часової дискретизації;
4) перевірити наявність аномальних викидів;
5) окремо обробляти торгові сесії, міжсесійні розриви та різкі цінові стрибки.

Для приведення даних до сталої частоти можуть використовуватися стандартні засоби роботи з часовими рядами, зокрема pandas.resample, який призначений для перетворення частоти часових даних за умови наявності datetime-подібного індексу. У загальному вигляді процедуру дискретизації можна подати так:

$$F^{(\Delta t)} = \{f_{t_1}, f_{t_2}, \ldots, f_{t_q}\}, \quad t_{i+1} - t_i = \Delta t,$$

де Δt – фіксований часовий крок, наприклад 5 хв, 15 хв або 30 хв.

Для внутрішньоденних даних не рекомендується механічно з'єднувати останнє значення попередньої торгової сесії з першим значенням наступної без урахування нічного розриву. Такий розрив може спотворювати скінченні різниці та призводити до надмірного розходження елементів PPS. Практично доцільним є один із трьох варіантів: прогнозування окремо в межах кожної торгової сесії; введення спеціальної ознаки початку сесії; або використання нормалізації відносно останнього відомого значення.

У розроблених методах доцільно розрізняти параметри, описані в табл. 5.6.:

Таблиця 5.6 – Приклад стартових налаштувань параметрів

Параметр	**Практичне значення**	**Пояснення**
L	$13 \ldots 15$	приблизна довжина одного торгового дня або близького локального фрагмента
M	$8 \ldots 10$	достатньо для формування PPS без надмірного посилення шуму
m	адаптивно	визначається за Δ-критерієм, DBSCAN, аналізом екстремуму або PPS-Net
α	$0,05 \ldots 0,25$	регулює ширину допустимого діапазону після екстремуму PPS
Δt	5, 15, 30 хв	залежить від частоти вхідних біржових даних

Збільшення M дає змогу врахувати поліноми вищого степеня, однак у зашумлених біржових рядах це може призводити до різкого зростання амплітуди окремих елементів PPS. Тому для короткострокового прогнозування в умовах обмеженого обсягу даних доцільно обмежувати M і застосовувати адаптивний вибір інформативної підмножини прогнозів.

У роботі розглянуто кілька варіантів використання PPS. Їх практичне призначення відрізняється залежно від обсягу даних, рівня шуму, потреби в інтерпретованості та вимог до швидкодії.

Найпростішим варіантом практичного використання PPS є формування остаточного прогнозу як середнього значення перших m елементів послідовності. Цей варіант доцільно застосовувати тоді, коли необхідна проста, швидка й інтерпретована процедура прогнозування без навчання параметричної моделі. Його перевагою є низька обчислювальна складність і можливість використання на малих вибірках.

Δ-критерій доцільно використовувати тоді, коли необхідно автоматично вибрати кількість елементів PPS на основі локальної узгодженості прогнозних значень. Цей метод рекомендовано використовувати як базовий адаптивний варіант, оскільки він не потребує навчальної вибірки, має чітку алгоритмічну інтерпретацію та може бути реалізований у режимі потокового оновлення даних.

Методи кластеризації значень PPS доцільно застосовувати тоді, коли в послідовності поліноміальних прогнозів спостерігається групування близьких значень. У цьому випадку остаточний прогноз може визначатися як центр найбільш інформативного кластера. Для малих PPS-послідовностей практично доцільним є використання DBSCAN, оскільки цей алгоритм не потребує наперед задавати кількість кластерів.

Метод аналізу екстремуму PPS доцільно застосовувати в ситуаціях, коли послідовність поліноміальних прогнозних значень має виражений локальний мінімум або максимум, після якого значення залишаються в допустимому діапазоні. Цей метод є доцільним тоді, коли потрібно зберегти інтерпретованість

прогнозу та уникнути використання надто віддалених елементів PPS, які можуть бути спричинені посиленням шуму при переході до поліномів вищого степеня.

Якщо доступна достатня кількість вікон для навчання, розроблені поліноміальні прогнозні значення можуть використовуватися як вхідні або проміжні ознаки нейромережевої архітектури PPS-Net. Такий підхід відповідає сучасній тенденції побудови спеціалізованих нейромережевих архітектур для часових рядів. PPS-Net доцільно використовувати тоді, коли потрібно не лише отримати прогноз, а й навчити модель автоматично визначати відносну інформативність різних елементів PPS. При цьому важливо уникати надмірного ускладнення архітектури, оскільки для малих біржових вибірок надто глибокі моделі можуть демонструвати нестабільне узагальнення.

З погляду програмної інженерії розроблені методи реалізовано у вигляді окремих модулів:

1) модуль завантаження даних;
2) модуль очищення та нормалізації;
3) модуль формування локальних вікон;
4) модуль побудови PPS;
5) модуль адаптивного вибору прогнозу;
6) модуль оцінювання похибки;
7) модуль збереження результатів експерименту;
8) модуль візуалізації прогнозів.

Практична структура програмного комплексу має зображена на рис. 5.9.

Під час переходу від експериментального коду до програмного комплексу важливо враховувати проблему технічного боргу в ML-системах. У дослідженні Д. Скаллі (D. Sculley) та його співавторів обґрунтовано, що системи машинного навчання (ML-системи) характеризуються наявністю специфічних джерел технічного боргу [167]. Автори довели, що такі приховані дефекти зумовлені складними залежностями на рівні даних, нестабільністю конфігураційних параметрів, наявністю системних зворотних зв’язків, а також динамічними змінами у зовнішньому середовищі (concept drift), що суттєво ускладнює

довгострокову підтримку та масштабування подібних комплексів. Для розроблених PPS-методів це означає, що параметри L, M, α, спосіб нормалізації, джерело даних і версія алгоритму мають явно фіксуватися в експериментальних протоколах.

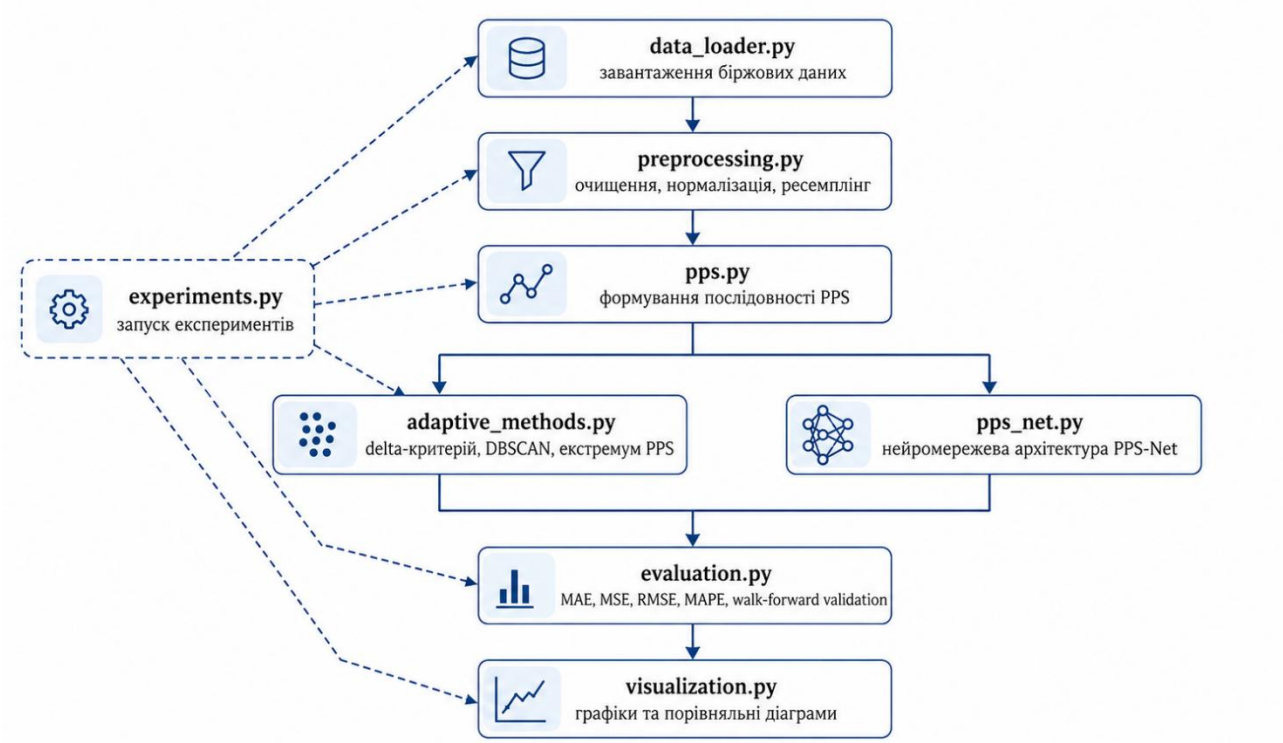


Рис. 5.9. Структура програмного комплексу для експериментального дослідження

Рекомендовано зберігати для кожного експерименту $\Theta = \{L, M, \alpha, \Delta t, D, A, V\}$, де де D – ідентифікатор набору даних, A – використаний алгоритм, V – версія програмної реалізації. Це забезпечує відтворюваність експериментів і дає можливість порівнювати результати різних методів на однакових умовах.

Розроблені методи не слід розглядати як універсальну торгову стратегію. Вони призначені насамперед для задач інтелектуального аналізу, прогнозування та порівняльного дослідження моделей часових рядів. Практичне використання прогнозів у фінансових рішеннях потребує додаткового врахування транзакційних витрат, ліквідності, спреду, ризик-менеджменту та регуляторних обмежень, зокрема:

1) поліноміальні прогнозні значення можуть різко розходитися за наявності шуму;
2) надто велике *M* може погіршувати якість прогнозу;
3) методи чутливі до часової дискретизації даних;
4) результати на одному фінансовому інструменті не гарантують аналогічної якості на інших інструментах;
5) нейромережева PPS-Net потребує достатньої кількості навчальних вікон;
6) для коректного порівняння необхідно використовувати однакову walk-forward схему оцінювання.

З урахуванням цих обмежень PPS-методи найбільш доцільно застосовувати як компонент програмної системи аналізу біржової динаміки, а не як ізольований механізм прийняття фінансових рішень.

Висновки до п’ятого розділу

Обґрунтовано вибір експериментальних даних для перевірки запропонованих методів. Як об’єкт експериментального дослідження використано часові ряди біржової динаміки, що характеризуються шумом, локальною мінливістю, нестаціонарністю, короткочасними трендами та нерівномірністю зміни цінових значень. Показано, що саме такі властивості ускладнюють застосування класичних методів прогнозування та зумовлюють доцільність використання адаптивних алгоритмів, які працюють із локальними фрагментами ряду. Для проведення експериментів сформовано послідовність локальних вікон, у межах яких виконувалося одноетапне прогнозування наступного значення біржового показника.

У результаті проведення обчислювальних експериментів із використанням кластерного аналізу встановлено, що значення PPS можуть утворювати локально згруповані області, які містять найбільш узгоджені прогнозні значення. Це підтверджує доцільність використання кластеризації як засобу адаптивного вибору інформативної підмножини поліноміальних прогнозів. Застосування підходів на основі найщільнішого інтервалу та DBSCAN дало змогу формувати остаточне прогнозне значення не за всією PPS, а лише за тією її частиною, яка

має найбільшу внутрішню узгодженість. Такий підхід зменшує вплив віддалених або нестабільних елементів PPS, що можуть виникати внаслідок посилення шуму при переході до поліномів вищого степеня.

Описано експериментальні дані, використані для перевірки ефективності розроблених методів прогнозування. Як основний набір обрано внутрішньоденні 30-хвилинні котирування акцій Netflix, Inc. з тікером NFLX. Цільовою змінною визначено ціну закриття $Close_t$, що дозволило подати задачу у вигляді одновимірного прогнозування наступного значення біржового часового ряду. Дані було перетворено на послідовність ковзних локальних ретроспективних фрагментів довжини $L = 15$, а для формування PPS використовувалося $M = 10$ останніх значень. Сформовано 770 експериментальних прикладів, з яких 577 використано для навчання, а 193 – для тестування. Така структура даних забезпечує коректну перевірку алгоритмів в умовах прогнозування та обмеженого обсягу доступних спостережень.

Показано, що нейромережева архітектура PPS-Net є доцільною для задач, у яких необхідно автоматизувати вибір найбільш інформативних елементів PPS. Формування остаточного прогнозу як зваженої комбінації поліноміальних прогнозних значень дає змогу уникнути жорсткого попереднього задання параметра *m* та передати функцію адаптивного вибору на рівень навченої моделі. Водночас встановлено, що ефективність PPS-Net залежить від обсягу навчальної вибірки, способу нормалізації даних, вибраної довжини локального вікна та кількості елементів PPS.

Для оцінювання якості розроблених методів використано процедуру послідовного прогнозування з урахуванням часової структури даних. Якість прогнозування оцінювалася за стандартними метриками MAE, MSE, RMSE та MAPE, що дало змогу порівнювати результати різних варіантів запропонованих методів у єдиній експериментальній схемі.

Розглянуто особливості програмної реалізації розроблених методів. Запропоновано модульну структуру програмного комплексу, що включає компоненти завантаження біржових даних, попередньої обробки, формування

PPS, реалізації адаптивних методів, побудови PPS-Net, оцінювання якості прогнозування, запуску експериментів та візуалізації результатів.

Результати експериментальних досліджень підтвердили працездатність і практичну доцільність запропонованих методів інтелектуального аналізу часових рядів біржової динаміки на основі поліноміального підходу. Розроблені алгоритми забезпечують формування послідовності альтернативних поліноміальних прогнозів, адаптивний вибір їх інформативної частини та можливість інтеграції з нейромережевими засобами прогнозування. Це дає підстави розглядати запропонований підхід як перспективний напрям побудови програмних систем прогнозування фінансових часових рядів в умовах неповноти, зашумленості та недетермінованості біржових процесів.

ВИСНОВКИ

1. На основі системного аналізу сучасних теоретичних засад інтелектуального аналізу та прогнозування фінансових часових рядів сформульовано постановку задачі дослідження біржової динаміки в умовах обмеженого обсягу даних, високої волатильності, шуму та недетермінованості біржових процесів. В межах цього доведено доцільність поєднання поліноміальної екстраполяції, аналізу послідовності поліноміальних прогнозних значень (PPS) та методів машинного навчання для розв’язання задач короткострокового прогнозування часових рядів біржової динаміки.

Це дає можливість враховувати локальну структуру ряду, виявляти області прогнозних значень і підвищувати адаптивність прогнозних моделей в умовах обмеженої вибірки. Сформовані теоретичні засади створюють основу для подальшої розробки методів інтелектуального аналізу та прогнозування часових рядів біржової динаміки.

2. Формалізовано методичні підходи до прогнозування часових рядів біржової динаміки на основі поліноміальної екстраполяції. Введено поняття послідовності поліноміальних прогнозних значень (PPS). У результаті дослідження обґрунтовано доцільність переходу від використання одиничного поліноміального прогнозу до комплексного аналізу послідовності поліноміальних прогнозних значень PPS.

Показано, що така послідовність відображає локальні властивості біржового часового ряду, зокрема збіжність, розбіжність, екстремуми, монотонні ділянки та області згущення прогнозних значень. Встановлено, що аналіз структури PPS дає змогу оцінювати чутливість прогнозу до зміни параметра поліноміальної екстраполяції та обґрунтовано вибирати оптимальне значення m, що підвищує адаптивність короткострокового прогнозування біржової динаміки.

3. Удосконалено сучасні методи адаптивного прогнозування часових рядів біржової динаміки, що передбачають побудову послідовності PPS, аналіз її структурних властивостей, вибір оптимального параметра m або інформативної підмножини прогнозів та формування остаточного прогнозного значення.

Запропоновані методи ґрунтуються на використанні внутрішніх характеристик PPS без залучення додаткових зовнішніх факторів, що підвищує їхню придатність для застосування в умовах обмеженості даних.

Показано, що використання усереднення відібраної підмножини поліноміальних прогнозів, аналіз локальних екстремумів і виявлення кластерних структур у PPS дають змогу підвищити точність прогнозування, особливо для малих фрагментів біржових часових рядів. Визначено межі застосовності запропонованих методів, зокрема випадки швидкої розбіжності PPS, відсутності виражених областей згущення або нестабільної локальної структури прогнозних значень.

Це дає можливість адаптивно визначати параметри поліноміальної екстраполяції відповідно до локальної поведінки часового ряду біржового процесу, зменшувати вплив випадкових коливань на остаточний прогноз і підвищувати обґрунтованість вибору прогнозної моделі в умовах невизначеності та високої волатильності фінансових даних.

4. Розроблено архітектуру нейронної мережі для прогнозування часових рядів біржової динаміки, що ґрунтується на поєднанні початкових значень локального фрагмента ряду та послідовності поліноміальних прогнозних значень PPS. Особливістю запропонованої архітектури є введення спеціального блоку формування PPS, блоку нейромережевого аналізу її структури та блоку адаптивного зважування поліноміальних прогнозів для різних значень параметра m.

Встановлено, що блок адаптивного зважування поліноміальних прогнозів дозволяє формувати остаточне прогнозне значення не шляхом вибору одного фіксованого порядку полінома, а як зважену комбінацію елементів PPS, що забезпечує поєднання аналітичних переваг поліноміального підходу з навчальною здатністю нейронної мережі та підвищує гнучкість моделі в умовах локальної нестійкості біржової динаміки.

Це дає можливість здійснювати інтерпретоване прогнозування часових рядів біржової динаміки, оскільки розподіл ваг за різними значеннями m

дозволяє визначити, які саме поліноміальні прогнози мали найбільший вплив на формування результату. Такий підхід підвищує обґрунтованість прогнозу, сприяє адаптивному врахуванню локальної структури ряду та створює основу для подальшого розвитку нейромережевих моделей, орієнтованих на використання поліноміального підходу.

5. Сформовано науково-прикладні засади оцінювання запропонованих методів прогнозування часових рядів біржової динаміки, що ґрунтуються на поєднанні walk-forward перевірки, one-step-ahead прогнозування та аналізу точності за метриками MAE, MSE, RMSE і MAPE.

Проведено експериментальну перевірку адаптивної поліноміальної екстраполяції, усереднення послідовності поліноміальних прогнозних значень PPS, екстремального критерію, кластерного підходу, DBSCAN-аналізу та запропонованої нейромережевої архітектури на реальних часових рядах біржової динаміки.

Встановлено, що оцінювання запропонованих методів доцільно здійснювати не лише за числовими показниками похибки прогнозу, а й з урахуванням структури PPS, наявності кластерів прогнозних значень та поведінки ваг у нейромережевій архітектурі. Такий підхід дозволяє глибше інтерпретувати результати прогнозування, визначати внесок окремих поліноміальних прогнозів у формування остаточного результату та виявляти випадки зниження надійності моделі.

Запропоновані засади дають змогу оцінювати ефективність прогнозних моделей у режимі послідовного надходження даних, що відповідає реальним умовам прогнозування біржових процесів. Це дає можливість визначати межі застосовності запропонованих методів для різних класів задач прогнозування біржової динаміки, зокрема в умовах короткого горизонту прогнозування, обмеженого обсягу вхідних даних, локальної структурованості PPS або, навпаки, її швидкої розбіжності.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Cont R. Empirical properties of asset returns: stylized facts and statistical issues. Quantitative Finance. 2001. Vol. 1, No. 2. P. 223–236. DOI: https://doi.org/10.1080/713665670.

2. Ratliff-Crain E., Van Oort C. M., Koehler M. T. K., Tivnan B. F. Revisiting Cont's stylized facts for modern stock markets. Quantitative Finance. 2025. Vol. 25, No. 9. P. 1343–1373. DOI: 10.1080/14697688.2025.2530637.

3. Mandelbrot B. B. The variation of certain speculative prices. The Journal of Business. 1963. Vol. 36, No. 4. P. 394–419. DOI: 10.1086/294632.

4. Fama E. F. The behavior of stock-market prices. The Journal of Business. 1965. Vol. 38, No. 1. P. 34–105. DOI: 10.1086/294743.

5. Fama E. F. Efficient capital markets: a review of theory and empirical work. The Journal of Finance. 1970. Vol. 25, No. 2. P. 383–417. DOI: 10.1111/j.1540-6261.1970.tb00518.x.

6. Engle R. F. Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. Econometrica. 1982. Vol. 50, No. 4. P. 987–1007. DOI: 10.2307/1912773.

7. Bollerslev T. Generalized autoregressive conditional heteroskedasticity. Journal of Econometrics. 1986. Vol. 31, No. 3. P. 307–327. DOI: 10.1016/0304-4076(86)90063-1.

8. Derbentsev V., Matviychuk A., Datsenko N., Bezkorovainyi V., Azaryan A. Machine learning approaches for financial time series forecasting. Proceedings of the 3rd International Workshop on Computer Modeling and Intelligent Systems. 2020. P. 434–450.

9. Derbentsev V., Matviychuk A., Soloviev V. N. Forecasting of Cryptocurrency Prices Using Machine Learning. In: Advanced Studies of Financial Technologies and Cryptocurrency Markets. Singapore : Springer, 2020. P. 211–231.

10. Poon S.-H., Granger C. W. J. Forecasting volatility in financial markets: a review. Journal of Economic Literature. 2003. Vol. 41, No. 2. P. 478–539. DOI: 10.1257/002205103765762743.

11. Schmitt T. A., Chetalova D., Schäfer R., Guhr T. Non-stationarity in financial time series: generic features and tail behavior. EPL (Europhysics Letters). 2013. Vol. 103, No. 5. Art. 58003. DOI: 10.1209/0295-5075/103/58003.

12. Gama J., Žliobaitė I., Bifet A., Pechenizkiy M., Bouchachia A. A survey on concept drift adaptation. ACM Computing Surveys. 2014. Vol. 46, No. 4. Art. 44. P. 1–37. DOI: 10.1145/2523813.

13. Calvet L. E., Fisher A. J. Multifractality in asset returns: theory and evidence. The Review of Economics and Statistics. 2002. Vol. 84, No. 3. P. 381–406. DOI: 10.1162/003465302320259420.

14. Di Matteo T., Aste T., Dacorogna M. M. Long-term memories of developed and emerging markets: using the scaling analysis to characterize their stage of development. Journal of Banking & Finance. 2005. Vol. 29, No. 4. P. 827–851. DOI: 10.1016/j.jbankfin.2004.08.004.

15. Gatheral J., Jaisson T., Rosenbaum M. Volatility is rough. Quantitative Finance. 2018. Vol. 18, No. 6. P. 933–949. DOI: 10.1080/14697688.2017.1393551.

16. Fukasawa M., Takabatake T., Westphal R. Is volatility rough? arXiv preprint. 2019. arXiv:1905.04852. DOI: 10.48550/arXiv.1905.04852.

17. Madhavan A. Market microstructure: a survey. Journal of Financial Markets. 2000. Vol. 3, No. 3. P. 205–258. DOI: 10.1016/S1386-4181(00)00007-0.

18. Wood R. A., McInish T. H., Ord J. K. An investigation of transactions data for NYSE stocks. The Journal of Finance. 1985. Vol. 40, No. 3. P. 723–739. DOI: 10.1111/j.1540-6261.1985.tb04996.x.

19. Andersen T. G., Bollerslev T. Intraday periodicity and volatility persistence in financial markets. Journal of Empirical Finance. 1997. Vol. 4, No. 2–3. P. 115–158. DOI: 10.1016/S0927-5398(97)00004-2.

20. Jain P. C., Joh G.-H. The dependence between hourly prices and trading volume. Journal of Financial and Quantitative Analysis. 1988. Vol. 23, No. 3. P. 269–283. DOI: 10.2307/2331067.

21. Barndorff-Nielsen O. E., Shephard N. Econometric analysis of realized volatility and its use in estimating stochastic volatility models. Journal of the Royal

Statistical Society: Series B (Statistical Methodology). 2002. Vol. 64, No. 2. P. 253–280. DOI: 10.1111/1467-9868.00336.

22. Hansen P. R., Lunde A. Realized variance and market microstructure noise. Journal of Business & Economic Statistics. 2006. Vol. 24, No. 2. P. 127–161. DOI: 10.1198/073500106000000071.

23. Aït-Sahalia Y., Yu J. High frequency market microstructure noise estimates and liquidity measures. The Annals of Applied Statistics. 2009. Vol. 3, No. 1. P. 422–457. DOI: 10.1214/08-AOAS200.

24. Bontempi G., Ben Taieb S., Le Borgne Y.-A. Machine learning strategies for time series forecasting. Business Intelligence. Lecture Notes in Business Information Processing. Berlin ; Heidelberg : Springer, 2013. Vol. 138. P. 62–77. DOI: 10.1007/978-3-642-36318-4_3.

25. Cerqueira V., Torgo L., Mozetič I. Evaluating time series forecasting models: an empirical study on performance estimation methods. Machine Learning. 2020. Vol. 109. P. 1997–2028. DOI: 10.1007/s10994-020-05910-7.

26. Masini R. P., Medeiros M. C., Mendes E. F. Machine learning advances for time series forecasting. Journal of Economic Surveys. 2023. Vol. 37, No. 1. P. 76–111. DOI: 10.1111/joes.12429.

27. Hyndman R. J., Athanasopoulos G. Forecasting: Principles and Practice [Електронний ресурс]. 3rd ed. Melbourne : OTexts, 2021. URL: https://otexts.com/fpp3/ (дата звернення: 08.10.2025).

28. Box G. E. P., Jenkins G. M., Reinsel G. C., Ljung G. M. Time Series Analysis: Forecasting and Control. 5th ed. Hoboken : John Wiley & Sons, 2015. 712 p.

29. Bates J. M., Granger C. W. J. The Combination of Forecasts. Operational Research Quarterly. 1969. Vol. 20, № 4. P. 451–468.

30. Timmermann A. Forecast Combinations. In: Handbook of Economic Forecasting. Amsterdam : Elsevier, 2006. Vol. 1. P. 135–196.

31. Brezinski C., Redivo-Zaglia M. Extrapolation and Rational Approximation: The Works of the Main Contributors. Cham : Springer, 2020. DOI: 10.1007/978-3-030-58418-4.

32. Ivakhnenko A. G., Lapa V. G. Cybernetics and Forecasting Techniques. New York : American Elsevier Publishing Company, 1967. 168 p.

33. Ivakhnenko A. G. Polynomial Theory of Complex Systems. IEEE Transactions on Systems, Man, and Cybernetics. 1971. Vol. SMC-1, № 4. P. 364–378.

34. Derbentsev V., Matviychuk A., Datsenko N., Bezkorovainyi V., Azaryan A. Machine Learning Approaches for Financial Time Series Forecasting. CEUR Workshop Proceedings. 2020. Vol. 2713. P. 434–450.

35. Zaychenko Y., Zaychenko H. Fuzzy GMDH and Its Application to Forecasting Financial Processes. System Research and Information Technologies. 2019. № 1. DOI: 10.20535/SRIT.2308-8893.2019.1.07.

36. Tsay R. S. Analysis of Financial Time Series. 3rd ed. Hoboken : John Wiley & Sons, 2010. 720 p.

37. Winters P. R. Forecasting sales by exponentially weighted moving averages. Management Science. 1960. Vol. 6, No. 3. P. 324–342. DOI: https://doi.org/10.1287/mnsc.6.3.324.

38. Dickey D. A., Fuller W. A. Distribution of the estimators for autoregressive time series with a unit root. Journal of the American Statistical Association. 1979. Vol. 74, No. 366. P. 427–431. DOI: https://doi.org/10.1080/01621459.1979.10482531.

39. Kwiatkowski D., Phillips P. C. B., Schmidt P., Shin Y. Testing the null hypothesis of stationarity against the alternative of a unit root: how sure are we that economic time series have a unit root? Journal of Econometrics. 1992. Vol. 54, No. 1–3. P. 159–178. DOI: https://doi.org/10.1016/0304-4076(92)90104-Y.

40. Akaike H. A new look at the statistical model identification. IEEE Transactions on Automatic Control. 1974. Vol. 19, No. 6. P. 716–723. DOI: https://doi.org/10.1109/TAC.1974.1100705.

41. Schwarz G. Estimating the dimension of a model. The Annals of Statistics. 1978. Vol. 6, No. 2. P. 461–464. DOI: https://doi.org/10.1214/aos/1176344136.

42. Hyndman R. J., Khandakar Y. Automatic time series forecasting: the forecast package for R. Journal of Statistical Software. 2008. Vol. 27, No. 3. P. 1–22. DOI: https://doi.org/10.18637/jss.v027.i03.

43. Engle R. F. Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. Econometrica. 1982. Vol. 50, No. 4. P. 987–1007. DOI: https://doi.org/10.2307/1912773.

44. Bollerslev T. Generalized autoregressive conditional heteroskedasticity. Journal of Econometrics. 1986. Vol. 31, No. 3. P. 307–327. DOI: https://doi.org/10.1016/0304-4076(86)90063-1.

45. Nelson D. B. Conditional heteroskedasticity in asset returns: a new approach. Econometrica. 1991. Vol. 59, No. 2. P. 347–370. DOI: https://doi.org/10.2307/2938260.

46. Glosten L. R., Jagannathan R., Runkle D. E. On the relation between the expected value and the volatility of the nominal excess return on stocks. The Journal of Finance. 1993. Vol. 48, No. 5. P. 1779–1801. DOI: https://doi.org/10.1111/j.1540-6261.1993.tb05128.x.

47. Sims C. A. Macroeconomics and reality. Econometrica. 1980. Vol. 48, No. 1. P. 1–48. DOI: https://doi.org/10.2307/1912017.

48. Engle R. F., Granger C. W. J. Co-integration and error correction: representation, estimation, and testing. Econometrica. 1987. Vol. 55, No. 2. P. 251–276. DOI: https://doi.org/10.2307/1913236.

49. Kalman R. E. A new approach to linear filtering and prediction problems. Journal of Basic Engineering. 1960. Vol. 82, No. 1. P. 35–45. DOI: https://doi.org/10.1115/1.3662552.

50. Durbin J., Koopman S. J. Time Series Analysis by State Space Methods. 2nd ed. Oxford : Oxford University Press, 2012. 368 p.

51. Hamilton J. D. A new approach to the economic analysis of nonstationary time series and the business cycle. Econometrica. 1989. Vol. 57, No. 2. P. 357–384. DOI: https://doi.org/10.2307/1912559

52. Hastie T., Tibshirani R., Friedman J. The elements of statistical learning: Data mining, inference, and prediction. 2nd ed. New York : Springer, 2009. 745 p. DOI: 10.1007/978-0-387-84858-7.

53. Bishop C. M. Pattern recognition and machine learning. New York : Springer, 2006. 738 p. ISBN 978-0-387-31073-2.

54. Hoerl A. E., Kennard R. W. Ridge regression: Biased estimation for nonorthogonal problems. Technometrics. 1970. Vol. 12, No. 1. P. 55–67. DOI: 10.1080/00401706.1970.10488634.

55. Tibshirani R. Regression shrinkage and selection via the Lasso. Journal of the Royal Statistical Society. Series B (Methodological). 1996. Vol. 58, No. 1. P. 267–288. DOI: 10.1111/j.2517-6161.1996.tb02080.x.

56. Cover T. M., Hart P. E. Nearest neighbor pattern classification. IEEE Transactions on Information Theory. 1967. Vol. 13, No. 1. P. 21–27. DOI: 10.1109/TIT.1967.1053964.

57. Cortes C., Vapnik V. Support-vector networks. Machine Learning. 1995. Vol. 20. P. 273–297. DOI: 10.1007/BF00994018.

58. Kim K.-J. Financial time series forecasting using support vector machines. Neurocomputing. 2003. Vol. 55, No. 1–2. P. 307–319. DOI: 10.1016/S0925-2312(03)00372-2.

59. Breiman L., Friedman J. H., Olshen R. A., Stone C. J. Classification and regression trees. Belmont, CA : Wadsworth International Group, 1984.

60. Breiman L. Random forests. Machine Learning. 2001. Vol. 45, No. 1. P. 5–32. DOI: 10.1023/A:1010933404324.

61. Freund Y., Schapire R. E. A decision-theoretic generalization of on-line learning and an application to boosting. Journal of Computer and System Sciences. 1997. Vol. 55, No. 1. P. 119–139. DOI: 10.1006/jcss.1997.1504.

62. Chen T., Guestrin C. XGBoost: A scalable tree boosting system. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. 2016. P. 785–794. DOI: 10.1145/2939672.2939785.

63. Atsalakis G. S., Valavanis K. P. Surveying stock market forecasting techniques — Part II: Soft computing methods. Expert Systems with Applications. 2009. Vol. 36, No. 3. P. 5932–5941. DOI: 10.1016/j.eswa.2008.07.006.

64. Patel J., Shah S., Thakkar P., Kotecha K. Predicting stock and stock price index movement using trend deterministic data preparation and machine learning techniques. Expert Systems with Applications. 2015. Vol. 42, No. 1. P. 259–268. DOI: 10.1016/j.eswa.2014.07.040.

65. Ballings M., Van den Poel D., Hespeels N., Gryp R. Evaluating multiple classifiers for stock price direction prediction. Expert Systems with Applications. 2015. Vol. 42, No. 20. P. 7046–7056. DOI: 10.1016/j.eswa.2015.05.013.

66. Krauß C., Do X. A., Huck N. Deep neural networks, gradient-boosted trees, random forests: Statistical arbitrage on the S&P 500. European Journal of Operational Research. 2017. Vol. 259, No. 2. P. 689–702. DOI: 10.1016/j.ejor.2016.10.031.

67. Henrique B. M., Sobreiro V. A., Kimura H. Literature review: Machine learning techniques applied to financial market prediction. Expert Systems with Applications. 2019. Vol. 124. P. 226–251. DOI: 10.1016/j.eswa.2019.01.012.

68. Gu S., Kelly B., Xiu D. Empirical asset pricing via machine learning. The Review of Financial Studies. 2020. Vol. 33, No. 5. P. 2223–2273. DOI: 10.1093/rfs/hhaa009.

69. Bergmeir C., Hyndman R. J., Koo B. A note on the validity of cross-validation for evaluating autoregressive time series prediction. Computational Statistics & Data Analysis. 2018. Vol. 120. P. 70–83. DOI: 10.1016/j.csda.2017.11.003.

70. Cerqueira V., Torgo L., Mozetič I. Evaluating time series forecasting models: An empirical study on performance estimation methods. Machine Learning. 2020. Vol. 109. P. 1997–2028. DOI: 10.1007/s10994-020-05910-7.

71. Rouf N., Malik M. B., Arif T., Sharma S., Singh S., Aich S., Kim H.-C. Stock market prediction using machine learning techniques: A decade survey on methodologies, recent developments, and future directions. Electronics. 2021. Vol. 10, No. 21. Article 2717. DOI: 10.3390/electronics10212717.

72. Mintarya L. N., Halim J. N. M., Angie C., Achmad S., Kurniawan A. Machine learning approaches in stock market prediction: A systematic literature review. Procedia Computer Science. 2023. Vol. 216. P. 96–102. DOI: 10.1016/j.procs.2022.12.115.

73. Vuong P. H., Phu L. H., Nguyen T. H. V., Duy L. N., Bao P. T., Trinh T. D. A bibliometric literature review of stock price forecasting: From statistical model to deep learning approach. Science Progress. 2024. Vol. 107, No. 1. P. 1–31. DOI: 10.1177/00368504241236557.

74. Bustos O., Pomares-Quimbaya A., Stellian R. Machine learning, stock market forecasting, and market efficiency: A comparative study. International Journal of Data Science and Analytics. 2025. Vol. 20. P. 6815–6839. DOI: 10.1007/s41060-025-00854-4.

75. Sezer O. B., Gudelek M. U., Ozbayoglu A. M. Financial time series forecasting with deep learning: A systematic literature review: 2005–2019. Applied Soft Computing. 2020. Vol. 90. Article 106181. DOI: 10.1016/j.asoc.2020.106181.

76. Zhang G., Patuwo B. E., Hu M. Y. Forecasting with artificial neural networks: The state of the art. International Journal of Forecasting. 1998. Vol. 14, No. 1. P. 35–62. DOI: 10.1016/S0169-2070(97)00044-7.

77. Bengio Y., Simard P., Frasconi P. Learning long-term dependencies with gradient descent is difficult. IEEE Transactions on Neural Networks. 1994. Vol. 5, No. 2. P. 157–166. DOI: 10.1109/72.279181.

78. Hochreiter S., Schmidhuber J. Long short-term memory. Neural Computation. 1997. Vol. 9, No. 8. P. 1735–1780. DOI: 10.1162/neco.1997.9.8.1735.

79. Cho K., van Merriënboer B., Gulcehre C., Bahdanau D., Bougares F., Schwenk H., Bengio Y. Learning phrase representations using RNN Encoder–Decoder for statistical machine translation. Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing. Doha, 2014. P. 1724–1734. DOI: 10.3115/v1/D14-1179.

80. Bai S., Kolter J. Z., Koltun V. An empirical evaluation of generic convolutional and recurrent networks for sequence modeling. arXiv preprint. 2018. arXiv:1803.01271. DOI: 10.48550/arXiv.1803.01271.

81. Vaswani A., Shazeer N., Parmar N., Uszkoreit J., Jones L., Gomez A. N., Kaiser Ł., Polosukhin I. Attention is all you need. Advances in Neural Information Processing Systems. 2017. Vol. 30. P. 5998–6008.

82. Lim B., Zohren S. Time-series forecasting with deep learning: A survey. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences. 2021. Vol. 379, No. 2194. Article 20200209. DOI: 10.1098/rsta.2020.0209.

83. Sezer O. B., Gudelek M. U., Ozbayoglu A. M. Financial time series forecasting with deep learning: A systematic literature review: 2005–2019. Applied Soft Computing. 2020. Vol. 90. Article 106181. DOI: 10.1016/j.asoc.2020.106181.

84. Hewamalage H., Bergmeir C., Bandara K. Recurrent neural networks for time series forecasting: Current status and future directions. International Journal of Forecasting. 2021. Vol. 37, No. 1. P. 388–427. DOI: 10.1016/j.ijforecast.2020.06.008.

85. Fischer T., Krauss C. Deep learning with long short-term memory networks for financial market predictions. European Journal of Operational Research. 2018. Vol. 270, No. 2. P. 654–669. DOI: 10.1016/j.ejor.2017.11.054.

86. Nelson D. M. Q., Pereira A. C. M., de Oliveira R. A. Stock market's price movement prediction with LSTM neural networks. 2017 International Joint Conference on Neural Networks. Anchorage, 2017. P. 1419–1426. DOI: 10.1109/IJCNN.2017.7966019.

87. Salinas D., Flunkert V., Gasthaus J., Januschowski T. DeepAR: Probabilistic forecasting with autoregressive recurrent networks. International Journal of Forecasting. 2020. Vol. 36, No. 3. P. 1181–1191. DOI: 10.1016/j.ijforecast.2019.07.001.

88. Lim B., Arık S. Ö., Loeff N., Pfister T. Temporal fusion transformers for interpretable multi-horizon time series forecasting. International Journal of

Forecasting. 2021. Vol. 37, No. 4. P. 1748–1764. DOI: 10.1016/j.ijforecast.2021.03.012.

89. Oreshkin B. N., Carpov D., Chapados N., Bengio Y. N-BEATS: Neural basis expansion analysis for interpretable time series forecasting. International Conference on Learning Representations. 2020.

90. Bloomberg Terminal. Bloomberg Professional Services. URL: https://professional.bloomberg.com/products/bloomberg-terminal/ (дата звернення 20.05.2026).

91. LSEG Workspace. London Stock Exchange Group. URL: https://www.lseg.com/en/data-analytics/products/workspace (дата звернення 20.05.2026).

92. FactSet: Financial Data, Market Analytics & AI Solutions. FactSet Research Systems Inc. URL: https://www.factset.com/ (дата звернення 20.05.2026).

93. TradingView Subscriptions: Pricing and Features. TradingView. URL: https://www.tradingview.com/pricing/ (дата звернення 20.05.2026).

94. MetaTrader 5 Trading Platform for Forex, Stocks, Futures. MetaQuotes Ltd. URL: https://www.metatrader5.com/en (дата звернення 20.05.2026).

95. NinjaTrader Trading Platform. NinjaTrader. URL: https://ninjatrader.com/trading-platform/ (дата звернення 20.05.2026).

96. AmiBroker: Technical Analysis Software. AmiBroker. URL: https://www.amibroker.com/ (дата звернення 20.05.2026).

97. QuantConnect LEAN Engine. QuantConnect Documentation. URL: https://www.quantconnect.com/docs/v2/lean-engine (дата звернення 21.05.2026).

98. Backtrader: Python Backtesting and Trading Framework. Backtrader. URL: https://www.backtrader.com/ (дата звернення 21.05.2026).

99. VectorBT: Python Package for Quantitative Analysis. VectorBT Documentation. URL: https://vectorbt.dev/ (дата звернення 21.05.2026).

100. PyTorch: Open Source Machine Learning Framework. PyTorch Foundation. URL: https://pytorch.org/ (дата звернення 21.05.2026).

101. MATLAB Pricing and Licensing. MathWorks. URL: https://www.mathworks.com/pricing-licensing.html (дата звернення 21.05.2026).

102. Time-Series. ClickHouse Documentation. URL: https://clickhouse.com/docs/use-cases/time-series (дата звернення 21.05.2026).

103. Challu C., Olivares K. G., Oreshkin B. N., Garza F., Mergenthaler-Canseco M., Dubrawski A. N-HiTS: Neural hierarchical interpolation for time series forecasting. Proceedings of the AAAI Conference on Artificial Intelligence. 2023. Vol. 37, No. 6. P. 6989–6997. DOI: 10.1609/aaai.v37i6.25854.

104. Zhou H., Zhang S., Peng J., Zhang S., Li J., Xiong H., Zhang W. Informer: Beyond efficient Transformer for long sequence time-series forecasting. Proceedings of the AAAI Conference on Artificial Intelligence. 2021. Vol. 35, No. 12. P. 11106–11115. DOI: 10.1609/aaai.v35i12.17325.

105. Zeng A., Chen M., Zhang L., Xu Q. Are Transformers effective for time series forecasting? Proceedings of the AAAI Conference on Artificial Intelligence. 2023. Vol. 37, No. 9. P. 11121–11128. DOI: 10.1609/aaai.v37i9.26317.

106. Nie Y., Nguyen N. H., Sinthong P., Kalagnanam J. A time series is worth 64 words: Long-term forecasting with Transformers. International Conference on Learning Representations. 2023.

107. Wu H., Hu T., Liu Y., Zhou H., Wang J., Long M. TimesNet: Temporal 2D-variation modeling for general time series analysis. International Conference on Learning Representations. 2023.

108. Liu Y., Hu T., Zhang H., Wu H., Wang S., Ma L., Long M. iTransformer: Inverted Transformers are effective for time series forecasting. International Conference on Learning Representations. 2024.

109. Ansari A. F., Stella L., Turkmen C., Zhang X., Mercado P., Shen H., Shchur O., Rangapuram S. S., Arango S. P., Kapoor S., Zschiegner J., Maddix D. C. et al. Chronos: Learning the language of time series. Transactions on Machine Learning Research. 2024. DOI: 10.48550/arXiv.2403.07815.

110. Hyndman R. J., Koehler A. B. Another look at measures of forecast accuracy. International Journal of Forecasting. 2006. Vol. 22, No. 4. P. 679–688. DOI: 10.1016/j.ijforecast.2006.03.001.

111. Makridakis S., Spiliotis E., Hollyman R., Petropoulos F., Swanson N., Gaba A. The M6 forecasting competition: bridging the gap between forecasting and investment decisions. International Journal of Forecasting. 2025. Vol. 41, No. 4. P. 1315–1354. DOI: 10.1016/j.ijforecast.2024.11.002.

112. Makridakis S., Spiliotis E., Assimakopoulos V. The M4 Competition: Results, findings, conclusion and way forward. International Journal of Forecasting. 2018. Vol. 34, No. 4. P. 802–808.

113. Токарєва К. А. Гібридні та ансамблеві методи та моделі машинного навчання прогнозування фінансових часових рядів : дис. д-ра філософії : 113 Прикладна математика. Чернівці : Чернівецький національний університет імені Юрія Федьковича, 2024.

114. Бідюк П. І. Системний підхід до прогнозування на основі моделей часових рядів. Системні дослідження та інформаційні технології. 2003. № 3. С. 88–110.

115. Бідюк П. І., Федоров А. В. Ймовірнісне прогнозування процесів ціноутворення на фондових ринках. Системні дослідження та інформаційні технології. 2009. № 1. С. 65–73.

116. Шубенкова І. А., Петрова С. К., Бідюк П. І. Системний підхід до моделювання та прогнозування на основі регресійних моделей і фільтра Калмана. Системні дослідження та інформаційні технології. 2017. № 2. С. 52–61. DOI: 10.20535/SRIT.2308-8893.2017.2.05.

117. Прогнозування фінансових процесів: практикум : навч. посіб. [Електронний ресурс] / уклад.: Н. В. Кузнєцова, П. І. Бідюк ; КПІ ім. Ігоря Сікорського. Київ : КПІ ім. Ігоря Сікорського, 2023. 106 с. URL: https://ela.kpi.ua/handle/123456789/63940 (дата звернення: 10.06.2026).

118. Турбал Ю. В., Кубай О. В. Аналіз ступеня стохастичності процесів на основі послідовності поліноміальних прогнозів (PPS). Комп’ютерно-інтегровані

технології: освіта, наука, виробництво. 2026. № 63. С. 277–286. DOI: 10.36910/6775-2524-0560-2026-63-31.

119. Fan J., Gijbels I. Local Polynomial Modelling and Its Applications. London : Chapman & Hall, 1996. 360 p.

120. Savitzky A., Golay M. J. E. Smoothing and Differentiation of Data by Simplified Least Squares Procedures. Analytical Chemistry. 1964. Vol. 36, No. 8. P. 1627–1639. DOI: 10.1021/ac60214a047.

121. Flores A., Tito-Chura H., Yana-Mamani V., Rosado-Chavez C., Ecos-Espino A. Weighted Averages and Polynomial Interpolation for PM2.5 Time Series Forecasting. Computers. 2024. Vol. 13, No. 9. Art. 238. DOI: 10.3390/computers13090238.

122. Єріна А. М., Мазуренко О. К. Статистичний аналіз часових рядів : навчальний посібник. Київ : ВПЦ «Київський університет», 2022. 164 с.

123. Сенишин О. С. Екстраполяційні методи прогнозування як інструмент передбачення оптимальних обсягів споживання продукції вітчизняного продовольчого комплексу. Молодіжний економічний дайджест. 2014. № 1. С. 26–32.

124. Turbal Y., Bomba A., Turbal M., Abd Alkaleg Hsen Drivi. Some Aspects of Extrapolation Based on Interpolation Polynomials. Physico-Mathematical Modelling and Informational Technologies. 2021. No. 33. P. 175–180. DOI: 10.15407/fmmit2021.33.175.

125. Turbal Y., Bomba A., Turbal M., Turbal B., Smirnov D. Forecasting Algorithm Based on Intellectual Analysis of Polynomial Extrapolation. Lecture Notes in Data Engineering, Computational Intelligence, and Decision-Making. Cham : Springer, 2025. Vol. 244. P. 98–113. DOI: 10.1007/978-3-031-88483-2_5.

126. Trefethen L. N. Approximation Theory and Approximation Practice. Philadelphia : Society for Industrial and Applied Mathematics, 2013. ISBN 978-1-61197-239-9.

127. Turbal Y., Kubai O. A forecasting method based on clustering the polynomial extrapolation sequence. Computer Systems and Information Technologies. 2026. No. 2. P. 48–60. DOI: 10.31891/csit-2026-2-5.

128. Turbal, Y., Shlikhta, G., Turbal, M., & Turbal, B. (2023) The polynomial forecasts improvement based on the algorithm of optimal polynomial degree selecting. Eastern-European Journal of Enterprise Technologies, vol. 6, no. 4 (126), pp. 17–26.

129. Turbal Y., Turbal M., Kubai O., Smirnov D., Melnychuk M. Метод прогнозування на основі усереднення поліноміальної екстраполяційної послідовності. Моделювання, керування та інформаційні технології. 2025. № 8. С. 326–329. DOI: 10.31713/MCIT.2025.102.

130. Demanet L., Townsend A. Stable extrapolation of analytic functions. Foundations of Computational Mathematics. 2019. Vol. 19. P. 297–331. DOI: 10.48550/arXiv.1605.09601.

131. Castle J. L., Clements M. P., Hendry D. F. Forecasting principles from experience with forecasting competitions. Forecasting. 2021. Vol. 3, No. 1. P. 138–165. DOI: 10.3390/forecast3010010.

132. Cont R. Volatility clustering in financial markets: Empirical facts and agent-based models. Long Memory in Economics / ed. by A. Teyssière, K. Kirman. Berlin : Springer, 2007. P. 289–309. DOI: 10.2139/ssrn.1411462.

133. Burden R. L., Faires J. D. Numerical Analysis. 9th ed. Boston : Cengage Learning, 2010.

134. El-Yaniv R., Faynburd A. Autoregressive short-term prediction of turning points using support vector regression. arXiv:1209.0127. 2012.

135. Yazdani F., Khashei M., Hejazi S. R. Using a Graph-based Method for Detecting the Optimal Turning Points of Financial Time Series. Financial Research Journal. 2022. Vol. 24, No. 1. P. 18–36. DOI: 10.22059/frj.2021.327413.1007219.

136. Zaychenko Y., Zaichenko H., Kuzmenko O. Investigation of Artificial Intelligence Methods in the Short-Term and Middle-Term Forecasting in Financial Sphere. CEUR Workshop Proceedings. 2022. Vol. 3347. P. 80–89.

137. Süli E., Mayers D. F. An Introduction to Numerical Analysis. Cambridge : Cambridge University Press, 2003. 444 p.

138. Wang X., Hyndman R. J., Li F., Kang Y. (2023) Forecast combinations: An over 50-year review. International Journal of Forecasting, vol. 39, no. 4, pp. 1518–1547.

139. López-Oriona Á., Montero-Manso P., Vilar J. A. (2025) Time series clustering based on prediction accuracy of global forecasting models. Knowledge-Based Systems, vol. 312, article 113221.

140. Şen B., Ye H., Sugihara G. (2024) Detecting stochasticity in population time series using a non-parametric test of intrinsic predictability. Methods in Ecology and Evolution, vol. 15, no. 9, pp. 1741–1754.

141. Hassani H., Silva E. S., Ghodsi M. (2025) White Noise and Its Misapplications: Impacts on Time Series Model Adequacy and Forecasting. Stats, vol. 7, no. 1, pp. 8–23.

142. Karmaker S., Hossain M., Rahman M. et al. A Review of Time-Series Forecasting Algorithms for Industrial Manufacturing Systems. 2024. https://doi.org/10.3390/machines12060380.

143. Shah, Dhawani & Thaker, Manishkumar. (2024). A Review of Time Series Forecasting Methods. International journal of research and analytical reviews. 11. 749. 10.1729/Journal.38816.

144. Giantsidi A., Tarantola A. (2025) Deep learning for financial forecasting: A review of recent trends. Finance Research Letters, vol. 74, article 106885.

145. Li X., Zhang Y., Chen H. et al. (2025) Enhancing financial time series forecasting with hybrid Deep Learning: CEEMDAN-Informer-LSTM model. Applied Soft Computing, vol. 168, article 112345.

146. Saunoriene L., Mockus J., Katkevičius A. (2025) Short-term time series prediction based on evolutionary interpolation of Chebyshev polynomials with internal smoothing. Soft Computing, vol. 29, no. 1, pp. 1–15.

147. Salinas D., Flunkert V., Gasthaus J. (2017) DeepAR: Probabilistic Forecasting with Autoregressive Recurrent Networks. arXiv preprint arXiv:1704.04110.

148. Турбал Ю., Кубай О. (2026). Про особливості використання послідовностей поліноміальних прогнозів у алгоритмах інтелектуального аналізу даних за умов існування збіжності. Вимірювальна та обчислювальна техніка в технологічних процесах, (2), С. 431–441. https://doi.org/10.31891/2219-9365-2026-86-51

149. Khizar Farooq, Aljethi Reem Abdullah, Ejaz Hussain, Muhammad Amin S. Murad. (2025) Chaos, Lyapunov exponent, and sensitivity demonstration of the coupled nonlinear integrable model with soliton solutions[J]. AIMS Mathematics, 10(10): 22929-22957. doi: 10.3934/math.20251019.

150. Marwan N., Romano M. C., Thiel M., Kurths J. (2007) Recurrence plots for the analysis of complex systems // Physics Reports. Vol. 438, No. 5–6. P. 237–329. DOI: 10.1016/j.physrep.2006.11.001.

151. Riley, W. J. (2008). Handbook of frequency stability analysis. Gaithersburg: National Institute of Standards and Technology. NIST Special Publication 1065. https://doi.org/10.6028/NIST.SP.1065

152. Turbal, Y., Bomba, A., Sokh, A., Radoveniuk, O., Turbal, M. (2019). Pyramidal method of small time series extrapolation. International Journal of Computing Science and Mathematics, 10(4), 122–130.

153. Nazareth N., Ramana Reddy Y. V. Financial applications of machine learning: A literature review. Expert Systems with Applications. 2023. Vol. 219. Article 119640. DOI: 10.1016/j.eswa.2023.119640.

154. Oreshkin B. N., Carpov D., Chapados N., Bengio Y. N-BEATS: Neural basis expansion analysis for interpretable time series forecasting. arXiv. 2019. arXiv:1905.10437. DOI: 10.48550/arXiv.1905.10437.

155. Chrysos G. G., Moschoglou S., Bouritsas G., Panagakis Y., Deng J., Zafeiriou S. Π-Nets: Deep Polynomial Neural Networks. Proceedings of the

IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2020. P. 7325–7335.

156. Ghazali R., Hussain A. J., El-Deredy W. Application of Ridge Polynomial Neural Networks to Financial Time Series Prediction. 2006 IEEE International Joint Conference on Neural Networks Proceedings. Vancouver, BC, Canada, 16–21 July 2006. P. 913–920. DOI: 10.1109/IJCNN.2006.246783.

157. Lo A. W., MacKinlay A. C. Stock market prices do not follow random walks: evidence from a simple specification test. The Review of Financial Studies. 1988. Vol. 1, No. 1. P. 41–66.

158. Zhang C., Sjarif N. N. A., Ibrahim R. Deep learning models for price forecasting of financial time series: a review of recent advancements: 2020–2022. WIREs Data Mining and Knowledge Discovery. 2023.

159. Бідюк П. І., Калініна І. О., Гожий О. П. **Байєсівський аналіз даних** : монографія. Херсон : ФОП Вишемирський В. С., 2021. 208 с.

160. Кузнєцова Н. В. Методи та моделі коротко- і середньострокового прогнозування фінансових процесів. Київ : КПІ ім. Ігоря Сікорського, 2023.

161. Netflix, Inc. (NFLX) Stock Historical Prices & Data. Yahoo Finance. Електронний ресурс.

162. Netflix – Historical Stock Quote. Netflix Investor Relations. Електронний ресурс.

163. yfinance documentation: download market data from Yahoo! Finance API. Електронний ресурс.

164. O'Hara M. High frequency market microstructure. Journal of Financial Economics. 2015. Vol. 116, No. 2. P. 257–270. DOI: 10.1016/j.jfineco.2015.01.003.

165. Tashman L. J. Out-of-sample tests of forecasting accuracy: an analysis and review. International Journal of Forecasting. 2000. Vol. 16, No. 4. P. 437–450.

166. Bergmeir C., Benítez J. M. On the use of cross-validation for time series predictor evaluation. Information Sciences. 2012. Vol. 191. P. 192–213. DOI: 10.1016/j.ins.2011.12.028.

167. Sculley D., Holt G., Golovin D., Davydov E., Phillips T., Ebner D., Chaudhary V., Young M., Crespo J.-F., Dennison D. Hidden technical debt in machine learning systems. *Advances in Neural Information Processing Systems*. 2015. P. 2503–2511.

ДОДАТКИ

Додаток А

Порівняльна характеристика статистичних методів для аналізу фінансових часових рядів

№ з/п	Метод	Основна ідея	Переваги	Обмеження для коротких внутрішньоденних рядів
1.	**Наївний прогноз**	Наступне значення дорівнює останньому відомому	Простота; відсутність навчання; обов'язковий базовий орієнтир	Не враховує локальний тренд і структуру коливань
2.	**Ковзне середнє**	Усереднення останніх (m) значень	Зменшення впливу шуму	Запізнення під час зміни тренду; необхідність вибору ширини вікна
3.	**Експоненціальне згладжування**	Експоненційне зменшення ваг попередніх значень	Невелика кількість параметрів; рекурентне оновлення	Обмежена здатність описувати складні локальні коливання
4.	**AR**	Лінійна залежність від попередніх значень	Інтерпретованість; простота реалізації	Потребує стаціонарності; нестійкість оцінок за малого обсягу даних
5.	**ARMA**	Поєднання авторегресії та залежності від похибок	Гнучкіше відтворення автокореляційної структури	Необхідність вибору порядків (p) і (q); вимога стаціонарності
6.	**ARIMA**	Диференціювання нестаціонарного ряду з подальшим ARMA-моделюванням	Універсальність; розвинена методологія ідентифікації	Параметри можуть швидко втрачати актуальність за зміни локальної динаміки
7.	**SARIMA**	Урахування сезонності	Придатність для повторюваних циклів	Потребує достатньої кількості повторень сезонного шаблону
8.	**ARCH, GARCH**	Моделювання змінної умовної дисперсії	Опис кластеризації волатильності; оцінювання ризику	Переважно прогнозують волатильність, а не майбутню ціну
9.	**ARIMAX**	Урахування зовнішніх факторів	Можливість використання додаткової інформації	Ризик витоку інформації; необхідність прогнозування зовнішніх змінних
10.	**VAR, VECM**	Спільне моделювання кількох часових рядів	Урахування взаємозв'язків між показниками	Значна кількість параметрів; високі вимоги до обсягу вибірки
11.	**Моделі простору станів**	Оцінювання прихованих компонент процесу	Адаптивність; рекурентне оновлення оцінок	Необхідність коректного вибору структури та параметрів шуму
12.	**Марковські моделі**	Перемикання між прихованими режимами	Урахування структурних змін	Складність оцінювання на коротких вибірках

Додаток Б

Порівняльна характеристика методів машинного навчання для аналізу фінансових часових рядів

Метод	Тип задачі	Основні переваги	Основні обмеження	Доцільність використання
Лінійна регресія	Регресія	Простота, швидкість, інтерпретованість	Обмежена здатність відображати нелінійні залежності	Базова модель порівняння
Ridge-регресія	Регресія	Стійкість за наявності корельованих ознак, зменшення перенавчання	Необхідність вибору коефіцієнта регуляризації	Інтерпретована модель для малих вибірок
LASSO-регресія	Регресія, відбір ознак	Автоматизоване скорочення кількості ознак	Нестійкість вибору ознак за сильної кореляції	Формування компактного набору предикторів
Логістична регресія	Класифікація	Інтерпретованість, оцінювання ймовірності	Переважно лінійна межа поділу	Базова модель прогнозування напрямку
(k)-NN	Регресія, класифікація	Урахування локальних аналогій, простота	Чутливість до масштабу та розмірності ознак	Аналіз локально подібних станів
SVM, SVR	Регресія, класифікація	Ядерне моделювання нелінійних залежностей, придатність для вибірок помірного обсягу	Складність налаштування гіперпараметрів	Порівняльна нелінійна модель
Дерево рішень	Регресія, класифікація	Інтерпретованість правил, урахування нелінійностей	Нестійкість окремого дерева	Допоміжна модель
Random Forest	Регресія, класифікація	Зменшення дисперсії, моделювання нелінійних взаємодій	Обмежена екстраполяція, складніша інтерпретація	Сильна базова ансамблева модель
Градієнтний бустинг, XGBoost	Регресія, класифікація	Висока гнучкість для табличних даних, урахування взаємодій	Ризик перенавчання, потреба в налаштуванні	Розширена модель порівняння
Кластерний аналіз	Навчання без учителя	Виявлення груп, режимів і локальної узгодженості	Не формує прогноз без додаткового правила	Аналіз структури PPS
Нейронні мережі	Регресія, класифікація	Моделювання складних нелінійних залежностей	Потреба в даних, ризик перенавчання, складність інтерпретації	Окремий напрям дослідження

Додаток В

Порівняльна характеристика нейронних мереж для аналізу фінансових часових рядів

Архітектура	Основний механізм	Переваги	Обмеження для коротких біржових рядів
MLP	Нелінійне відображення локального вікна	Простота, невисокі обчислювальні витрати, придатність до малих вибірок	Відсутність внутрішньої пам’яті, необхідність явно задавати довжину вікна
RNN	Рекурентне оновлення прихованого стану	Природне врахування часової послідовності	Зникнення або вибухове зростання градієнтів
LSTM	Комірка пам’яті та система вентилів	Моделювання залежностей різної тривалості	Велика кількість параметрів, ризик перенавчання, низька інтерпретованість
GRU	Спрощені вентильні рекурентні блоки	Менша складність порівняно з LSTM	Результат істотно залежить від даних і гіперпараметрів
CNN / TCN	Каузальні та дилатовані згортки	Паралельне навчання, виділення локальних шаблонів	Необхідність обґрунтувати рецептивне поле та масштаб аналізу
Transformer	Self-attention	Моделювання глобальних залежностей, паралельні обчислення	Висока параметрична та обчислювальна складність, ризик моделювання шуму
DeepAR	Авторегресійна RNN із прогнозуванням розподілу	Оцінювання невизначеності	Переважна орієнтація на множини пов’язаних рядів
TFT	Поєднання LSTM, уваги та відбору ознак	Багатогоризонтність, часткова інтерпретованість	Складність і потреба у значному обсязі даних
N-BEATS	Базисне розкладання та залишкові зв’язки	Структуроване формування прогнозу, інтерпретовані компоненти	Потреба адаптувати базиси до конкретного процесу
N-HiTS	Ієрархічна інтерполяція та багатомасштабний аналіз	Ефективність для довгих горизонтів	Надлишкова складність для деяких однокрокових задач
PatchTST	Поділ ряду на фрагменти-токени	Робота з довгим контекстом, скорочення витрат уваги	Потреба в достатній кількості даних для навчання
iTransformer	Увага між змінними багатовимірного ряду	Урахування міжознакових зв’язків	Обмежена доцільність для короткого одновимірного ряду
Chronos	Попереднє навчання та токенізація значень	Zero-shot прогнозування, масштабованість	Обмежена прозорість і залежність від доменної відповідності

Додаток Г

Порівняльна характеристика програмних комплексів для аналізу фінансових часових рядів

Група	Приклади	Переваги	Недоліки	Умови поширення	Придатність для дисертаційного дослідження
Професійні фінансові термінали	Bloomberg Terminal, LSEG Workspace, FactSet	якісні дані, новини, аналітика, професійні workflow	висока вартість, закритість, обмеження на дані	комерційна підписка	корисні як джерело даних і приклад інституційних систем
Торгово-аналітичні платформи	TradingView, MetaTrader 5, NinjaTrader, AmiBroker	графіки, індикатори, сигнали, торгові стратегії	обмежена наукова гнучкість, залежність від платформи	freemium, безкоштовні або комерційні ліцензії	корисні для прикладного аналізу, але не як основне дослідницьке середовище
Backtesting / quant-середовища	QuantConnect/LEAN, Backtrader, vectorbt	формалізоване тестування, Python, walk-forward, trading simulation	потреба у програмуванні, ризик методичних помилок	open-source, freemium, комерційні PRO-версії	дуже придатні для експериментальної перевірки
Наукові середовища	Python, R, MATLAB	статистика, ML, DL, реалізація власних моделей	потреба в самостійному налаштуванні pipeline	Python/R – переважно open-source; MATLAB – комерційний	найбільш придатні для розробки PPS, PPS-Net і порівняльних експериментів
Time-series databases	kdb+, ArcticDB, InfluxDB, ClickHouse, TimescaleDB	масштабованість, tick data, швидкі часові запити	складність адміністрування, надмірність для малих вибірок	open-source, commercial, enterprise	перспективні для масштабування на високочастотні дані

Додаток Д

Результати кластерного аналізу з використанням методів DI та DBSCAN

Метод	Точне значення	Прогнозне значення	Абсолютна похиька	Відносна похибка, %	Ліва межа	Права межа	Кількість кластерів	Ширина кластера	Режим кластера
Densest Interval (DI)	-0,97845	-0,97845	0,00001	0,00065	-0,97846	-0,97843	2	0,00003	mean
DBSCAN	-0,97845	-0,99442	0,01597	1,63220	-1,16445	-0,94898	15	0,21547	mean

Таблиця 5.2. Результати кластерного аналізу з використанням методів DI та DBSCAN для стохастичних даних.

Метод	Точне значення	Прогнозне значення	Абсолютна похиька	Відносна похибка, %	Ліва межа	Права межа	Кількість кластерів	Ширина кластера	Режим кластера
Densest Interval (DI)	1,070824	0,217155	0,853669	79,72079	0,209855	0,224455	2	0,01460	mean
DBSCAN	1,070824	0,174162	0,896661	83,73567	0,088178	0,224455	3	0,13628	mean

Таблиця 5.3. Результати кластерного аналізу з використанням методів DI та DBSCAN для даних про акцію (NFLX, 30 хв).

Метод	Точне значення	Прогнозне значення	Абсолютна похиька	Відносна похибка, %	Ліва межа	Права межа	Кількість кластерів	Ширина кластера	Режим кластера
Densest Interval (DI)	96,6101	96,1695	0,4406	0,45606	96,1096	96,2294	2	0,1198	mean
DBSCAN	96,6101	96,2294	0,3807	0,39406	95,9195	96,6298	5	0,7103	mean

Додаток Е

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ ВОДНОГО ГОСПОДАРСТВА ТА ПРИРОДОКОРИСТУВАННЯ
(НУВГП)
вул. Соборна, 11, м. Рівне, 33028, тел. (0362) 63 30 98, факс (0362) 63 32 09,
e-mail: mail@nuwm.edu.ua, web: https://nuwm.edu.ua
код ЄДРПОУ 02071116

Від 10.04.2026 № 011-10 На № від

ДОВІДКА
про використання у навчальному процесі
Національного університету водного господарства та природокористування
результатів досліджень і розробок, одержаних при виконанні дисертаційної роботи
КУБАЯ ОЛЕКСАНДРА ВАСИЛЬОВИЧА
на здобуття ступеня доктора філософії за спеціальністю
122 – Комп'ютерні науки

Науково-практичні результати дисертаційного дослідження Кубая Олександра Васильовича щодо обґрунтування методів інтелектуального аналізу та прогнозування часових рядів біржової динаміки на основі поліноміального підходу використовуються в освітньому процесі навчально-наукового інституту будівництва, архітектури та дизайну, навчально-наукового інституту кібернетики, інформаційних технологій та інженерії Національного університету водного господарства та природокористування в процесі підготовки фахівців за освітньо-науковим ступенем «доктор філософії» за спеціальністю F3 «Комп'ютерні науки», спеціальністю G19 «Будівництво та цивільна інженерія», а також фахівців за освітнім ступенем «бакалавр» за спеціальністю F3 «Комп'ютерні науки».

Пропозиції впроваджено в освітній процес для удосконалення структури та змісту навчально-методичного забезпечення таких навчальних дисциплін, як:

«Прикладна інформатика»:
- Тема 4: «Методи та підходи до побудови прогнозів часових рядів»;
- Тема 5: «Методи прогнозування на основі многочленів»;
- Тема 6: «Пірамідальні методи прогнозування».

«Нейронні мережі та генетичні алгоритми»:
- Тема 8: «Спеціалізовані нейронні мережі».

Використання результатів дисертаційного дослідження Кубая О. В. свідчить про їх завершеність, актуальність і можливість впровадження в освітній процес закладів вищої освіти.

Проректор з наукової роботи
та міжнародних зв'язків Національного
університету водного господарства та
природокористування
д. е. н., професор Наталія САВІНА

Юрій ТУРБАЛ +380980841880

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ

НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ ВОДНОГО ГОСПОДАРСТВА ТА ПРИРОДОКОРИСТУВАННЯ

вул. Соборна, 11, м. Рівне, 33028, тел. (0362) 63-30-98, факс (0362) 63-32-09, mail@nuwm.edu.ua

Від 20.04.26 № 204-26

На № _____ від _____

ДОВІДКА

про використання досліджень, отриманих при виконанні дисертаційної роботи
Кубая Олександра Васильовича
на здобуття ступеня доктора філософії за спеціальністю 122 – комп'ютерні науки,
в науково-дослідних роботах, виконаних у науково-дослідний частині
Національного університету водного господарства та природокористування

Видана ***Кубаю Олександру Васильовичу*** з підтвердженням того, що результати досліджень, викладених в дисертаційній роботі на тему ***«Методи інтелектуального аналізу та прогнозування часових рядів біржової динаміки на основі поліноміального підходу»*** використовувались при виконанні наукової теми кафедри комп'ютерних наук та прикладної математики Національного університету водного господарства та природокористування **«Математичне та комп'ютерне моделювання процесів та систем»** (номер державної реєстрації 0125U001313), де автором запропоновано архітектуру нейронної мережі для прогнозування часових рядів біржової динаміки, яка характеризується введенням спеціального блоку формування послідовності поліноміальних прогнозних значень, блоку нейромережевого аналізу її структури та блоку адаптивного зважування поліноміальних прогнозів. Це дає можливість використовувати не лише початкові значення вихідного часового ряду, а й множину поліноміально обчислених прогнозних значень, забезпечуючи таким чином інтерпретоване формування прогнозу.

Крім того, результати досліджень використовувалися при виконанні науково-дослідних робіт в науково-дослідній частині Національного університету водного господарства та природокористування, а саме: **«Вдосконалення інформаційних технологій обробки даних на основі оптимізованих комп'ютерних систем»** (номер державної реєстрації 0121U111219), де автором удосконалено інформаційну технологію інтелектуальної обробки та прогнозування часових рядів біржової динаміки. Запропонований підхід передбачає формування послідовності поліноміальних прогнозних значень, аналіз її структурних властивостей, виділення інформативної підмножини прогнозів, адаптивний вибір оптимальної кількості її елементів та обчислення результуючого прогнозного значення. Використання розроблених алгоритмів дає змогу підвищити ефективність обробки фінансових даних у комп'ютерних системах шляхом автоматизованого налаштування параметрів поліноміальної екстраполяції відповідно до локальної поведінки часового ряду, зменшення впливу випадкових коливань і підвищення обґрунтованості вибору прогнозної моделі в умовах невизначеності та високої волатильності біржових процесів.

Проректор з наукової роботи
та міжнародних зв'язків НУВГП
д. е. н., професор Наталія САВІНА

Куницький 0967675013

ТОВАРИСТВО З ОБМЕЖЕНОЮ ВІДПОВІДАЛЬНІСТЮ
«ВІВА СОФТ»

Адреса: 33001, м. Рівне, вул. Дубенська, БОС 9/9
Р/р: UA173333910000026004054739351
Банк: ПАТ КБ «ПРИВАТБАНК»
МФО: 333391
ЄДРПОУ: 42765487

ДОВІДКА

про впровадження результатів дисертаційного дослідження
Кубая Олександра Васильовича
«Методи інтелектуального аналізу та прогнозування часових рядів біржової динаміки на основі поліноміального підходу»
на здобуття ступеня доктора філософії за спеціальністю 122 – Комп'ютерні науки.

Результати, які отримані в результаті дисертаційного дослідження Кубая Олександра Васильовича «Методи інтелектуального аналізу та прогнозування часових рядів біржової динаміки на основі поліноміального підходу» впроваджено у виробничу діяльність ТОВ «ВІВА СОФТ».

Практичне використання отриманих результатів здійснюється у процесі розроблення та вдосконалення програмних засобів інтелектуальної обробки даних, призначених для аналізу та прогнозування часових рядів фінансово-економічних показників.

У виробничій діяльності підприємства використано результати дисертаційного дослідження, які:

- передбачають формування локального вікна даних, побудову послідовності поліноміальних прогнозних значень та автоматизований вибір параметрів прогнозування відповідно до локальної структури досліджуваного процесу;
- забезпечують можливість оцінювання характеру зміни прогнозів залежно від параметра поліноміальної екстраполяції, виявлення областей згущення, локальних екстремумів, монотонних ділянок, ознак збіжності або розбіжності прогнозних значень;
- дають змогу формувати остаточне прогнозне значення шляхом адаптивного відбору та усереднення найбільш узгоджених елементів послідовності (PPS);
- враховують структурні властивості послідовності поліноміальних прогнозних значень та дозволяють виявляти випадки зниження надійності результатів за умов високої волатильності, шуму або переважання стохастичної складової в досліджуваних даних;
- програмно реалізують формування локальних вибірок, обчислення поліноміальних прогнозів, аналіз структури (PPS), вибір параметрів прогнозування та візуалізацію отриманих результатів.

Також в процесі розроблення програмного забезпечення здійснюється апробація нейронної мережі PPS-Net прогнозування часових рядів, що містить спеціальний блок формування послідовності поліноміальних прогнозних значень (PPS), блок нейромережевого аналізу її структури та блок адаптивного зважування окремих прогнозів. Її використання забезпечує інтерпретоване формування результуючого прогнозного значення та дозволяє визначати внесок поліноміальних прогнозів різних порядків у кінцевий результат.

11.05.2026 р.

Директор _____ (Райкович О.В.)

Додаток Ж

Список публікацій здобувача за темою дисертації

Статті у фахових наукових виданнях України

1. Турбал Ю., Турбал М., Кубай О., Смірнов Д., Мельничук М. (2025) Метод прогнозування на основі усереднення поліноміальної екстраполяційної послідовності. Моделювання, керування та інформаційні технології. № 8. С. 326–329. DOI: 10.31713/MCIT.2025.102.

2. Турбал Ю., Кубай О. (2025) Теорема Ковера та задача відокремлення множин за допомогою поліноміальних порогових функцій (PTF – Polynomial Threshold Functions). Вісник Національного університету водного господарства та природокористування. Технічні науки. №2. С. 178–191. DOI: https://doi.org/10.31713/vt2202516.

3. Turbal Y., Kubai O. (2026) A forecasting method based on clustering the polynomial extrapolation sequence. Computer Systems and Information Technologies. No. 2. P. 48–60. DOI: 10.31891/csit-2026-2-5

4. Турбал Ю., Кубай О. (2026). Про особливості використання послідовностей поліноміальних прогнозів у алгоритмах інтелектуального аналізу даних за умов існування збіжності. Вимірювальна та обчислювальна техніка в технологічних процесах, (2), С. 431–441. https://doi.org/10.31891/2219-9365-2026-86-51

5. Турбал Ю.,Кубай, О. (2026). Аналіз ступеня стохастичності процесів на основі послідовності поліноміальних прогнозів (PPS). Комп’ютерно-інтегровані технології: освіта, наука, виробництво, (63), 277-286. https://doi.org/10.36910/6775-2524-0560-2026-63-31

6. Турбал Ю., Кубай О. (2026) Архітектура спеціалізованої нейронної мережі для прогнозування фінансових часових рядів на основі послідовності поліноміальних прогнозів. Вісник Національного університету водного господарства та природокористування. Технічні науки. №1. С. 365–384. DOI: https://doi.org/10.31713/vt2202516.

Публікації в матеріалах конференції, тези доповідей

7. Кубай О.В Особливості цифровізації правових відносин в умовах воєнного стану. Матеріали регіональної науково-практичної конференції «Актуальні проблеми національного адміністративного права в умовах воєнного стану», 2.03.2023 р.,

8. Кубай О.В., Пилипюк Ю.Ю. Роль та використання інформаційно-правових цифрових сервісів в умовах воєнного стану. Матеріали регіональної науково-практичної конференції «Актуальні проблеми національного адміністративного права в умовах воєнного стану», 2.03.2023 р.,

9. Кубай О.В., Турбал Ю.В. Використання штучних нейронних мереж в бізнес процесах: передумови, проблеми, перспективи Міжнародна науково-практична конференція молодих науковців, аспірантів і здобувачів вищої освіти «Проблеми та перспективи розвитку сучасної науки», 11-12 травня 2023 р, м. Рівне, 532 с.

10. Кубай О. Регулювання використання штучного інтелекту в США як модель для України. Системи та засоби штучного інтелекту: тези доповідей Міжнародної наукової конференції «Штучний інтелект: досягнення, виклики та ризики». – Київ: ІПШІ «Наука і освіта», 15-16.03.2024. – 550 с.

11. Кубай О.В. Верифікація та валідація різнотипних даних як ключові аспекти для їх подальшої автоматизованої обробки». Міжнародна науково-практична конференція молодих науковців, аспірантів і здобувачів вищої освіти «Проблеми та перспективи розвитку сучасної науки», 9-10 травня 2024 р, м. Рівне.