

<https://doi.org/10.31891/2219-9365-2026-85-8>

УДК 004.75

ЛУТЮК Лев

Хмельницький національний університет

<https://orcid.org/0009-0004-7497-3427>

e-mail: lev_lutiuk@ukr.net

КАШТАЛЬЯН Антоніна

Хмельницький національний університет

<https://orcid.org/0000-0002-4925-9713>

e-mail: yantonina@ukr.net

КОВАЛЬЧУК Василь

Хмельницький національний університет

<https://orcid.org/0009-0008-7013-5919>

e-mail: vasuli458@gmail.com

ВИБІР АРХІТЕКТУР ГЛИБОКИХ НЕЙРОННИХ МЕРЕЖ У ЗАДАЧАХ ПРОГНОЗУВАННЯ ФІНАНСОВИХ РИНКІВ ЗА УМОВ ВИСОКОЇ ВОЛАТИЛЬНОСТІ

Прогнозування фінансових ринків є однією з ключових задач сучасного фінансового аналізу, ускладненою високою волатильністю, нелінійною поведінкою ринку та впливом численних економічних, політичних і поведінкових факторів. У таких умовах традиційні статистичні методи часто виявляються недостатніми, що зумовлює зростання застосування підходів глибокого навчання.

У статті розглядається проблема вибору архітектур глибоких нейронних мереж для прогнозування фінансових ринків за умов високої волатильності. Досліджуються рекурентні, згорткові, часові згорткові, гібридні моделі, а також архітектури з механізмами уваги та нейронні мовні моделі. Особлива увага приділяється проблемі запізнювання прогнозів та її впливу на точність передбачень у динамічному ринковому середовищі.

Експериментальні результати демонструють переваги гібридних моделей та моделей з механізмами уваги у забезпеченні високої точності прогнозів і адаптивності моделей у високоволатильних умовах.

Отримані результати можуть бути використані для обґрунтування вибору архітектур нейронних мереж та розробки інформаційних технологій прогнозування фінансових ринків, спрямованих на підвищення точності та стійкості прогнозів.

Ключові слова: інформаційні технології, нейронні мережі, фінансові часові ряди, висока волатильність, проблема запізнювання прогнозів, гібридні моделі, нейронні мовні моделі

LUTIUK Lev, KASHTALIAN Antonina, KOVALCHUK Vasily

Khmelnyskyi National University

SELECTION OF DEEP NEURAL NETWORK ARCHITECTURES FOR FORECASTING FINANCIAL MARKETS UNDER CONDITIONS OF HIGH VOLATILITY

The paper addresses the problem of selecting deep neural network architectures for financial market forecasting under conditions of high volatility, nonstationarity, and a high level of noise in financial time series. Traditional statistical and econometric approaches often fail to capture nonlinear dependencies and rapidly changing market regimes, which necessitates the use of deep learning models capable of adaptive representation learning and multiscale temporal analysis.

The study provides a systematic comparative analysis of recurrent neural networks (RNN), LSTM and GRU architectures, convolutional and temporal convolutional networks (CNN, TCN), attention-based models, transformer architectures, and neural language models for integrating textual financial information. Particular attention is paid to the prediction lag problem, which significantly reduces the practical value of forecasts in highly dynamic market environments.

It is shown that recurrent architectures effectively model long-term dependencies but may exhibit delayed responses to rapid short-term fluctuations. CNN and TCN models demonstrate higher robustness to noise and better extraction of local patterns, while transformer-based models enable global context modeling and faster adaptation to new information. Neural language models allow the integration of news flows and sentiment analysis into forecasting systems, improving responsiveness to market events.

Experimental results obtained on volatile financial time series indicate that hybrid architectures combining convolutional, recurrent, and attention mechanisms provide the best balance between forecasting accuracy, robustness, and computational efficiency. Such models simultaneously capture short-term dynamics, long-term trends, and contextual information, reducing prediction lag and improving generalization under changing market regimes.

The proposed systematization of deep neural network architectures and their characteristics can be used to justify the selection of forecasting models and to develop information technologies for financial market prediction aimed at increasing accuracy, stability, and adaptability in high-volatility conditions.

Keywords: information technologies, neural networks, financial time series, high volatility, prediction lag problem, hybrid models, neural language models

Стаття надійшла до редакції / Received 30.01.2026

Прийнята до друку / Accepted 16.02.2026

Опубліковано / Published 05.03.2026



This is an Open Access article distributed under the terms of the [Creative Commons CC-BY 4.0](https://creativecommons.org/licenses/by/4.0/)

© Лутюк Лев, Каштальян Антоніна, Ковальчук Василь

ПОСТАНОВКА ПРОБЛЕМИ У ЗАГАЛЬНОМУ ВИГЛЯДІ ТА ЇЇ ЗВ'ЯЗОК ІЗ ВАЖЛИВИМИ НАУКОВИМИ ЧИ ПРАКТИЧНИМИ ЗАВДАННЯМИ

Прогнозування фінансових ринків є однією з найбільш складних і водночас важливих задач у сучасних інформаційних технологіях фінансового аналізу. Поведінка фінансових інструментів формується під впливом численних економічних, політичних, соціальних і поведінкових факторів, що зумовлює високий рівень волатильності та нестабільності ринкових процесів. За таких умов фінансові часові ряди характеризуються нелінійністю, нестационарністю та значною зашумленістю, що істотно ускладнює побудову точних і стійких прогнозних моделей.

Традиційні статистичні та економетричні методи прогнозування, як правило, ґрунтуються на припущеннях щодо лінійності процесів або стабільності їхніх параметрів у часі, що обмежує можливості їх застосування в умовах різких ринкових коливань. У результаті такі підходи часто не здатні адекватно відобразити складні взаємозв'язки між змінними та швидко адаптуватися до змін ринкового середовища, що призводить до зниження точності прогнозів.

Розвиток методів глибокого навчання та зростання обчислювальних можливостей відкрили нові перспективи для прогнозування фінансових ринків. Глибокі нейронні мережі здатні ефективно працювати з великими обсягами даних, моделювати нелінійні залежності та враховувати часову структуру фінансових рядів. Водночас практичне застосування таких моделей супроводжується низкою проблем, серед яких особливе місце займає складність вибору відповідної архітектури нейронної мережі з урахуванням специфіки фінансових даних і рівня ринкової волатильності.

Різні архітектури глибоких нейронних мереж, зокрема рекурентні, згорткові, часові згорткові, гібридні моделі та моделі з механізмами уваги, демонструють відмінні властивості при роботі з фінансовими часовими рядами. Неправильний вибір архітектури або її параметрів може призводити до перенавчання, втрати здатності до узагальнення, а також до виникнення проблеми запізнювання прогнозів, що є критичним у високодинамічному ринковому середовищі.

Таким чином, актуальною науково-практичною проблемою є обґрунтований вибір архітектур глибоких нейронних мереж у задачах прогнозування фінансових ринків за умов високої волатильності, що вимагає врахування властивостей фінансових часових рядів, особливостей різних типів нейронних мереж та їхньої здатності забезпечувати точні й стійкі прогнози в умовах змінного ринкового середовища.

АНАЛІЗ ДОСЛІДЖЕНЬ ТА ПУБЛІКАЦІЙ

У сучасних наукових дослідженнях значна увага приділяється використанню методів машинного навчання у фінансовому прогнозуванні, плануванні та аналітиці. У роботах [1,2] розглядаються підходи до застосування алгоритмів машинного навчання в задачах фінансового аналізу та підтримки управлінських рішень, де підкреслюється їх здатність автоматично обробляти великі обсяги різнорідних даних. Водночас зазначається, що класичні методи машинного навчання переважно орієнтовані на прогнозування і можуть мати обмеження при використанні в умовах високої нестабільності та змінності фінансових процесів.

Окремий напрям досліджень присвячений нейронним мережам та їх застосуванню у фінансових ринках. У працях [3-6] здійснено огляд сучасних архітектур глибокого навчання та проаналізовано їх роль у прогнозуванні фінансових часових рядів і оптимізації інвестиційних портфелів. Наголошується, що рекурентні нейронні мережі та їх модифікації, зокрема LSTM і GRU, є ефективними для моделювання часових залежностей, однак можуть демонструвати зниження якості прогнозів у разі незначних або поступових коливань ринку.

У роботах [7,8] описується явище запізнення прогнозів, за якого передбачувані значення систематично відстають від реальних. Експериментальні результати свідчать, що більшість популярних рекурентних архітектур схильні до цієї проблеми, що пов'язується з обмеженнями нелінійних функцій та деградацією здатності моделей до адекватної апроксимації складних ринкових процесів. Це знижує їх придатність для застосування в умовах високої волатильності та при необхідності оперативного реагування на зміни.

Значний пласт досліджень зосереджений на прогнозуванні дохідності акцій. У роботі [9] підкреслюється важливість вибору адекватної моделі для оцінювання майбутньої прибутковості фінансових інструментів. У дослідженнях [10-11] наведено результати застосування рекурентних нейронних мереж для прогнозування дохідності, тоді як ефективність часових згорткових мереж досліджувалася менш системно [12-14]. Водночас багатогоризонтне прогнозування, яке є особливо важливим для аналізу ринків з високою волатильністю, залишається недостатньо вивченим [15-18].

Отже, проведений огляд наукових досліджень свідчить про значний інтерес до застосування глибоких нейронних мереж у задачах прогнозування фінансових ринків. Існуючі роботи демонструють високу ефективність різних архітектур нейронних мереж при аналізі фінансових часових рядів, водночас результати таких досліджень суттєво залежать від типу використовуваної моделі, горизонту прогнозування та характеристик ринкового середовища. Незважаючи на наявність численних підходів, питання обґрунтованого вибору архітектур глибоких нейронних мереж з урахуванням високої волатильності фінансових ринків

залишається недостатньо систематизованим і потребує подальшого дослідження. Це зумовлює доцільність порівняльного розгляду сучасних архітектур глибокого навчання, представлення узагальненої таблиці характеристик та проведення експериментальних досліджень для визначення їхніх переваг і обмежень у практичному застосуванні.

ВИКЛАД ОСНОВНОГО МАТЕРІАЛУ

Фінансові часові ряди характеризуються низкою специфічних властивостей, які істотно ускладнюють задачу прогнозування, особливо в умовах підвищеної волатильності ринку. На відміну від багатьох технічних або фізичних процесів, фінансові дані є нестационарними, шумними та схильними до різких структурних змін, що зумовлено впливом макроекономічних факторів, поведінки учасників ринку, інформаційних подій і зовнішніх шоків.

Однією з ключових характеристик фінансових часових рядів є нестационарність, яка проявляється у зміні статистичних властивостей даних з часом. Середнє значення, дисперсія та кореляційна структура можуть суттєво відрізнятися залежно від ринкової фази, що робить складним використання моделей із фіксованими параметрами. В умовах високої волатильності нестационарність посилюється, оскільки періоди відносної стабільності швидко змінюються фазами різких коливань, викликаних економічними новинами, політичними подіями або спекулятивною активністю.

Іншою важливою властивістю фінансових рядів є наявність значного рівня шуму, який маскує корисні сигнали. Короткострокові флуктуації цін часто не несуть фундаментального змісту, проте можуть суттєво впливати на процес навчання моделей прогнозування. У періоди високої волатильності співвідношення між сигналом і шумом погіршується, що підвищує ризик перенавчання та знижує здатність моделі до узагальнення.

Фінансові часові ряди також характеризуються явищем кластеризації волатильності, коли періоди високих коливань змінюються серіями менш інтенсивних змін. Це означає, що волатильність має властивість зберігатися протягом певного часу, формуючи залежності, які є складними для моделювання традиційними лінійними методами. Для нейронних мереж це створює додаткові вимоги щодо здатності моделі адаптуватися до різних режимів ринку без втрати стабільності прогнозів.

Особливої уваги потребує проблема швидкої зміни ринкових режимів, коли історичні залежності втрачають актуальність у дуже короткі проміжки часу. У таких умовах модель прогнозування повинна не лише враховувати довгострокові тренди, а й оперативно реагувати на локальні зміни. Недостатня чутливість до нової інформації призводить до ефекту запізнювання прогнозу, що є критичним для практичного використання моделей у трейдингу та управлінні ризиками.

Таким чином, фінансові часові ряди в умовах високої волатильності формують складне середовище для прогнозування, у якому поєднуються нестационарність, шумність, кластеризація волатильності та часта зміна ринкових режимів. Це зумовлює необхідність використання таких архітектур нейронних мереж, які здатні одночасно моделювати коротко- та довгострокові залежності, швидко адаптуватися до нових умов і зберігати стійкість до шумових компонентів даних.

Властивості фінансових часових рядів, особливо за умов високої волатильності, формують специфічні вимоги до архітектури моделей прогнозування. Ефективність нейронної мережі в цьому контексті визначається не лише її здатністю до апроксимації складних нелінійних залежностей, а й стійкістю до шуму, адаптивністю до зміни ринкових режимів і швидкістю реакції на нову інформацію.

Однією з ключових архітектурних вимог є здатність моделі одночасно враховувати короткострокові та довгострокові залежності у часових рядах. Короткострокові коливання відображають реакцію ринку на оперативні події та інформаційні сигнали, тоді як довгострокові тренди формуються під впливом фундаментальних економічних факторів. Архітектури, орієнтовані лише на один із цих аспектів, зазвичай демонструють обмежену ефективність у нестабільних ринкових умовах.

Важливою вимогою є мінімізація запізнювання прогнозу. У середовищі з високою волатильністю затримка реакції моделі навіть на декілька часових кроків може призводити до суттєвих похибок і зниження практичної цінності прогнозів. З цієї точки зору архітектура повинна забезпечувати швидке оновлення внутрішніх представлень і ефективну обробку нових спостережень без надмірної інерційності.

Ще одним критичним аспектом є стійкість моделі до шуму та аномальних значень. Фінансові дані часто містять випадкові викиди, зумовлені спекулятивною активністю або короткочасними ринковими дисбалансами. Архітектура нейронної мережі повинна мати вбудовані механізми, які дозволяють відокремлювати значущі патерни від випадкових флуктуацій, не втрачаючи при цьому чутливості до реальних змін ринкової динаміки.

Адаптивність до зміни ринкових режимів є ще однією суттєвою вимогою. Фінансові ринки можуть переходити від фаз низької волатильності до кризових станів у дуже короткі проміжки часу. У таких умовах архітектура моделі повинна забезпечувати можливість ефективного перенавчання або оновлення параметрів без втрати накопичених знань про попередні режими функціонування ринку.

Не менш важливим фактором є обчислювальна ефективність моделі. Глибокі нейронні мережі з великою кількістю параметрів можуть демонструвати високу точність, однак їх застосування в реальних фінансових системах часто обмежується вимогами до швидкодії та доступних обчислювальних ресурсів. Тому архітектурні рішення повинні враховувати компроміс між складністю моделі та її практичною придатністю для використання в режимі, близькому до реального часу.

Таким чином, архітектура нейронної мережі для прогнозування фінансових ринків за умов високої волатильності повинна поєднувати здатність до моделювання багатомасштабних часових залежностей, швидку реакцію на нову інформацію, стійкість до шуму, адаптивність до змін ринкових режимів та обчислювальну ефективність. Саме з позиції цих вимог у наступних розділах буде проведено аналіз основних архітектур глибоких нейронних мереж, що застосовуються у задачах прогнозування фінансових ринків.

Рекурентні нейронні мережі (RNN) традиційно розглядаються як базова архітектура для роботи з часовими рядами завдяки здатності моделювати послідовні залежності між спостереженнями. У фінансових задачах це дозволяє враховувати вплив попередніх значень ринку на поточні показники, що є критично важливим для прогнозування цін акцій, валютних курсів або обсягів торгів. Проте стандартні RNN демонструють обмежену ефективність на довгих часових рядах, оскільки їх навчання ускладнюється явищем затухання градієнта, що знижує здатність мережі утримувати інформацію про події минулого.

Для подолання цих обмежень були розроблені модифікації RNN, зокрема LSTM (Long Short-Term Memory) та GRU (Gated Recurrent Unit). LSTM володіє складною системою комірок пам'яті та механізмами контролю інформаційного потоку через гейти, що дозволяє ефективно моделювати довгострокові залежності та зберігати інформацію про минулі події протягом тривалого часу. Це особливо важливо для фінансових ринків, де довгострокові тенденції формуються під впливом фундаментальних факторів та накопичених історичних даних.

GRU, будучи спрощеною альтернативою LSTM, поєднує функції входу та виходу гейтів у єдиний механізм, що забезпечує менш складну структуру та швидше навчання при збереженні здатності до моделювання довгострокових залежностей. В умовах високої волатильності GRU часто демонструє кращий компроміс між точністю прогнозу та обчислювальними витратами, дозволяючи оперативніше реагувати на зміни ринкових умов.

Основним викликом для рекурентних архітектур у задачах прогнозування фінансових ринків є проблема запізнення прогнозу, коли зміни, що відбулися на ринку, відображаються у прогнозах із затримкою. Це явище проявляється особливо сильно в умовах високої волатильності, коли короткострокові коливання швидко змінюють ринкову динаміку. Недостатня чутливість стандартних RNN до таких змін обмежує їх практичну цінність для трейдингу та прийняття інвестиційних рішень у реальному часі.

Для зменшення цих обмежень використовують різноманітні стратегії вдосконалення архітектури та навчання. До них належать поєднання RNN із зовнішніми механізмами уваги, гібридизація з CNN або TCN, а також застосування адаптивних алгоритмів оптимізації для швидкого оновлення вагів при появі нових даних. Ці підходи дозволяють рекурентним мережам ефективніше реагувати на короткострокові коливання, одночасно зберігаючи здатність моделювати довгострокові тренди.

Таким чином, рекурентні архітектури, особливо у модифікаціях LSTM і GRU, зберігають важливе значення для прогнозування фінансових часових рядів, але їх ефективність у умовах високої волатильності значною мірою залежить від здатності архітектури адаптуватися до динамічних змін, зменшувати запізнення прогнозу та підтримувати баланс між точністю та обчислювальними ресурсами. Це формує підґрунтя для порівняльного аналізу інших архітектур, що розглядаються у наступних розділах.

Згорткові нейронні мережі (CNN) традиційно застосовуються для обробки просторових даних, проте їх здатність виявляти локальні патерни робить їх корисними й для аналізу фінансових часових рядів. У цьому контексті CNN дозволяють ефективно виявляти короткострокові закономірності у динаміці цін, обсягах торгів та інших ринкових показниках. Виявлення локальних патернів є особливо важливим в умовах високої волатильності, коли ринок демонструє часті та швидкі коливання, а традиційні рекурентні архітектури можуть не встигати оперативно реагувати на зміни.

Основною перевагою CNN у фінансовому прогнозуванні є здатність до автоматичного виділення найбільш значущих ознак із великого обсягу даних, що знижує вплив шуму та випадкових коливань. Це дозволяє моделі фокусуватися на релевантних паттернах і підвищує точність короткострокових прогнозів. Проте CNN мають обмеження щодо моделювання довгострокових залежностей, оскільки стандартні згорткові шари орієнтовані на локальні інтервали даних.

Для подолання цього обмеження застосовуються тимчасові згорткові мережі (TCN), які поєднують переваги CNN із можливістю ефективного аналізу часових рядів. TCN використовують каузальні та розширені згортки, що забезпечує збереження послідовності подій та дозволяє моделювати довгострокові залежності без обмежень, характерних для класичних RNN. Каузальні згортки гарантують, що прогноз для конкретного моменту часу формується виключно на основі минулих даних, що є критично важливим для фінансових прогнозів у реальному часі.

TCN демонструють високу стійкість до шуму та здатність швидко реагувати на локальні зміни ринку, що робить їх особливо корисними в умовах високої волатильності. Водночас архітектура TCN дозволяє досягти ефективного компромісу між точністю прогнозу та обчислювальними витратами, оскільки паралельне обчислення згорткових шарів пришвидшує навчання та прогнозування порівняно з рекурентними підходами.

Особливо ефективним є поєднання CNN або TCN із рекурентними архітектурами у гібридних моделях. Такі моделі спочатку використовують згорткові шари для виділення локальних патернів, а потім передають отримані ознаки до рекурентних або LSTM/GRU шарів для моделювання довгострокових залежностей. Це дозволяє одночасно враховувати коротко- та довгострокові коливання, підвищуючи точність прогнозів та зменшуючи запізнення, що є критичним у високоволатильних ринкових умовах.

Таким чином, CNN та TCN доповнюють рекурентні архітектури, забезпечуючи ефективну обробку локальних закономірностей і підвищуючи швидкість адаптації до швидкоплинних змін ринку. Їх інтеграція у гібридні моделі відкриває можливості для побудови прогнозних систем, здатних одночасно моделювати короткі флуктуації та довгострокові тенденції фінансових часових рядів.

Архітектури на основі механізмів уваги та трансформерів здобули значну популярність у задачах обробки послідовних даних завдяки здатності моделювати глобальні залежності та ефективно захоплювати контекст на різних часових масштабах. На відміну від рекурентних та згорткових мереж, трансформери не обмежені обробкою даних послідовно, що дозволяє одночасно враховувати всі часові точки в ряді та прискорює навчання моделі.

Механізм самоуваги (self-attention) є ключовим компонентом трансформерів. Він дозволяє моделі визначати, які минулі спостереження мають найбільший вплив на прогноз у конкретний момент часу. Для фінансових часових рядів це особливо важливо, оскільки ціни акцій та інші ринкові показники можуть залежати як від недавніх змін, так і від подій, що сталися багато часу тому. Самоувага забезпечує здатність моделі одночасно моделювати короткострокові коливання та довгострокові тренди, що підвищує точність прогнозів у умовах високої волатильності.

Трансформери також забезпечують стійкість до шуму та аномалій у даних завдяки здатності зважувати значущість кожного вхідного сигналу. Це дозволяє виділяти ключові патерни та ігнорувати випадкові коливання, що особливо важливо для фінансових ринків із високою частотою незначних флуктуацій. Крім того, архітектура трансформера сприяє зменшенню запізнення прогнозу, оскільки механізм уваги може оперативніше реагувати на нові дані без потреби в послідовному оновленні стану, характерному для RNN.

Особливу увагу у фінансовому прогнозуванні приділяють поєднанню трансформерів із іншими архітектурами у гібридні моделі. Наприклад, комбінація CNN для виділення локальних патернів, LSTM або GRU для моделювання довгострокових залежностей і трансформерів для глобального контекстного аналізу дозволяє створювати системи, здатні одночасно враховувати короткострокові флуктуації, середньострокові тренди та історичні закономірності на великому часовому масштабі. Такі моделі демонструють високу адаптивність у мінливих ринкових умовах, знижують похибку прогнозу та підвищують практичну цінність прогнозів для трейдерів та інвесторів.

Підсумовуючи, архітектури з механізмами уваги та трансформери є перспективним підходом для прогнозування фінансових ринків, оскільки вони забезпечують одночасну обробку локальних і глобальних залежностей, стійкість до шуму, швидку адаптацію до змін ринкових режимів і ефективне моделювання складних часових структур. Це робить їх незамінним інструментом для побудови сучасних систем прогнозування, особливо в умовах високої волатильності фінансових часових рядів.

В умовах високої волатильності фінансові ринки реагують не лише на історичні ціни та обсяги торгів, а й на інформаційний потік у вигляді новин, аналітичних оглядів, публікацій у соціальних мережах та фінансових звітів компаній. Аналіз текстових даних стає критично важливим для прогнозування ринкових тенденцій, і тут нейронні мовні моделі на основі архітектури трансформерів демонструють високу ефективність.

Ключовим механізмом у таких моделях є здатність до контекстного аналізу та моделювання семантичних взаємозв'язків між словами і фразами. Це дозволяє виділяти значущу інформацію, наприклад оцінювати тональність новин, визначати потенційний вплив макроекономічних або корпоративних подій на ринок і прогнозувати реакцію ціни активу. Моделі типу GPT (Generative Pre-trained Transformer) та BERT (Bidirectional Encoder Representations from Transformers) широко застосовуються для задач генерації тексту, класифікації настроїв та витягування ключової інформації.

GPT є генеративною моделлю, яка дозволяє створювати текст на основі попередньо навченої інформації. Це може бути корисно для автоматичного створення аналітичних оглядів або коротких викладок фінансових новин, що допомагає трейдерам та аналітикам швидко орієнтуватися у великому потоці інформації. BERT, завдяки двосторонньому контекстному кодуванню, ефективно витягує зв'язки між словами та фразами, що дозволяє здійснювати глибоку класифікацію та оцінку тональності текстів. Наприклад, аналіз

позитивного або негативного тону повідомлень може сигналізувати про можливі коливання цін на фінансових ринках.

Поєднання нейронних мовних моделей із традиційними моделями для прогнозування часових рядів, такими як LSTM, GRU, CNN чи трансформери для числових даних, дозволяє створювати гібридні системи прогнозування, що враховують як історичні дані, так і текстову інформацію. Такий підхід підвищує точність прогнозів, особливо у високоволатильних умовах, де ринок реагує миттєво на новини або неочікувані події.

Завдяки своїй здатності швидко обробляти великі обсяги текстових даних, нейронні мовні моделі забезпечують зменшення запізнювання прогнозу та підвищують адаптивність системи. Це робить їх ефективним інструментом для трейдерів, що працюють у режимі високочастотного або алгоритмічного трейдингу, а також для аналітичних платформ, що відстежують новинні потоки в реальному часі.

Отже, нейронні мовні моделі значно розширюють можливості прогнозування фінансових ринків, дозволяючи інтегрувати якісну текстову інформацію в моделі часових рядів, підвищувати точність прогнозів та оперативно реагувати на швидкі зміни ринкової динаміки.

Сучасні підходи до прогнозування фінансових ринків демонструють, що жодна окрема архітектура нейронних мереж не здатна оптимально вирішувати всі завдання у високоволатильних умовах. Рекурентні мережі (RNN, LSTM, GRU) ефективно моделюють послідовні залежності та довгострокові тренди, але можуть демонструвати запізнювання при різких короткострокових коливаннях. CNN та TCN успішно виділяють локальні патерни і швидко реагують на короткострокові зміни, проте обмежені у моделюванні глобальних залежностей. Архітектури на основі механізмів уваги та трансформери забезпечують аналіз всього контексту часового ряду, а нейронні мовні моделі дозволяють інтегрувати інформаційний потік з текстових джерел.

Гібридні моделі поєднують переваги різних архітектур, створюючи системи, здатні одночасно враховувати короткострокові та довгострокові закономірності, локальні патерни та глобальний контекст, а також текстову інформацію. Наприклад, комбінація CNN або TCN для виявлення локальних трендів, LSTM або GRU для довгострокових залежностей і трансформерів для глобального контекстного аналізу дозволяє моделі адаптуватися до динамічних змін ринку, знижувати похибку прогнозу та підвищувати практичну цінність результатів. Додавання нейронних мовних моделей дозволяє системі миттєво реагувати на новини, звіти та соціальні сигнали, що особливо важливо при високій волатильності.

Ключовими аспектами інтеграції є вибір оптимальної послідовності шарів, механізмів уваги та методів регуляризації для зменшення перенавчання. Важливе значення має також налаштування вагових коефіцієнтів та алгоритмів оптимізації для забезпечення швидкого адаптивного навчання в умовах змінного ринку. Такі гібридні системи дозволяють отримати комплексну модель, що поєднує високу точність прогнозів, стійкість до шуму та оперативну адаптацію до нових даних.

Узагальнюючи розглянуті підходи до побудови та застосування архітектур глибоких нейронних мереж у задачах прогнозування фінансових ринків, можна відзначити, що кожен тип моделі характеризується власними перевагами та обмеженнями, зумовленими як структурними особливостями нейронної мережі, так і специфікою фінансових часових рядів. Ефективність прогнозування значною мірою залежить від здатності моделі враховувати нелінійність, часову залежність і рівень волатильності ринкових даних. У зв'язку з цим доцільним є систематизоване порівняння основних архітектур глибоких нейронних мереж (табл. 1) з позиції їх практичного застосування у високоволатильному фінансовому середовищі, що дозволяє обґрунтувати вибір найбільш доцільних моделей для подальших експериментальних досліджень.

Для оцінювання ефективності різних архітектур глибоких нейронних мереж у задачах прогнозування фінансових ринків було проведено серію експериментальних досліджень на основі фінансових часових рядів, що характеризуються підвищеним рівнем волатильності. В якості вхідних даних використовувалися історичні котирування фінансових інструментів, попередньо нормалізовані та підготовлені з урахуванням часової структури даних.

У межах експериментів розглядалися рекурентні нейронні мережі, моделі типу LSTM та GRU, згорткові нейронні мережі, часові згорткові мережі, гібридні архітектури, а також нейронні мовні моделі для обробки текстових фінансових даних. Для забезпечення коректності порівняння моделі навчалися за однакових умов, з використанням спільних наборів даних та єдиних критеріїв оцінювання якості прогнозів, таких як MSE, MAE та R^2 .

Результати експериментів показали, що класичні рекурентні архітектури демонструють обмежену ефективність у середовищах з високою волатильністю, особливо при збільшенні горизонту прогнозування. Моделі LSTM та GRU забезпечують більш стабільні результати завдяки здатності враховувати довгострокові часові залежності, проте їх обчислювальна складність зростає зі збільшенням розміру вхідних даних.

Часові згорткові мережі та гібридні архітектури продемонстрували найвищу стійкість до шуму та запізнювання прогнозів, що є критичним фактором у задачах фінансового прогнозування за умов високої волатильності. Застосування механізмів уваги та нейронних мовних моделей додатково підвищує адаптивність моделей до змін ринкового середовища, дозволяючи акцентувати увагу на найбільш інформативних фрагментах числових та текстових даних.

Таблиця 1

Порівняльна характеристика архітектур глибоких нейронних мереж у задачах прогнозування фінансових ринків

Архітектура нейронної мережі	Основні особливості	Переваги у фінансовому прогнозуванні	Обмеження та недоліки	Доцільність застосування за високої волатильності
RNN	Послідовна обробка часових даних, наявність зворотних зв'язків	Моделювання короткострокових залежностей	Проблема зникання градієнта, нестійкість при довгих послідовностях	Обмежена
LSTM	Комірки пам'яті та керування інформаційними потоками	Захоплення довгострокових залежностей, стабільніші прогнози	Висока обчислювальна складність	Висока
GRU	Спрощена структура порівняно з LSTM	Швидше навчання при збереженні точності	Менша гнучкість у складних залежностях	Висока
CNN	Згорткові фільтри для виділення локальних патернів	Ефективне вилучення ознак, стійкість до шуму	Обмежене врахування довготривалої часової залежності	Середня
TCN	Причинні згортки та дилатація	Відсутність проблеми відставання, паралельна обробка	Потреба ретельного налаштування гіперпараметрів	Висока
Гібридні моделі	Комбінація CNN, RNN, LSTM, TCN	Поєднання переваг різних архітектур	Ускладнена реалізація та інтерпретація	Дуже висока
Моделі з механізмом уваги	Динамічне виділення значущих фрагментів	Адаптація до зміни ринкових умов	Значні вимоги до даних і ресурсів	Дуже висока
Нейронні мовні моделі	Контекстний аналіз текстових даних, трансформери	Інтеграція інформаційного потоку новин, оцінка тональності, швидка реакція на події	Високі обчислювальні витрати, потреба в великих датасетах	Висока

Отримані експериментальні результати підтверджують доцільність обґрунтованого вибору архітектур глибоких нейронних мереж залежно від характеристик фінансових даних та умов ринку, а також свідчать про перспективність використання гібридних та увагою керованих моделей у практичних інформаційних технологіях прогнозування фінансових ринків.

ВИСНОВКИ З ДАНОГО ДОСЛІДЖЕННЯ

І ПЕРСПЕКТИВИ ПОДАЛЬШИХ РОЗВІДОК У ДАНОМУ НАПРЯМІ

У статті розглянуто сучасні підходи до застосування глибоких нейронних мереж у задачах прогнозування фінансових ринків за умов високої волатильності. Показано, що ефективність прогнозування фінансових часових рядів значною мірою залежить від обґрунтованого вибору архітектури нейронної мережі з урахуванням нелінійності, часової залежності та зашумленості ринкових даних.

На основі проведеного узагальнення та систематизованого порівняння архітектур глибоких нейронних мереж визначено їхні основні переваги й обмеження у контексті практичного застосування у високоволатильному фінансовому середовищі. Запропонована зведена характеристика архітектур дозволяє формалізувати процес вибору моделей для побудови інформаційних технологій прогнозування фінансових ринків.

Результати експериментальних досліджень підтверджують, що рекурентні та згорткові моделі забезпечують різний рівень точності прогнозування залежно від горизонту прогнозу та характеристик ринкових даних. Найбільш стійкі результати продемонстрували часові згорткові та гібридні архітектури, які поєднують здатність до моделювання складних часових залежностей і підвищену стійкість до шуму та проблеми відставання прогнозів.

Отримані результати свідчать про доцільність використання гібридних моделей та моделей з механізмами уваги у задачах прогнозування фінансових ринків за умов високої волатильності. Подальші дослідження доцільно спрямувати на розширення експериментальної бази, оптимізацію гіперпараметрів моделей та інтеграцію розглянутих архітектур у прикладні інформаційні системи фінансового аналізу.

Література

1. Ferrara, M., Ciano, T., Nava, C.R. and Cananà, L., 2023. Multi-criteria decision analysis: Hesitant fuzzy methodology towards expert systems for analyzing financial markets dynamics. *Soft Computing*, pp.1-16.
2. Santur, Y., 2023. A novel financial forecasting approach using deep learning framework. *Computational Economics*, 62(3), pp.1341-1392. <https://doi.org/10.1007/s10614-024-10587-4>
3. Bibinger, M., 2024. Probabilistic models and statistics for electronic financial markets in the digital age. *Jahresbericht der Deutschen Mathematiker-Vereinigung*, pp.1-37.

4. Orlando, G., Bufalo, M. and Stoop, R., 2022. Financial markets' deterministic aspects modeled by a low-dimensional equation. *Scientific reports*, 12(1), p.1693. <https://doi.org/10.1038/s41598-022-05765-z>
5. Santoro, D. and Grilli, L., 2022. Generative Adversarial Network to evaluate quantity of information in financial markets. *Neural Computing and Applications*, 34(20), pp.17473-17490. <https://doi.org/10.1007/s00521-022-07401-3>
6. Khattak, B.H.A., Shafi, I., Rashid, C.H., Safran, M., Alfarhood, S. and Ashraf, I., 2024. Profitability trend prediction in crypto financial markets using Fibonacci technical indicator and hybrid CNN model. *Journal of Big Data*, 11(1), p.58.
7. Yilmaz, F.M. and Yildiztepe, E., 2024. Statistical evaluation of deep learning models for stock return forecasting. *Computational Economics*, 63(1), pp.221-244. <https://doi.org/10.1177/00368504241275>
8. Liu, B. and Lai, M., 2024. Advanced Machine Learning for Financial Markets: A PCA-GRU-LSTM Approach. *Journal of the Knowledge Economy*, pp.1-35.
9. Zhan, Z. and Kim, S.K., 2024. Versatile time-window sliding machine learning techniques for stock market forecasting. *Artificial Intelligence Review*, 57(8), p.209. <https://doi.org/10.1007/s10462-024-10851-x>
10. Grilli, L. and Santoro, D., 2022. Forecasting financial time series with Boltzmann entropy through neural networks. *Computational Management Science*, 19(4), pp.665-681. <https://doi.org/10.1007/s10287-022-00430-2>
11. Sahiner, M., 2022. Forecasting volatility in Asian financial markets: evidence from recursive and rolling window methods. *SN Business & Economics*, 2(10), p.157. <https://doi.org/10.1007/s43546-022-00329-9>
12. Manjunath, S. and Halasuru Manjunath, P., 2022, February. A Novel Approach for Financial Markets Forecasting Using Deep Learning with Long Short Term Networks. In *International Conference on Remote Engineering and Virtual Instrumentation* (pp. 456-462). Cham: Springer International Publishing. <https://doi.org/10.30699/IJF.2022.313164.1286>
13. Zhu, M., Song, Y. and Zheng, X., 2024. Volatility Dynamics and Mixed Jump-GARCH Model Based Jump Detection in Financial Markets. *Computational Economics*, pp.1-33. <https://doi.org/10.3390/math13030347>
14. Chen, C.W. and Chien, C.T., 2024. Improving Quantile Forecasts via Realized Double Hysteretic GARCH Model in Stock Markets. *Computational Economics*, pp.1-25.
15. Pokou, F., Sadefo Kamdem, J. and Benhmad, F., 2024. Hybridization of ARIMA with learning models for forecasting of stock market time series. *Computational Economics*, 63(4), pp.1349-1399. <https://doi.org/10.1007/s10614-023-10499-9>
16. Orra, A., Sahoo, K. and Choudhary, H., 2023. Machine Learning-Based Hybrid Models for Trend Forecasting in Financial Instruments. In *Soft Computing for Problem Solving: Proceedings of the SocProS 2022* (pp. 337-353). Singapore: Springer Nature Singapore. https://doi.org/10.1007/978-981-19-6525-8_26
17. Razletovskaia, V.V. and Stepanov, I.M., 2023. Coordination of the Development of Financial and Intellectual Technologies in the Global and National Markets. In *Anti-Crisis Approach to the Provision of the Environmental Sustainability of Economy* (pp. 235-244). Singapore: Springer Nature Singapore.
18. Fen, H., 2023, July. Application of Machine Learning Algorithm in Risk Prediction of Financial Markets. In *International Conference on Frontier Computing* (pp. 463-470). Singapore: Springer Nature Singapore.

References

1. Ferrara, M., Ciano, T., Nava, C.R. and Cananà, L., 2023. Multi-criteria decision analysis: Hesitant fuzzy methodology towards expert systems for analyzing financial markets dynamics. *Soft Computing*, pp.1-16.
2. Santur, Y., 2023. A novel financial forecasting approach using deep learning framework. *Computational Economics*, 62(3), pp.1341-1392. <https://doi.org/10.1007/s10614-024-10587-4>
3. Bibinger, M., 2024. Probabilistic models and statistics for electronic financial markets in the digital age. *Jahresbericht der Deutschen Mathematiker-Vereinigung*, pp.1-37.
4. Orlando, G., Bufalo, M. and Stoop, R., 2022. Financial markets' deterministic aspects modeled by a low-dimensional equation. *Scientific reports*, 12(1), p.1693. <https://doi.org/10.1038/s41598-022-05765-z>
5. Santoro, D. and Grilli, L., 2022. Generative Adversarial Network to evaluate quantity of information in financial markets. *Neural Computing and Applications*, 34(20), pp.17473-17490. <https://doi.org/10.1007/s00521-022-07401-3>
6. Khattak, B.H.A., Shafi, I., Rashid, C.H., Safran, M., Alfarhood, S. and Ashraf, I., 2024. Profitability trend prediction in crypto financial markets using Fibonacci technical indicator and hybrid CNN model. *Journal of Big Data*, 11(1), p.58.
7. Yilmaz, F.M. and Yildiztepe, E., 2024. Statistical evaluation of deep learning models for stock return forecasting. *Computational Economics*, 63(1), pp.221-244. <https://doi.org/10.1177/00368504241275>
8. Liu, B. and Lai, M., 2024. Advanced Machine Learning for Financial Markets: A PCA-GRU-LSTM Approach. *Journal of the Knowledge Economy*, pp.1-35.
9. Zhan, Z. and Kim, S.K., 2024. Versatile time-window sliding machine learning techniques for stock market forecasting. *Artificial Intelligence Review*, 57(8), p.209. <https://doi.org/10.1007/s10462-024-10851-x>
10. Grilli, L. and Santoro, D., 2022. Forecasting financial time series with Boltzmann entropy through neural networks. *Computational Management Science*, 19(4), pp.665-681. <https://doi.org/10.1007/s10287-022-00430-2>
11. Sahiner, M., 2022. Forecasting volatility in Asian financial markets: evidence from recursive and rolling window methods. *SN Business & Economics*, 2(10), p.157. <https://doi.org/10.1007/s43546-022-00329-9>

5. Manjunath, S. and Halasuru Manjunath, P., 2022, February. A Novel Approach for Financial Markets Forecasting Using Deep Learning with Long Short Term Networks. In International Conference on Remote Engineering and Virtual Instrumentation (pp. 456-462). Cham: Springer International Publishing. <https://doi.org/10.30699/IJF.2022.313164.1286>
6. 13. Zhu, M., Song, Y. and Zheng, X., 2024. Volatility Dynamics and Mixed Jump-GARCH Model Based Jump Detection in Financial Markets. Computational Economics, pp.1-33. <https://doi.org/10.3390/math13030347>
7. Chen, C.W. and Chien, C.T., 2024. Improving Quantile Forecasts via Realized Double Hysteretic GARCH Model in Stock Markets. Computational Economics, pp.1-25.
8. 15. Pokou, F., Sadefo Kamdem, J. and Benhmad, F., 2024. Hybridization of ARIMA with learning models for forecasting of stock market time series. Computational Economics, 63(4), pp.1349-1399. <https://doi.org/10.1007/s10614-023-10499-9>
9. Orra, A., Sahoo, K. and Choudhary, H., 2023. Machine Learning-Based Hybrid Models for Trend Forecasting in Financial Instruments. In Soft Computing for Problem Solving: Proceedings of the SocProS 2022 (pp. 337-353). Singapore: Springer Nature Singapore. https://doi.org/10.1007/978-981-19-6525-8_26
10. Razletovskaia, V.V. and Stepnov, I.M., 2023. Coordination of the Development of Financial and Intellectual Technologies in the Global and National Markets. In Anti-Crisis Approach to the Provision of the Environmental Sustainability of Economy (pp. 235-244). Singapore: Springer Nature Singapore.
11. Fen, H., 2023, July. Application of Machine Learning Algorithm in Risk Prediction of Financial Markets. In International Conference on Frontier Computing (pp. 463-470). Singapore: Springer Nature Singapore.