

**МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
КИЇВСЬКИЙ НАЦІОНАЛЬНИЙ ЕКОНОМІЧНИЙ УНІВЕРСИТЕТ
ІМЕНІ ВАДИМА ГЕТЬМАНА**

**Навчально-науковий інститут
Інститут інформаційних технологій в економіці**

Кафедра математичного моделювання та статистики

Освітньо-професійна програма «Економічна кібернетика та ДатаСайнс»

Галузь знань 05 «Соціальні та поведінкові науки»
Спеціальність 051 «Економіка»

Форма навчання: очна (денна)

КВАЛІФІКАЦІЙНА МАГІСТЕРСЬКА РОБОТА

на тему: **«ПРОГНОЗУВАННЯ ЦІН НА АКЦІЇ МЕТОДАМИ ГЛИБОКОГО
НАВЧАННЯ»**

Здобувача Зубової Марини Сергіївни

Науковий керівник: Олена ЮНЬКОВА

**Робота допущена до захисту перед екзаменаційною комісією
з атестації здобувачів вищої освіти (ЕК)**

Завідувач кафедри:

Київ 2024

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
КИЇВСЬКИЙ НАЦІОНАЛЬНИЙ ЕКОНОМІЧНИЙ УНІВЕРСИТЕТ
ІМЕНІ ВАДИМА ГЕТЬМАНА
Навчально-науковий інститут
«Інститут інформаційних технологій в економіці»
Кафедра математичного моделювання та статистики

**ОСВІТНЬО-ПРОФЕСІЙНА
ПРОГРАМА**

**ЕКОНОМІЧНА КІБЕРНЕТИКА та
ДАТА САЙНС**

Галузь знань 05 соціальні та поведінкові науки

Спеціальність 051 економіка

ПОГОДЖЕНО

Керівник проектної групи (гарант)
освітньо-професійної програми
О.В.Піскунова

ЗАТВЕРДЖУЮ

Завідувач кафедри математичного
моделювання та статистики
Г.І.Великоіваненко

(підпис)

20__ р.

(підпис)

20__ р.

ІНДИВІДУАЛЬНЕ ЗАВДАННЯ

здобувачу вищої освіти Зубовій Марині Сергіївні

очної (денної) форми навчання

на підготовку кваліфікаційної магістерської роботи

на тему «Прогнозування цін на акції методами глибокого навчання»

Тему затверджено наказом ректора Університету від "19" вересня 2024 р. №
СТ

Кваліфікаційна магістерська робота виконується на матеріалах сучасних
інтернет-видань, підручників та датасету, отриманого з сайту [Yahoo Finance](https://finance.yahoo.com/)

План кваліфікаційної магістерської роботи

Розділ 1	Теоретичні аспекти прогнозування динаміки фінансових ринків
Розділ 2	Застосування глибокого навчання у фінансовому прогнозуванні
Розділ 3	Практична реалізація прогнозних моделей для цін на акції на основі глибокого навчання

Об'єкт дослідження:	Ринок акцій міжнародних корпорацій, зокрема, їх фінансові показники та динаміка котирувань в умовах зовнішніх шоків
Предмет дослідження:	Методи прогнозування цін на акції, що застосовуються для оцінювання майбутніх трендів на фінансовому ринку
Мета кваліфікаційної бакалаврської роботи:	Прогнозування цін на акції компанії Apple за допомогою методів глибокого навчання та аналіз фінансових показників і тенденцій ринку акцій міжнародних корпорацій в умовах економічної та політичної нестабільності, глобальних криз і економічних шоків

Конкретні завдання, які здобувач повинен виконати для досягнення поставленої мети:

У розділі 1	провести аналіз наукових джерел, присвячених прогнозуванню динаміки фінансового ринку для визначення основних підходів і методів дослідження та вибрати найбільш релевантні економіко-математичні методи для аналізу та прогнозування ціни акцій компанії Apple;
У розділі 2	провести аналіз вхідних даних для побудови моделі; розробити та реалізувати математичні моделі прогнозування рівня цін на акції
У розділі 3	здійснити практичну апробацію моделей та порівняти їх точність і ефективність; розробити рекомендації щодо застосування моделей на основі отриманих результатів

Завдання підготував
науковий керівник

(підпис)

Олена ЮНЬКОВА

(ініціали, прізвище)

«____» _____ 20__ р.

Завдання одержав
здобувач

(підпис)

Марина ЗУБОВА

(ініціали, прізвище)

«____» _____ 20__ р.

Реферат

Тема дипломної роботи: *Прогнозування цін на акції методами глибокого навчання*

Актуальність теми зумовлена необхідністю розробки більш точних і ефективних інструментів для аналізу фінансових ринків. Сучасні фінансові ринки характеризуються високою динамічністю, великою кількістю змінних та складними взаємозв'язками, що ускладнює прогнозування цін на акції за допомогою традиційних статистичних методів, таких як лінійна регресія та ARIMA. Обмеження цих методів полягають у їх здатності працювати тільки з лінійними залежностями, що є недостатнім для аналізу складних фінансових часових рядів. З огляду на це, методи глибокого навчання, зокрема рекурентні нейронні мережі (RNN), Long Short-Term Memory (LSTM) та їх гібридні архітектури, набувають особливої популярності завдяки здатності ефективно обробляти великі обсяги даних та враховувати часові залежності, що робить їх надзвичайно перспективними для прогнозування цін на акції.

Метою даного дослідження є розробка та оцінка ефективності методів прогнозування цін на акції з використанням сучасних алгоритмів глибокого навчання. В межах цієї роботи аналізуються можливості застосування LSTM, CNN та гібридних моделей, для прогнозування фінансових часових рядів, а також проводиться порівняння з традиційними статистичними підходами, такими як ARIMA і експоненційне згладжування.

Завданнями дослідження є: аналіз фінансових показників акцій на основі історичних даних, порівняння ефективності моделей прогнозування, розробка та застосування методів очищення та підготовки даних, а також вибір і використання метрик для оцінки точності прогнозів. В рамках роботи використано методи очищення даних, нормалізації, видалення аномалій, а також метрики, такі як MAE, MAPE та RMSE, для оцінки якості моделей.

Результати експерименту показали, що методи глибокого навчання, зокрема LSTM та гібридні моделі, демонструють високу точність прогнозування цін на акції в умовах високої волатильності ринку та складних нелінійних залежностей. Вони

перевершують класичні статистичні моделі, такі як ARIMA, за точністю та стабільністю результатів, що робить їх перспективними для використання в реальних фінансових стратегіях.

Практичне значення роботи полягає в розробці ефективних інструментів для прогнозування цін на акції, що може бути корисно для інвесторів, фінансових аналітиків та інших учасників ринку, які працюють в умовах невизначеності та нестабільності.

Ключові слова: прогнозування цін на акції, глибоке навчання, рекурентні нейронні мережі, LSTM, фінансові ряди, машинне навчання, ARIMA, лінійна регресія, фінансова аналітика.

Відгук

**про кваліфікаційну магістерську роботу
здобувача навчально-наукового інституту
«Інститут інформаційних технологій в економіці»
освітньо-професійної програми «Економічна кібернетика і Дата Сайнс»
Зубової Марини Сергіївни
на тему «Прогнозування цін на акції методами глибокого навчання»**

Актуальність теми: Підвищення точності прогнозування динаміки фінансових показників є актуальною задачею, адже визначення трендів і довгострокових змін на фінансовому ринку дає можливість компаніям передбачати потенційні ризики, інвесторам розробляти ефективні інвестиційні стратегії, а всім учасникам ринку мати чіткіші очікування, що зменшують рівень невизначеності на волатильних фінансових ринках.

Позитивні риси кваліфікаційної магістерської роботи: Застосовані в роботі методи глибокого навчання для прогнозування ціни на акції є науково новим підходом завдяки здатності цих методів моделювати складні, нелінійні процеси, враховувати довготривалі та короткотривалі залежності, адаптуватися до раптових ринкових змін і масштабувати аналіз. Це є особливо важливим для фінансового середовища, де точність і швидкість прогнозів є критичними для успішної економічної діяльності та управління ризиками.

Наявність самостійних розробок автора: У роботі застосовано комплексний підхід, що охоплює формування статистичної бази дослідження, застосування моделей глибокого навчання з різною архітектурою, оптимізацію гіперпараметрів моделей та їх тестування.

Цінність теоретичних висновків та практичних рекомендацій Застосування методів глибокого навчання в аналізі та прогнозуванні ціни на акції дають змогу ефективно враховувати вплив численних факторів та непередбачуваних змін на ринку. Запропоновані підходи можуть бути підґрунтям для розробки більш адаптованих і точних моделей прогнозування динаміки фінансового ринку, які дозволять знижувати ризики, підвищувати прибутковість інвестиційних стратегій, а також оптимізувати процеси управління портфелями. Робота пройшла апробацію на XII всеукраїнській науково-практичній конференції «Форум молодих економістів-кібернетиків «Моделювання економіки: проблеми, тенденції, досвід», що проходила 22-23 листопада 2024 року у Львівському національному університеті імені Івана Франка.

Наявність недоліків: В роботі варто було б глибше проаналізувати вплив параметрів гібридного методу на їхню точність.

Загальна оцінка кваліфікаційної магістерської роботи та її допущення до захисту перед ЕК: Кваліфікаційна магістерська робота Зубової М.С. демонструє високий рівень обізнаності автора в темі, здатність до аналізу та систематизації матеріалів і є самостійним науковим дослідженням. Набуті автором компетентності відповідають вимогам освітньої програми, а кваліфікаційна магістерська робота може бути рекомендована до захисту з високою оцінкою.

Науковий керівник: кандидат фізико-математичних наук,
доцент, професор кафедри математичного

моделювання та статистики

О.О. Юнькова

3MICT

[illegible]

ВСТУП

Актуальність теми. Прогнозування цін на акції є однією з ключових задач у фінансовому аналізі, що зумовлено багатьма причинами. В сучасних умовах фондовий ринок є основним механізмом залучення інвестицій, які стимулюють розвиток компаній і економік. Зростаюча глобалізація фінансових ринків посилює їхній вплив на загальний стан економіки, оскільки зміни на цих ринках можуть впливати на інші. Динаміка цін на акції відображає стан економіки і потребує надійного прогнозування: розуміння перспектив розвитку економіки допомагає ідентифікувати потенційно вигідні активи для інвестування. Інвестори та фінансові установи, які прагнуть мінімізувати ризики втрат і підвищити ефективність своїх інвестиційних рішень, зацікавлені в точних прогнозах, щоб аналізувати потенційні ризики та розробляти ефективні інвестиційні стратегії.

Інформаційна відкритість та висока швидкість передачі даних змушують учасників ринку оперативно реагувати на нову інформацію. Завдяки появі нових технологій (штучного інтелекту, машинного навчання, великих даних) аналізування та прогнозування фінансових показників дозволяє враховувати як історичні дані, так і поточні новини, соціальні настрої та інші фактори, що впливають на ринок. Сучасні методи глибокого навчання, зокрема рекурентні нейронні мережі (RNN) та Long Short-Term Memory (LSTM), здатні обробляти великі обсяги даних і враховувати часові залежності, що робить їх ефективними інструментами для прогнозування цін на акції в умовах високої волатильності. Це дає змогу знизити ризики та підвищити точність інвестиційних рішень у сучасних умовах економічної нестабільності.

Аналіз останніх досліджень і публікацій. Питання прогнозування фінансових ринків, моделювання часових рядів, а також застосування методів машинного навчання та економіки поведінки для оцінки фінансової стабільності розглядалися в численних працях зарубіжних авторів таких, як праці Фама Ю., Мандельброта Б., Гамільтона Дж., Енгла Р., Бокса Дж., Дженкінса Г., Канемана Д., Шиллера Р., Тверські А., Лея К. М., Петрова А. Б.. Зокрема, у працях Мандельброта Б., Гудмана Р. Л., Шиллера Р., Канемана Д. аналіз та прогнозування фінансових ринків розглядаються

не лише з математичних позицій, зокрема методів фрактального аналізу, а також і з психологічних - економіки поведінки для розуміння фінансової нестабільності.

Метою кваліфікаційної магістерської роботи є аналіз фінансових показників і тенденцій ринку акцій міжнародних корпорацій в умовах економічної та політичної нестабільності, глобальних криз і зовнішніх економічних шоків, прогнозування цін на акції компанії Apple за допомогою методів глибокого навчання та.

Завданнями дослідження є:

- Проведення аналізу наукових джерел, присвячених прогнозуванню динаміки фінансових ринків
- Оцінювання впливу кризових економічних умов, на динаміку цін на акції міжнародних корпорацій.
- Аналіз використання методів глибокого навчання (зокрема LSTM та гібридних моделей, CNN, ARIMA, експоненціальне згладжування) для прогнозування цін на акції в умовах нестабільних ринків.
- Вивчення ефективності застосування різних моделей машинного навчання для фінансового прогнозування в умовах високої волатильності ринку.
- Прогнозування цін на акції методами глибокого навчання.
- Розробка рекомендацій щодо оптимізації процесів інвестиційного планування та прийняття фінансових рішень.

Об'єктом дослідження є динаміка ринку акцій, зокрема, їх фінансові показники та динаміка котирувань в умовах зовнішніх шоків.

Предметом дослідження є методи прогнозування цін на акції, що застосовуються для оцінювання майбутніх трендів на фінансовому ринку.

Методи дослідження. У даній роботі використано методи прогнозування часових рядів: метод автокореляції, інтеграції та ковзного середнього ARIMA, метод експоненціального згладжування, а також методи машинного навчання: штучні нейронні мережі, зокрема рекурентні нейронні мережі (LSTM), а також комбіновані моделі, які дозволяють адаптуватися до змін в умовах фінансової та економічної нестабільності; табличний та графічний методи аналізу.

Теоретична значущість отриманих результатів полягає в розширенні знань про застосування методів глибокого навчання для прогнозування фінансових рядів, зокрема цін акцій. *Методична значущість* полягає в розробці та адаптації алгоритмів для ефективного прогнозування на основі реальних фінансових даних. *Практична значущість* — в можливості використання отриманих моделей для підвищення точності прогнозування на фінансових ринках, що може допомогти інвесторам і аналітикам у прийнятті обґрунтованих рішень.

Інформаційною базою дослідження є наукові статті, публікації в рецензованих журналах, книги та технічні звіти з області глибокого навчання, машинного навчання та фінансового аналізу, а також відкриті фінансові дані, зокрема історичні ціни акцій, котирування та макроекономічні показники. Окрім того, використано ресурси з онлайн-платформ для аналізу фінансових ринків та інструментів програмного забезпечення для обробки даних і розробки моделей прогнозування, для виконання коду використовувався Google Colab.

Структура роботи. Робота складається зі вступу, трьох розділів, висновків, списку використаних джерел та додатків.

РОЗДІЛ 1 ТЕОРЕТИЧНІ АСПЕКТИ ПРОГНОЗУВАННЯ ДИНАМІКИ ФІНАНСОВИХ РИНКІВ

ОЗВИТОК НАУКОВОЇ ДУМКИ ЩОДО ПРОГНОЗУВАННЯ ФІНАНСОВИХ ПОКАЗНИКІВ

Прогнозування фінансових ринків є ключовим інструментом для аналізу тенденцій, ухвалення інвестиційних рішень та ефективного управління активами. Це міждисциплінарний процес, що інтегрує знання з економіки, математики, статистики, психології та інформатики. Еволюція підходів до прогнозування фінансових ринків відображає не лише зростання аналітичних можливостей, але й глибше розуміння динаміки ринкових процесів.

Витоки прогнозування фінансових ринків лежать у класичних економічних теоріях, розроблених у першій половині XX століття. У цей період переважала ідея, що ринки рухаються раціонально, і ціни формуються на основі фундаментальних економічних чинників. Найважливішими підходами досліджень були:

- Теорія раціональних очікувань, згідно якої ціни активів враховують всю доступну інформацію.
- Портфельна теорія Гаррі Марковіца (1952), яка базується на ідеї диверсифікації ризику та оптимізації інвестиційного портфеля.
- Фундаментальний аналіз, в якому оцінювання компаній здійснюється через аналіз їхніх фінансових показників, таких як прибуток, активи та обсяг продажів.

У середині XX століття дослідження фінансових ринків почали орієнтуватися на застосування математичних і статистичних моделей [2, с. 45]. Ключовими ідеями цього періоду стали: фрактальний аналіз Бенуа Мандельброта (1963), гіпотеза ефективного ринку (ЕМН) Юджина Фама (1965) та моделі часових рядів.

Фрактальний аналіз Бенуа Мандельброта, який він представив у 1963 році, значною мірою змінив уявлення про поведінку фінансових ринків [3, с. 112] та концепцію ризику в економіці. До його досліджень більшість економістів вважала,

що фінансові ринки підпорядковуються законам нормального розподілу, що означає наявність симетричних коливань цін навколо середнього значення. Однак Мандельброт довів, що це припущення є хибним і не відповідає реальності фінансових ринків. Він вперше застосував концепцію фракталів до аналізу ринкових змін, показавши, що ці зміни мають властивість самоподібності. Іншими словами, коливання цін на ринках на різних часових інтервалах мають схожий характер, що порушує принципи класичних статистичних теорій.

Мандельброт запровадив ідею, що ринки мають хаотичну природу, яка не піддається традиційним методам прогнозування, таким як нормальний розподіл. Фрактальний підхід показує, що на ринках існують певні закономірності, які повторюються на різних рівнях масштабу, але вони не є лінійними або прогнозованими за допомогою звичних статистичних моделей. Це відкриття мало велике значення для розуміння ризиків на фінансових ринках, оскільки показало, що ці ризики є набагато вищими, ніж передбачали класичні економічні моделі. Мандельброт довів, що ринки можуть зазнавати різких і непередбачуваних коливань, які не пояснюються традиційними методами статистичного аналізу. Цей підхід надав нові інструменти для аналізу ризику та волатильності, що стало важливим етапом у розвитку фінансової теорії та практики.

Теорія фракталів Мандельброта також призвела до перегляду традиційних підходів до управління ризиками, оскільки вона продемонструвала, що великі фінансові кризові явища, такі як різкі падіння ринку, є набагато більш ймовірними, ніж це припускали класичні моделі, що ґрунтуються на нормальному розподілі. Це відкриття мало серйозні наслідки для фінансових інститутів і дозволило створити нові методи оцінки ризиків, які враховують хаотичну і фрактальну природу ринкових процесів.

Гіпотеза ефективного ринку (ЕМН), запропонована Юджином Фамою у 1965 році, стала основною концепцією в теорії фінансів та економіки, зокрема у сфері аналізу ринкових цін і поведінки активів. Згідно з цією гіпотезою, фінансові ринки є настільки ефективними, що ціни активів у будь-який момент часу відображають всю наявну інформацію, яка доступна інвесторам. Іншими словами, ціни на акції,

облігації та інші фінансові інструменти відразу реагують на будь-які новини або зміни в економічних умовах, отже систематично отримувати надприбутки, використовуючи цю інформацію, неможливо.

Згідно з ЕМН((Efficient Market Hypothesis) або Гіпотеза ефективного ринку), якщо ринок є ефективним, прогнозування майбутніх цін за допомогою аналізу історичної інформації (технічного аналізу) або іншої доступної інформації є неможливим. Це пов'язано з тим, що всі відомості вже враховані в поточних цінах. Іншими словами, кожен новий сигнал, що з'являється на ринку, одразу враховується в ціні активу, тому передбачення змін поточної ціни на основі минулих цін або новин не забезпечує стабільного успіху.

Гіпотеза ефективного ринку має три основні форми: слабку, середню та сильну. Слабка форма припускає, що всі минулі ціни відображаються в поточних цінах. Середня форма враховує не тільки історичні ціни, а й публічну інформацію, таку як фінансові звіти та новини. Сильна форма стверджує, що навіть інформація, доступна тільки обмеженому колу осіб (інсайдерська інформація), вже врахована в цінах активів [1, с. 89].

Хоча гіпотеза ефективного ринку стала основою для багатьох досліджень і практичних рішень у фінансовому секторі, критики стверджують, що ринки не завжди ідеально ефективні, і часто ціни можуть бути спотворені через психологічні фактори, інформаційну асиметрію або інші недосконалості ринку.

Одним з важливих напрямків у фінансовій економіці є використання моделей часових рядів для аналізу та прогнозування змін у поведінці фінансових активів [4, с. 75], зокрема цін на акції, валюту та інші інструменти. Вивчення часового ряду дає змогу виявити тенденції, цикли, сезонні коливання та інші патерни, які можуть бути корисними для прогнозування майбутніх цінових змін. В економічному аналізі широко застосовуються різноманітні моделі часових рядів, серед яких найпоширенішими є ARIMA (AutoRegressive Integrated Moving Average) та GARCH

У 1970-1980-х роках значно зросла увага до психологічних чинників, які впливають на поведінку інвесторів та динаміку фінансових ринків [8, с. 201]. Цей

період став важливим етапом у розвитку нових підходів до економічного аналізу, оскільки дослідники почали визнавати, що традиційні моделі, які ґрунтуються на раціональних рішеннях і чисто математичних обчисленнях, не можуть повною мірою пояснити поведінку ринків. Психологічні фактори, такі як емоції, стрес і соціальний вплив, почали вважатися важливими для розуміння ринкових коливань, що стало основою для розвитку нової галузі — поведінкової економіки.

Роберт Шиллер, видатний економіст, у 1981 році показав, що ціни на фінансових ринках часто суттєво відхиляються від своїх фундаментальних значень, що зумовлено впливом психологічних факторів. Він довів, що коливання цін на ринку не можна пояснити лише змінами в основних економічних показниках, таких як прибуток компаній чи макроекономічні показники [9, с. 57]. Шиллер акцентував увагу на таких психологічних явищах, як *масова ейфорія*, коли інвестори необґрунтовано завищують вартість активів під впливом позитивних очікувань, та **паніка**, яка може призвести до надмірного зниження цін через негативні настрої на ринку.

Ці висновки стали основою для розвитку *поведінкової економіки*, яка досліджує, як психічні та емоційні чинники впливають на економічні рішення [10, с. . Шиллер підкреслив, що інвестори не завжди приймають рішення на основі логіки і раціональності, а часто керуються емоціями, соціальним тиском та іншими психологічними чинниками, які можуть призводити до ірраціональних ринкових рішень. Таке пояснення коливань на фінансових ринках стало важливим доповненням до традиційних моделей, що базуються на раціональній поведінці.

Поведінковий аналіз спрямований на вивчення того, як психологічні фактори та емоції впливають на прийняття інвестиційних рішень. Цей підхід акцентує увагу на тому, що інвестори не завжди діють раціонально, і їхні дії часто обумовлені когнітивними упередженнями, емоційними реакціями та соціальними впливами. Протилежно класичній економічній теорії, що передбачає раціональну поведінку агентів, поведінковий аналіз підкреслює, що ринки можуть бути неефективними і можуть відхилятися від своєї теоретичної рівноваги через психологічні чинники.

Одним із основних принципів поведінкового аналізу є те, що емоції мають суттєвий вплив на ринок. Страх і жадібність виступають як ключові рушійні сили, які формують ринкові рішення. В умовах невизначеності інвестори часто ухвалюють рішення, керуючись емоціями, такими як страх втратити прибуток чи бажання отримати швидкий виграш. Ці емоції можуть призвести до перевищення очікувань та виникнення ринкових аномалій. Когнітивні упередження, такі як упередженість підтвердження та ефект доступності, також значною мірою впливають на ринкову поведінку. Упередженість підтвердження проявляється в тому, що інвестори схильні шукати та надавати більше ваги інформації, яка підтверджує їхні попередні погляди, ігноруючи дані, що суперечать їхнім переконанням. Ефект доступності полягає в тому, що люди надають більше значення останнім подіям чи інформації, а не враховують всю історичну картину.

Поведінковий аналіз також зосереджується на явищі, яке називається "переживанням втрат". Виявлено, що втрати мають значно сильніший емоційний вплив на інвесторів, ніж рівнозначні за величиною прибутки. Це призводить до того, що інвестори часто схильні до консервативних рішень, аби уникнути можливих втрат, навіть коли потенційна вигода була б вищою. Крім того, масова поведінка на ринку, як правило, ірраціональна. Масова ейфорія чи паніка можуть призводити до формування "цінових бульбашок", коли ціни на активи значно перевищують їхню справжню вартість, або, навпаки, до різких обвалів ринку через панічні розпродажі [20, с. 123].

Інструменти поведінкового аналізу включають індекси настроїв, які дають змогу оцінити емоційний стан ринку. Наприклад, індекс страху та жадібності дозволяє вимірювати панічні чи жадібні настрої серед інвесторів, а також відстежувати зміни в емоційному фоні. Такі опитування, як опитування споживачів або професійних трейдерів, можуть допомогти визначити загальний настрій серед ринкових учасників і оцінити ймовірність майбутніх коливань цін. Окрім того, в умовах цифровізації, поведінкові аналітики активно використовують аналіз великих даних, зокрема інформацію з соціальних мереж, новинних ресурсів та інших джерел, щоб відстежувати настрої ринку. Соціальні медіа і новини часто відображають

емоційні реакції та колективні настрої, які можуть вказувати на наближення важливих змін на ринку.

Переваги поведінкового аналізу полягають у тому, що він допомагає пояснити чимало ринкових аномалій, які не можуть бути охарактеризовані традиційними економічними моделями. Цей підхід дозволяє враховувати людський фактор, який часто ігнорується іншими методами прогнозування, і дає змогу краще розуміти нерациональні рухи на ринку, такі як виникнення фінансових бульбашок або паніки. Однак цей підхід має й деякі обмеження. Одним з основних є складність кількісного вимірювання емоцій і настроїв, що робить інтерпретацію даних і прогнозування результатів менш точними. Також, поведінковий аналіз має обмеження у передбаченні масової поведінки, оскільки колективна поведінка може бути надзвичайно непередбачуваною, і реакція мас на певні події не завжди відповідає логічним чи економічним стандартам.

Інший важливий крок у вивченні психологічного аспекту економіки зробили психологи Даніел Канеман та Амос Тверські, які в 1979 році розробили *теорію перспектив*, що змінила підхід до аналізу рішень в умовах невизначеності [11, с. 78]. Теорія перспектив пояснює, як люди сприймають ризик і приймають рішення в ситуаціях, коли немає чіткої гарантії результату. В основі цієї теорії лежить ідея, що індивіди оцінюють вигоди та втрати не за їх абсолютною величиною, а з урахуванням психологічного впливу на їхні емоції.

Згідно з цією теорією, люди схильні *перебільшувати важливість потенційних втрат*, що призводить до надмірної обережності в ситуаціях ризику (ефект втрат), водночас як *потенційні вигоди* можуть сприйматися менш вагомими [11, с. 123]. Це пояснює, чому інвестори можуть ухвалювати рішення, які є ірраціональними з економічної точки зору, наприклад, продавати активи за низькими цінами через страх втрат або ухвалювати ризиковані рішення, коли перспективи значного прибутку виглядають надто привабливими.

Теорія перспектив Канемана та Тверські також описує феномен *автореференції*, коли люди схильні більше звертати увагу на інформацію, яка підтверджує їхні існуючі переконання, і ігнорувати протилежні дані. Це може

призвести до створення *пухких бульбашок на фінансових ринках*, коли інвестори, що сприймають лише позитивні сигнали, намагаються відмовитися від будь-яких застережень і закупають активи за завищеними цінами [11, с. 145].

Отже, дослідження Канемана і Тверські дали нові інструменти для розуміння того, як психологічні та емоційні чинники впливають на економічні рішення в умовах невизначеності, особливо в умовах, коли ринок демонструє надмірну реакцію на події або інформацію. Вони продемонстрували, що поведінка інвесторів не завжди є раціональною і що для прийняття економічних рішень необхідно враховувати не тільки чисто фінансові показники, а й психологічні чинники [11, с. 162].

Сучасні стратегії прогнозування активно використовують технічні індикатори, такі як ковзні середні, RSI (індекс відносної сили), MACD (ковзний середній індикатор конвергенції/дивергенції) та японські свічкові патерни, для аналізу тенденцій на фінансових ринках. Ці індикатори є важливими інструментами для визначення потенційних точок входу та виходу на ринок. Ковзні середні, наприклад, допомагають виявити тренди, а RSI дозволяє оцінити рівень перекупленості або перепроданості активів, що може сигналізувати про можливі зміни напрямку цін [14, с. 78].

Інтеграція цих технічних індикаторів у *моделі глибокого навчання* дозволяє автоматизувати аналіз ринку, покращуючи точність прогнозів. Використання комбінованих моделей, які поєднують класичні технічні індикатори з алгоритмами машинного навчання, дає змогу розширити можливості прогнозування, зокрема шляхом автоматизації виявлення важливих патернів та трендів на ринку. Це також дозволяє зменшити суб'єктивність рішень, властиву традиційним методам технічного аналізу.

Одним із нових і перспективних напрямів у прогнозуванні фінансових ринків є *квантова економіка*, що використовує принципи квантової механіки для розуміння складності та нестабільності фінансових ринків [15, с. 112]. У цьому контексті, квантова механіка може допомогти описати такі ринкові явища, як *непередбачуваність та хаотичні коливання* на фінансових ринках, а також вирішити деякі проблеми, пов'язані з *передбачуваністю* ринкових процесів.

Принципи *квантової суперпозиції* і *квантової невизначеності* можуть бути використані для створення моделей, які дозволяють більш точно описувати взаємодії між численними факторами, що впливають на ринок. Це може стати важливим кроком до подолання обмежень традиційних методів прогнозування, які не завжди можуть врахувати складність та багатогранність фінансових ринків. Зокрема, квантові алгоритми можуть використовуватися для моделювання взаємозв'язків між активами та ринковими умовами, що дозволить отримати більш гнучкі та адаптивні прогнози [16, с. 45].

Отже, розвиток новітніх технологій, таких як машинне навчання, великі дані та квантова механіка, відкриває нові можливості для більш точного та ефективного прогнозування фінансових ринків, що може значно покращити стратегії інвестування та управління ризиками в майбутньому.

Еволюція наукової думки в прогнозуванні фінансових ринків показує, як поступове накопичення знань та технологічний прогрес дозволяють глибше зрозуміти складність ринкових процесів. Від класичних економічних теорій до сучасних методів штучного інтелекту — прогнозування залишається багатогранним процесом, що постійно вдосконалюється для задоволення потреб інвесторів та фінансових установ [17, с. 56].

Прогнозування фінансових ринків є важливою складовою сучасної економічної практики, що дозволяє ухвалювати обґрунтовані інвестиційні рішення, оцінювати ризики та визначати перспективи розвитку. Для досягнення цих цілей існує кілька концептуальних підходів, які застосовуються залежно від специфіки задачі, доступних даних та досвіду аналітика. Основними з цих підходів є фундаментальний, технічний і поведінковий аналіз. Кожен з них має свої принципи, інструменти і методи, що дозволяють по-різному трактувати і прогнозувати поведінку ринків [18, с. 34].

Фундаментальний аналіз є одним з основних підходів до прогнозування фінансових ринків. Цей метод базується на оцінюванні внутрішньої вартості активів через глибоке вивчення економічних, фінансових і галузевих показників, що дає змогу визначити реальну вартість компаній та інших фінансових інструментів.

Припущенням, на якому ґрунтується фундаментальний аналіз, є те, що ринкова ціна активу з часом має наближатися до його реальної, фундаментальної вартості. Тобто, якщо ринкова ціна активу виявляється занадто низькою або занадто високою порівняно з його внутрішньою вартістю, ринок коригує це в довгостроковій перспективі [18, с. 45].

Фундаментальний аналіз передбачає кілька основних принципів. Одним з ключових є оцінка внутрішньої вартості активів, що ґрунтується на вивченні фінансових звітів компаній, таких як баланс, звіт про прибутки та збитки, звіт про рух грошових коштів. Ці документи дозволяють виявити справжній фінансовий стан компанії і на основі цього зробити висновки про її потенціал і вартості. Іншим важливим аспектом є аналіз макроекономічних факторів, зокрема таких показників, як валовий внутрішній продукт (ВВП), рівень безробіття, процентні ставки, інфляція, обмінні курси та політична стабільність. Ці фактори можуть істотно впливати на поведінку фінансових ринків і визначати потенціал для економічного зростання або спаду. Також важливою є оцінка галузевих тенденцій, що дозволяє зрозуміти, які чинники можуть вплинути на компанії в межах конкретної індустрії, що дає можливість точно прогнозувати їхні фінансові результати.

Інструменти, що використовуються в рамках фундаментального аналізу, зокрема коефіцієнти оцінки, допомагають виявити співвідношення між ринковою ціною активу і його фундаментальними характеристиками. Наприклад, коефіцієнт Р/Е (ціна/прибуток) дозволяє визначити, чи є активи переоціненими або недооціненими, виходячи з їхньої прибутковості. Коефіцієнт Р/BV (ціна/балансова вартість) порівнює ринкову вартість компанії з її балансовою вартістю, а рентабельність капіталу (ROE) вимірює здатність компанії генерувати прибуток за рахунок власного капіталу. Крім того, методи дисконтованих грошових потоків (DCF) є важливим інструментом, який дозволяє оцінити майбутні грошові потоки компанії та їхнє дисконтоване значення для визначення реальної вартості активу [18, с. 67].

Серед переваг фундаментального аналізу можна виділити його здатність надавати довгострокову перспективу. Цей підхід дозволяє визначити, чи є активи

недооціненими або переоціненими, що важливо для прийняття рішень про довгострокові інвестиції. Крім того, фундаментальний аналіз дозволяє інтегрувати макроекономічні фактори в загальний аналіз, що робить його більш всебічним і точним. Однак, цей підхід має й певні обмеження. Одним з них є складність врахування всіх чинників, які можуть вплинути на вартість активу, оскільки на ринку існує велика кількість змінних. Крім того, фундаментальний аналіз є чутливим до помилок у прогнозах економічних показників, таких як інфляція або процентні ставки, оскільки навіть незначні зміни в цих показниках можуть значно вплинути на кінцевий результат аналізу.

Технічний аналіз є підходом, що зосереджується на вивченні історичних даних ринку, зокрема цінових рухів, обсягів торгівлі та графічних патернів, для прогнозування майбутніх тенденцій. Цей метод базується на припущенні, що всі важливі фактори, які впливають на ринок, уже враховані в ціні активу. Згідно з технічним аналізом, ринкові ціни є результатом взаємодії багатьох факторів, включаючи економічні, політичні та психологічні чинники, тому їхня поточна величина повною мірою відображає всі наявні відомості про ринок. З огляду на це, технічні аналітики вважають, що майбутні рухи цін можуть бути передбачені шляхом вивчення історичних закономірностей та графічних патернів [19, с. 45].

Одним з основних принципів технічного аналізу є теза, що ціни враховують усе. Це означає, що ринкова ціна вже відображає всю доступну інформацію, що стосується як фундаментальних факторів, так і психологічних аспектів ринкової поведінки. Технічний аналіз також передбачає, що рухи цін є трендовими, тобто ринки мають тенденцію рухатися в певному напрямку, будь то вгору, вниз або в межах бічного тренду. Оскільки ці тренди повторюються, аналіз минулих ринкових ситуацій і патернів дозволяє прогнозувати їх повторення у майбутньому.

Головним інструментом технічного аналізу є графічні інструменти, які дозволяють візуалізувати ринкові тенденції та визначати потенційні точки входу або виходу з ринку. Лінійні графіки, бар-чарти та японські свічки є основними типами графіків, що використовуються для відображення цінових змін за певний період. Крім того, існує ряд графічних патернів, таких як "голова і плечі", "подвійна

вершина" чи "трикутники", які сигналізують про потенційні зміни тренду. Ці патерни часто повторюються в ринковій динаміці, що робить їх важливим інструментом прогнозування [19, с. 78].

Технічний аналіз також використовує різноманітні технічні індикатори, що дозволяють більш точно визначати стан ринку та його майбутні тенденції. Серед найбільш популярних індикаторів — ковзні середні, зокрема прості (SMA) та експоненційні (EMA), що допомагають згладжувати коливання цін і визначати загальні тренди. Індикатор відносної сили (RSI) вимірює величину зміни ціни в певний період, що дає змогу оцінити, чи є актив перекупленим або перепроданим. Індикатор MACD (збіжність/розбіжність ковзних середніх) дозволяє виявити потенційні зміни тренду на основі аналізу різниці між короткостроковими та довгостроковими ковзними середніми.

Рівні підтримки та опору є ще одним важливим інструментом у технічному аналізі. Ці рівні визначають цінові рівні, на яких актив має тенденцію зупинитися або змінювати свій напрямок. Рівень підтримки — це точка, де попит є достатньо сильним, щоб зупинити падіння ціни, тоді як рівень опору визначає ціну, на якій пропозиція може бути достатньою для того, щоб зупинити зростання [20, с. 95].

Однією з основних переваг технічного аналізу є його простота у використанні для короткострокових прогнозів. Візуалізація ринкових тенденцій через графіки дає змогу швидко і ефективно приймати рішення про вхід чи вихід з ринку, що робить цей підхід популярним серед трейдерів, які працюють на короткострокових часових горизонтах. Однак, технічний аналіз має свої обмеження. Одним із них є ігнорування фундаментальних факторів, таких як зміни в економічній політиці чи фінансових показниках компанії, що може призвести до недооцінки важливих змін, які впливають на ринок. Крім того, передбачення довгострокових змін на основі технічних даних може бути складним, оскільки ринкові патерни, які спостерігаються в минулому, не завжди повторюються в майбутньому, що робить цей метод менш ефективним для прогнозування змін на великих часових горизонтах.

З розвитком комп'ютерних технологій, зокрема в галузі обробки великих даних та машинного навчання, прогнозування фінансових ринків зазнало значних

змін. Технологічні досягнення у цих сферах дозволили створити нові підходи до аналізу фінансових ринків, що враховують їхню складність, динамічність і численні фактори, що впливають на цінову поведінку. Сучасні методи прогнозування зосереджуються на застосуванні складних алгоритмів і аналізі великих обсягів даних, що дає можливість отримати більш точні прогнози та ефективно оцінювати ринкові ризики.

Машинне навчання та штучний інтелект (ШІ) стали важливими інструментами для прогнозування фінансових ринків. Алгоритми машинного навчання, такі як нейронні мережі, зокрема довготривалі короткочасні пам'яті (LSTM) та згорткові нейронні мережі (CNN), дають змогу враховувати складну структуру фінансових даних та ідентифікувати приховані закономірності, які можуть бути неочевидними для традиційних методів аналізу. LSTM-мережі, наприклад, здатні працювати з часовими рядами, де важлива інформація може бути відстрочена в часі, що дозволяє моделювати залежності між даними, які виникають в різні моменти часу [12, с. 162].

Застосування *великих даних (Big Data)* стало ще однією революцією у фінансовому прогнозуванні. Тепер аналіз не обмежується лише традиційними економічними показниками, а включає альтернативні джерела інформації, такі як соціальні мережі, новини або публікації в ЗМІ. Використання цих джерел дозволяє отримати додаткові інсайти, які можуть впливати на ринкові настрої та поведінку інвесторів, що є важливим для прогнозування короткострокових коливань на ринку [13, с. 45].

традиційні статистичні методи прогнозування цін

Традиційні статистичні методи є основою прогнозування фінансових часових рядів і дозволяють отримувати достовірні результати при аналізі лінійних залежностей. Вони спираються на математичні моделі, які описують залежності між

змінними або динаміку часових рядів. Розглянемо три ключові методи: *ARIMA*, *GARCH* та лінійну регресію.

Модель *ARIMA* є однією з найпоширеніших моделей для прогнозування стаціонарних часових рядів. Вона базується на трьох компонентах: авторегресії (AR), інтеграції (I) та ковзного середнього (MA). Модель *ARIMA* дозволяє аналізувати залежності між поточним та попередніми значеннями рівнів ряду [5, с. 123], що дає змогу прогнозувати майбутні значення на основі минулих спостережень. Вона особливо корисна для прогнозування лінійних трендів та сезонних коливань у цінових даних.

Компонент AR відображає залежність поточного значення ряду від попередніх значень цього ж ряду, що дозволяє врахувати часові кореляції в даних. Інтеграція (I) має на меті перетворити нестационарний ряд у стаціонарний, тобто зробити його статистичні властивості сталими з часом, що є необхідним для точного прогнозування. Компонент MA моделює залежність поточного значення ряду від попередніх залишків або помилок, що виникають під час прогнозування, таким чином усуваючи вплив випадкових коливань і шуму. Завдяки своїй здатності поєднувати ці три компоненти, модель *ARIMA* є потужним інструментом для прогнозування, особливо у випадках, коли необхідно врахувати як історичні дані, так і випадкові зміни в часовому ряді [21, с. 145].

Модель *ARIMA* має кілька значних переваг, що робить її ефективним інструментом для аналізу та прогнозування часових рядів. Однією з головних переваг є її простота інтерпретації, оскільки результати моделі можна легко пояснити через взаємозв'язок між поточними значеннями, попередніми спостереженнями та випадковими коливаннями (шумом). Це дозволяє аналітикам чітко розуміти, як історичні дані впливають на прогноз. Також *ARIMA* є ефективною для короткострокового прогнозування, особливо у випадках, коли дані є стаціонарними або майже стаціонарними, що забезпечує високу точність прогнозів у таких умовах. Модель добре працює з лінійними часовими рядами, що мають чітко виражені тренди, що часто характерно для фінансових даних [22, с. 156].

Проте, як і будь-яка інша модель, ARIMA має певні обмеження. Одним з основних її недоліків є низька ефективність при роботі з нелінійними даними, оскільки модель не враховує складні нелінійні залежності. Це означає, що для даних з нелінійними взаємозв'язками ARIMA може не дати точних результатів. Також ARIMA є чутливою до викидів, і аномальні значення можуть значно спотворювати прогнози, знижуючи точність моделі. Окрім того, одна з вимог до моделі — це стаціонарність даних, тому перед її застосуванням необхідно трансформувати нестаціонарні ряди, що може вимагати додаткових маніпуляцій з даними.

Приклад застосування моделі ARIMA може бути розглянутий на прикладі прогнозування цін акцій компанії Apple. Першим кроком є застосування диференціювання для того, щоб перетворити початковий нестаціонарний ряд у стаціонарний. Далі, за допомогою графіків автокореляції та часткової автокореляції визначається порядок моделі (p , d , q). Після цього на основі отриманої моделі будується прогноз, який дозволяє оцінити ймовірні зміни цін у майбутньому.

Модель GARCH (Generalized Autoregressive Conditional Heteroskedasticity) використовується для моделювання волатильності фінансових часових рядів, таких як ціни акцій або валютні курси. Вона базується на припущенні, що волатильність змінюється з часом і має тенденцію до "кластеризації", тобто періоди високої волатильності зазвичай чергуються з періодами низької волатильності [6, с. 98]. Основною перевагою цієї моделі є її здатність точно відображати змінність волатильності, що важливо для прогнозування та оцінювання ризиків на фінансових ринках. Модель GARCH (1,1), яка є найпоширенішою формою цієї моделі, дозволяє ефективно прогнозувати кластеризацію волатильності, що є корисним для оцінки ризиків, особливо під час періодів ринкової нестабільності. Вона також має високу ефективність при аналізі ризиків, що робить її популярним інструментом дослідження у фінансових установах для оцінювання вартості під ризиком (VaR) [23, с. 182].

Перевагою GARCH є її здатність адаптуватися до мінливих умов на ринку, що дозволяє моделі ідеально описувати характеристики фінансових рядів, такі як "важкі хвости" та автокореляцію квадратів залишків. Однак ця модель має деякі обмеження,

зокрема її складність порівняно з іншими моделями, такими як ARIMA. Крім того, GARCH не враховує нелінійні залежності між змінними, що може обмежувати її застосування в умовах, де такі залежності мають значення. Прикладом застосування GARCH є прогнозування волатильності індексу S&P 500. Для цього будується модель GARCH на основі історичних даних, прогнозується умовна дисперсія на майбутні періоди, а отримані результати використовуються для оцінки ризику портфеля інвестора, що дозволяє приймати більш обґрунтовані інвестиційні рішення [24, с. 193].

Використання моделей часових рядів, таких як ARIMA та GARCH, дає змогу економістам і фінансовим аналітикам не тільки оцінювати ймовірні напрямки змінювання цін, але й краще розуміти природу волатильності та ризиків, характерних для фінансових ринків [7, с. 145]

Лінійна регресія є одним з основних статистичних методів, що відображає кореляційну залежність між змінними. У фінансовому контексті цей метод часто застосовується для прогнозування цін активів, аналізу трендів та оцінки впливу зовнішніх факторів на вартість фінансових інструментів. Класичне рівняння лінійної регресії, яке описує залежність між залежною змінною (наприклад, ціною акції) і незалежними змінними (наприклад, економічними показниками), дозволяє отримати кількісні оцінки впливу різних факторів на ринок [25, с. 112].

Одним з основних напрямків застосування лінійної регресії у фінансах є прогнозування трендів цін. Наприклад, регресійний аналіз може бути використаний для оцінювання впливу таких економічних показників, як рівень інфляції чи процентні ставки на динаміку фондових індексів. Крім того, лінійна регресія є корисною для аналізу чутливості цін активів до змін ключових факторів, що може допомогти інвесторам оцінити ризики та вигоди. Також цей метод дозволяє виявити взаємозв'язки між різними фінансовими змінними, такими як кореляція між обсягом торгів та волатильністю цін, що може допомогти у прийнятті більш обґрунтованих рішень.

Перевагою лінійної регресії є простота її реалізації та інтерпретації, що робить модель зручним інструментом для аналізу лінійних залежностей між змінними.

Однак, вона має й обмеження, серед яких найбільш значущими є неспроможність враховувати нелінійні зв'язки між змінними та чутливість до мультиколінеарності, тобто сильної кореляції між незалежними змінними, що може призвести до перекручених результатів. У прикладі застосування лінійної регресії для аналізу залежності ціни акції від обсягу продажів компанії, регресійна модель може бути використана для визначення коефіцієнтів, які допомагають прогнозувати змінювання ціни акції при варіаціях обсягів продажів, що є корисним для прийняття стратегічних інвестиційних рішень [26, с. 145].

Машинне та глибоке навчання у фінансовому прогнозуванні

Машинне та глибоке навчання стали важливими інструментами в сучасному фінансовому прогнозуванні завдяки своїй здатності аналізувати великі обсяги даних, виявляти складні нелінійні взаємозв'язки та створювати прогнози з високою точністю. Ці технології дозволяють набагато краще зрозуміти динамічні процеси на фінансових ринках, де традиційні статистичні методи можуть бути обмеженими. Однією з основних переваг машинного навчання є його висока адаптивність, яка дозволяє моделювати складні залежності між змінними. Це особливо важливо для фінансових ринків, де взаємозв'язки можуть бути нелінійними і змінюватися з часом. Наприклад, зміни в процентних ставках можуть впливати на валютні курси тільки за певних економічних умов, що важко описати за допомогою класичних статистичних моделей. Алгоритми, такі як дерева рішень або нейронні мережі, здатні моделювати ці складні залежності, успішно враховуючи вплив макроекономічних показників на ринки [27, с. 78].

Ще однією суттєвою перевагою машинного навчання є масштабованість. Алгоритми машинного навчання легко адаптуються до роботи з великими обсягами даних, що є ключовим аспектом у фінансовому прогнозуванні, де кожен день генерується величезна кількість інформації. Це включає в себе не лише історичні дані, але й інформацію з соціальних мереж, новини, а також ринкові транзакції.

Алгоритми, такі як Support Vector Machines (SVM) або нейронні мережі, здатні ефективно обробляти ці великі масиви даних. Для цього використовуються масштабовані платформи, такі як Apache Spark або TensorFlow, які дозволяють обробляти дані в реальному часі, забезпечуючи таким чином швидкий доступ до актуальної інформації для прогнозування [28, с. 112].

Ще однією важливою перевагою є автоматизація процесів прогнозування. Машинне навчання дозволяє мінімізувати людський вплив, що може призвести до упереджених або суб'єктивних результатів. Це особливо важливо в умовах, коли точність прогнозів є вирішальною для прийняття фінансових рішень. Алгоритми машинного навчання автоматично виявляють тренди та патерни в даних, що дозволяє значно знизити потребу в ручному втручанні. Наприклад, у фінансовому прогнозуванні можна використовувати методи обробки природної мови (NLP) для аналізу новин, що дозволяє алгоритмам автоматично визначати їхній вплив на ринки без необхідності втручання аналітика. Це підвищує ефективність та об'єктивність прогнозів [29, с. 89].

Додатково, машинне навчання дозволяє обробляти великі обсяги неструктурованих даних, таких як зображення, текст або відео. Цей аспект є особливо корисним для аналізу, наприклад, фінансових новин або графіків, що традиційно не піддаються автоматичному аналізу за допомогою класичних методів. Алгоритми машинного навчання також здатні швидко навчатися, постійно вдосконалюючи свої прогнози завдяки новим даним, що забезпечує їх здатність адаптуватися до змінюваних ринкових умов. Окрім того, алгоритми можуть працювати в режимі реального часу, що є критично важливим для високочастотної торгівлі, де необхідно приймати рішення за мілісекунди. В результаті, машинне навчання не тільки значно покращує точність прогнозів, але й дозволяє фінансовим установам здійснювати прогнози швидко і ефективно, що є основою для прийняття обґрунтованих і своєчасних рішень.

Машинне навчання знайшло широке застосування у фінансах, де використовується для вирішення різноманітних задач, таких як прогнозування цін акцій, виявлення шахрайства та розробка рекомендаційних систем. Одним із

найцікавіших і перспективних напрямків є застосування моделей глибокого навчання, зокрема Long Short-Term Memory (LSTM), для прогнозування цін акцій. LSTM є різновидом рекурентних нейронних мереж, які добре справляються з аналізом часових рядів, оскільки можуть зберігати інформацію про минулі значення протягом тривалого часу. Це дозволяє створювати прогнози, які враховують не лише поточні дані, але й історичну динаміку ринку. Моделі LSTM здатні ефективно обробляти великі обсяги даних, враховувати сезонні коливання, циклічні тренди та інші характеристики ринку, що забезпечує більш точні прогнози щодо руху цін на акції [30, с. 112].

Іншим важливим напрямком є використання алгоритмів класифікації для виявлення шахрайства. В умовах швидкого зростання обсягів фінансових транзакцій та онлайн-платежів, виявлення підозрілих операцій стало важливим завданням для фінансових установ. Алгоритми машинного навчання, такі як Logistic Regression або Random Forest, дозволяють автоматично ідентифікувати аномальні або підозрілі транзакції в реальному часі. Ці моделі здатні навчатися на великих наборах даних, виявляючи патерни, що вказують на можливі зловживання або шахрайські дії, такі як несанкціоновані платежі або спроби обману. Використання машинного навчання в цьому контексті значно підвищує ефективність систем виявлення шахрайства, знижує кількість хибнопозитивних результатів та дозволяє фінансовим установам оперативно реагувати на підозрілі активності [31, с. 98].

Ще однією важливою областю застосування машинного навчання є розробка рекомендаційних систем, які використовуються для надання індивідуальних фінансових порад клієнтам банків. Алгоритми машинного навчання аналізують поведінку клієнтів, їх фінансові звички, транзакції та інші дані, щоб запропонувати найбільш підходящі продукти або послуги. Це дозволяє банкам не лише підвищити рівень обслуговування, але й збільшити ефективність своїх маркетингових кампаній, орієнтуючи пропозиції на конкретні потреби та інтереси клієнтів. Застосування таких моделей дозволяє створювати персоналізовані рекомендації, що підвищує довіру до банківських установ і стимулює фінансову активність.

Загалом, машинне навчання відкриває нові можливості для фінансових установ, завдяки можливості працювати з великими, складними та швидко змінюваними даними. Адаптивність моделей машинного навчання дозволяє їм швидко реагувати на зміни в ринкових умовах, а масштабованість дає змогу працювати з величезними обсягами інформації. Завдяки автоматизації процесів аналізу та прогнозування, машинне навчання забезпечує більш точні та об'єктивні результати, що має велике значення для підвищення ефективності управління активами та мінімізації фінансових ризиків.

Технології машинного і глибокого навчання мають значний потенціал для застосування в фінансовому прогнозуванні, але при цьому існують обмеження, які можуть вплинути на їх ефективність і точність. Однією з основних проблем є інтерпретованість моделей. Особливо це стосується складних моделей, таких як глибокі нейронні мережі, які часто описуються як "чорні ящики". Це означає, що вони здатні надавати точні прогнози, але їхні рішення важко пояснити. У фінансовій сфері, де важлива не лише точність результату, але й розуміння причин його виникнення, така непрозорість може бути значною проблемою. Регулятори, інвестори та керівники ризиків вимагають прозорості в процесах прийняття рішень, тому пояснення того, як модель досягла своїх висновків, є критично важливим. Наприклад, якщо нейронна мережа прогнозує зростання акцій певної компанії, інвестор може бути не впевнений у правильності прогнозу, якщо не розуміє, які саме фактори вплинули на це рішення. Для покращення інтерпретованості застосовуються спеціальні методи, як LIME (Local Interpretable Model-agnostic Explanations) або SHAP (SHapley Additive exPlanations), які дозволяють оцінити вплив різних змінних на прогноз. Однак складність інтерпретації обмежує застосування таких моделей у критичних областях, таких як управління ризиками чи кредитний скоринг, де необхідна висока прозорість [32, с. 124].

Іншим важливим обмеженням є проблема перенавчання, або *overfitting*, яка виникає, коли модель надто добре пристосовується до навчальних даних, втрачаючи здатність до узагальнення на нових, невідомих даних. У фінансових ринках, де умови можуть змінюватися, перенавчена модель може демонструвати гарні результати на

історичних даних, але давати неточні прогнози на реальних ринкових даних. Це особливо небезпечно, коли фінансові рішення приймаються на основі таких недосконалих моделей, оскільки це може призвести до значних фінансових втрат. Наприклад, модель, яка навчається на історичних цінах акцій з урахуванням незначних коливань, може передбачити ціни акцій лише за схожих умов, не враховуючи можливість серйозних змін на ринку. Відомо, що складні моделі, такі як глибокі нейронні мережі з великою кількістю параметрів, особливо схильні до перенавчання, якщо даних для навчання недостатньо або вони мають надмірну варіативність. Для боротьби з перенавчанням застосовуються різні методи, такі як регуляризація (L1, L2), використання валідаційних наборів даних для перевірки результатів, спрощення архітектури моделі або застосування технік, таких як Dropout у нейронних мережах, які знижують ймовірність перенавчання [33, с. 76].

Технології машинного навчання також мають високу залежність від якості даних, з якими вони працюють. Моделі машинного навчання дуже чутливі до наявності пропущених значень, шуму та нерепрезентативності даних. У фінансовій сфері це є серйозною проблемою, оскільки точність прогнозів залежить від того, наскільки добре дані відображають реальну ситуацію на ринку. Наприклад, відсутність даних про обсяги торгів або ціни акцій у певні моменти часу може призвести до помилок у тренуванні моделі. Якщо ж у навчальному наборі даних відсутні кризові періоди, модель може не враховувати важливі ризики, пов'язані з можливими фінансовими спадами або зниженням ліквідності на ринку. Погана якість даних може значно знизити точність прогнозів, що може мати серйозні фінансові наслідки. Для вирішення цієї проблеми застосовуються методи попередньої обробки даних, такі як імпутація (відновлення пропущених значень), а також методи для виявлення та корекції аномальних або "шумних" даних. Використання великих репрезентативних наборів даних також допомагає зменшити вплив шуму і підвищити точність моделей.

Отже, хоча машинне і глибоке навчання є потужними інструментами для фінансового прогнозування, їх застосування має ряд обмежень, які необхідно враховувати для забезпечення надійності та точності результатів. Складність

інтерпретації, перенавчання та чутливість до якості даних є основними проблемами, що можуть вплинути на ефективність цих технологій у реальних фінансових умовах.

Алгоритми машинного навчання є основними інструментами для прогнозування на фінансових ринках, оскільки вони здатні обробляти великі обсяги даних, виявляти складні нелінійні взаємозв'язки та створювати точні прогнози. Особливо важливим є їх здатність адаптуватися до динамічних та непередбачуваних умов фінансових ринків, де на результат можуть впливати безліч факторів, що змінюються в часі. Серед широкого спектра алгоритмів, що використовуються у фінансових прогнозах, одними з найбільш ефективних є нейронні мережі.

Нейронні мережі (Artificial Neural Networks, ANN) є моделями, що натхненні роботою людського мозку. Вони складаються з великої кількості нейронів, які з'єднані між собою та взаємодіють один з одним, що дозволяє їм моделювати складні, нелінійні залежності в даних. Така структура дає нейронним мережам можливість навчатися на величезних наборах даних та виявляти приховані закономірності, які можуть бути неочевидними при використанні традиційних статистичних методів. Важливою особливістю нейронних мереж є їх здатність працювати з нелійними залежностями, що робить їх особливо ефективними для фінансових ринків, де динаміка може змінюватися під впливом численних, часто непередбачуваних факторів, таких як економічні кризи, зміни в політиці або глобальні події. Це дає змогу нейронним мережам прогнозувати не тільки за допомогою простих лінійних моделей, але й враховувати складні, складені взаємозв'язки між змінними [54, с. 102].

Одним з прикладів застосування нейронних мереж у фінансах є модель LSTM (Long Short-Term Memory), що є підтипом рекурентних нейронних мереж. Ці мережі особливо добре підходять для роботи з часовими рядами, оскільки мають здатність зберігати та опрацьовувати інформацію про попередні стани в серіях даних. У фінансових прогнозах LSTM використовуються для передбачення майбутніх значень таких змінних, як ціни акцій або криптовалют, аналізуючи історичні дані. Наприклад, на основі історичних цін на акції компанії, модель LSTM може прогнозувати їхню вартість на наступний день, враховуючи не тільки безпосередню залежність від попередніх значень, а й інші фактори, які можуть вплинути на ринок [34, с. 89].

Переваги нейронних мереж очевидні, оскільки вони здатні забезпечити високу точність прогнозування, особливо коли йдеться про складні та великі набори даних. Завдяки своїй гнучкості та здатності виявляти складні патерни, нейронні мережі можуть значно перевершувати традиційні моделі в аналізі фінансових ринків. Однак вони також мають певні недоліки, серед яких можна виділити високу обчислювальну складність, що вимагає значних ресурсів для тренування та використання моделей, особливо при обробці великих даних. Крім того, нейронні мережі схильні до перенавчання, якщо вони не використовують достатньо якісні або репрезентативні дані для навчання. Перенавчання означає, що модель може дуже точно прогнозувати на основі тренувальних даних, але не буде здатна давати точні прогнози для нових, невідомих даних, оскільки вона «запам'ятовує» особливості навчальних даних, а не вчиться робити узагальнення. Тому, щоб уникнути перенавчання, необхідно ретельно вибирати та обробляти дані, а також використовувати методи, що зменшують ризик перенавчання, такі як регуляризація та валідація на окремих наборах даних.

Random Forest (Случайний ліс) — це популярний ансамблевий метод машинного навчання, який базується на використанні декількох дерев рішень для підвищення точності прогнозів. У процесі роботи алгоритм створює велику кількість дерев, кожне з яких незалежно дає свій прогноз, а потім всі результати комбінуються для отримання фінального рішення. Такий підхід дозволяє мінімізувати помилки окремих моделей, оскільки кожне дерево вносить свій вклад у фінальний результат, а колективне голосування значно знижує ризик прийняття помилкових рішень. Однією з основних переваг Random Forest є його здатність ефективно працювати з великими та складними наборами даних, що містять багато змінних. Однак, для обробки великих обсягів даних та складних моделей, метод може вимагати значних обчислювальних ресурсів, що є важливим обмеженням при використанні цього алгоритму в реальних фінансових застосунках.

Одним із прикладів практичного використання Random Forest є аналіз впливу макроекономічних факторів на фінансові ринки, зокрема на ціни акцій. Алгоритм може аналізувати такі змінні, як ВВП, рівень безробіття, інфляція, та інші макроекономічні показники, щоб визначити їх значущість для прогнозу. За

допомогою Random Forest можна оцінити, які фактори мають найбільший вплив на зміни вартості акцій, а також врахувати взаємодію між різними економічними показниками. Завдяки цьому, цей метод є особливо корисним для розв'язання складних задач, пов'язаних з виявленням важливих змінних у великих наборах даних.

Незважаючи на численні переваги, такі як стійкість до шуму в даних та висока точність навіть при складних задачах, метод Random Forest має певні недоліки. Одним із основних є високі обчислювальні витрати, особливо при роботі з великими наборами даних. Оскільки цей метод створює безліч дерев і комбінує їх результати, процес навчання може займати значний час і вимагати великої кількості обчислювальних ресурсів. Крім того, Random Forest може бути менш ефективним, коли маємо справу з дуже великими наборами даних, де присутня велика кількість змінних. У таких випадках цей метод може стати менш ефективним через необхідність обробки надмірно складних структур даних.

Іншим важливим методом є градієнтний бустинг (Gradient Boosting), який поєднує декілька слабких моделей, зазвичай дерев рішень, для створення сильної моделі. Кожне нове дерево в цьому методі будується так, щоб виправити помилки, допущені попередньою моделлю. Процес побудови моделей у градієнтному бустингу є послідовним, де кожен етап прагне мінімізувати залишкові помилки попереднього кроку, що дозволяє поступово покращувати точність прогнозів. Серед найбільш відомих реалізацій цього методу можна виокремити такі бібліотеки, як XGBoost, LightGBM і CatBoost, які активно використовуються в практичних фінансових застосунках для підвищення точності прогнозів [35, с. 89].

Одним із прикладів застосування градієнтного бустингу є прогнозування ризику дефолту за кредитами. Метод допомагає виявити основні фактори ризику, такі як рівень доходу, кредитна історія, вік позичальника, та інші змінні, що впливають на ймовірність неплатоспроможності. Градієнтний бустинг є надзвичайно ефективним для таких задач завдяки своїй високій точності та здатності адаптуватися до різних типів даних, що робить його одним із найбільш поширених методів у фінансовому секторі для оцінки кредитних ризиків.

Однак градієнтний бустинг також має свої обмеження. Одним із основних недоліків є чутливість до параметрів моделі, що може вимагати детальної налаштування та підбору гіперпараметрів для досягнення оптимальних результатів. Крім того, для великих наборів даних час тренування моделі може бути довгим, оскільки кожна нова модель поступово коригує попередні помилки, що потребує великої кількості обчислювальних ресурсів. Це може бути проблемою, якщо потрібно обробляти дані в реальному часі або працювати з дуже великими наборами інформації.

Іншим важливим методом машинного навчання є Support Vector Machines (SVM), який є потужним алгоритмом для задач класифікації та регресії. Алгоритм працює за принципом пошуку гіперплощини, що найкраще розділяє дані на різні класи або передбачає їх значення в задачах регресії. Особливістю SVM є здатність працювати з нелінійними взаємозв'язками між змінними завдяки використанню ядерних функцій, які дозволяють знаходити оптимальну гіперплощину навіть для складних, не лінійних даних.

Прикладом використання SVM у фінансовому прогнозуванні є класифікація акцій на зростаючі та падаючі, на основі технічних індикаторів, таких як MACD або RSI. Метод може бути ефективним для виявлення патернів у ринкових даних та допомагати інвесторам приймати рішення про купівлю або продаж активів. Однією з основних переваг SVM є висока точність при роботі з невеликими наборами даних, що робить його корисним інструментом для аналізу окремих випадків, коли наявні дані є обмеженими. Однак SVM має свої обмеження: він не масштабується добре для великих наборів даних і чутливий до вибору ядра та параметрів моделі, що вимагає обережного налаштування для досягнення оптимальних результатів.

Вибір алгоритму машинного навчання для фінансових прогнозів є ключовим етапом, який залежить від характеру завдання, обсягу доступних даних, а також вимог до інтерпретованості результатів. Кожен з алгоритмів має свої переваги та обмеження, що визначають їх ефективність у різних контекстах.

Нейронні мережі, зокрема рекурентні нейронні мережі (LSTM), є потужним інструментом для роботи з складними часовими рядами, такими як прогнози цін

акцій або криптовалют. Їх здатність до виявлення нелінійних залежностей у великих наборах даних дозволяє моделювати складні динамічні процеси, характерні для фінансових ринків. Проте, нейронні мережі потребують значних обчислювальних ресурсів і можуть бути важкими для інтерпретації, що є важливим фактором у фінансових застосунках, де необхідно не тільки отримати точний прогноз, але й зрозуміти логіку прийняття рішень.

Random Forest є ефективним інструментом для аналізу важливості змінних і добре справляється з роботами з даними, що містять багато шуму або пропущених значень. Цей метод є стабільним і не чутливий до оверсамплінгу, що дозволяє йому працювати з великими наборами даних, зберігаючи високу точність прогнозів. Крім того, він дає можливість оцінити вплив окремих змінних на результат, що робить його дуже корисним для задач, де необхідно з'ясувати, які фактори є найбільш важливими для прийняття фінансових рішень. Однак, цей метод може бути менш ефективним у випадках з дуже великими даними або коли кількість змінних надто велика, що потребує значних обчислювальних витрат.

Гradientний бустинг (Gradient Boosting), зокрема такі реалізації, як XGBoost, LightGBM та CatBoost, забезпечують високу точність для більшості задач, таких як прогнозування ризику дефолту або класифікація фінансових інструментів. Ці методи поєднують переваги кількох слабких моделей, поступово покращуючи їх точність на кожному етапі. Завдяки своїй адаптивності та гнучкості, gradientний бустинг є чудовим вибором для широкого кола задач у фінансовому секторі. Однак цей метод вимагає значних обчислювальних ресурсів і чутливий до налаштування параметрів, що може бути проблемою при роботі з великими наборами даних.

Метод опорних векторів (SVM) є найкращим варіантом для класифікаційних задач, особливо коли маємо справу з відносно невеликими наборами даних з чіткою структурою. SVM добре працює з нелінійними залежностями та може бути ефективним для задач, таких як класифікація акцій за допомогою технічних індикаторів. Проте, SVM має обмеження, коли мова йде про масштабованість до великих обсягів даних, оскільки цей метод не завжди ефективний при обробці дуже

великих наборів з багатьма змінними. Його чутливість до параметрів і вибору ядра також може вимагати додаткових зусиль для налаштування.

Отже, ефективне прогнозування фінансових ринків часто передбачає комбінацію різних методів машинного навчання. Наприклад, для комплексних завдань, що включають аналіз часових рядів і виявлення складних патернів, можна використовувати нейронні мережі в поєднанні з Random Forest або градієнтним бустингом для підвищення точності та стабільності результатів. Вибір конкретного алгоритму залежить від специфіки завдання, доступних даних та вимог до результатів, при цьому комбінування методів може допомогти досягти найбільш точних і надійних прогнозів.

орівняльний аналіз методів прогнозування цін на фінансові активи

У процесі оцінювання ефективності різних методів прогнозування динаміки фінансових ринків важливо враховувати низку критеріїв, що визначають їхню здатність до адаптації, точності та ефективності. Точність прогнозування є основним критерієм, оскільки визначає адекватність передбачених результатів реальним даним. В умовах фінансових ринків, де навіть незначні відхилення можуть призвести до суттєвих фінансових втрат, висока точність є критично важливою. Тому важливо використовувати такі методи, які забезпечують максимальну точність на основі доступних даних.

Обчислювальна складність алгоритму також є важливим фактором. Вона характеризує, скільки ресурсів (часу та потужності обчислювальних систем) потрібно для обробки даних і виконання передбачення. Методи з високою складністю можуть бути менш придатними для використання в реальному часі або для аналізу великих наборів даних, оскільки це вимагатиме великих витрат ресурсів. Тому важливо балансувати між точністю та ефективністю обчислень, щоб забезпечити зручність використання моделі.

Інтерпретованість результатів є одним із важливих аспектів прогнозування, особливо в контексті фінансових ринків, де аналітики і інвестори повинні зрозуміти, як модель прийшла до своїх висновків. Пояснюваність результатів має особливе значення для фінансових аналітиків, які часто мають справу з необхідністю пояснення прогнозів клієнтам або керівництву. Якщо модель надає зрозумілі і логічні обґрунтування своїх рішень, це значно підвищує її довіру і зручність у застосуванні.

Стійкість до зовнішніх збурень також є важливою характеристикою для фінансових моделей. Визначає, наскільки надійно модель працює за умов наявності шуму, викидів або інших аномалій у даних. Фінансові ринки часто схильні до несподіваних змін, і модель повинна бути здатною ефективно реагувати на ці збурення, забезпечуючи стабільні результати навіть в умовах нестабільності.

Зважаючи на ці критерії, можна здійснити порівняльний аналіз основних методів прогнозування фінансових ринків. Метод ARIMA (Авторегресія з інтеграцією та зміщенням середнього) має середню точність і низьку обчислювальну складність. Це робить його ідеальним для простих часових рядів з чіткими лінійними тенденціями, проте він не підходить для складних і нелінійних ринків. ARIMA має високу інтерпретованість, але її адаптивність до довгострокових змін є обмеженою.

Модель GARCH (Generalized Autoregressive Conditional Heteroskedasticity) забезпечує високу точність для прогнозування волатильності фінансових активів, особливо в періоди ринкової турбулентності. Однак її складність у моделюванні та недостатня інтерпретованість результатів обмежують її застосування на практиці. GARCH добре підходить для короткострокових прогнозів і для виявлення змін у рівнях ризику, проте не є найкращим варіантом для довгострокового прогнозування.

Лінійна регресія є дуже простою в реалізації та має високу інтерпретованість, що робить її корисною для вивчення залежностей між окремими змінними. Проте її точність є обмеженою, особливо коли мова йде про нелінійні зв'язки, які є поширеними на фінансових ринках.

Нейронні мережі демонструють високу точність і здатність обробляти великі набори даних, а також адаптуватися до нових умов. Однак їхня висока обчислювальна складність і низька інтерпретованість можуть бути суттєвими

перешкодами, особливо для користувачів, які потребують пояснень результатів. Нейронні мережі є найбільш потужним інструментом для обробки складних нелінійних залежностей і для роботи з великими даними.

Метод випадкових лісів (Random Forest) забезпечує високу точність, стійкість до шуму та хорошу адаптивність. Це робить його корисним для задач, де є велика кількість факторів, що впливають на результат, але цей метод може бути менш точним у випадках із дуже складними нелінійними залежностями.

Висновки, зроблені з цього аналізу, підтверджують, що вибір методу прогнозування залежить від конкретних вимог задачі та доступних ресурсів. Для задач з чіткими лінійними залежностями та високими вимогами до інтерпретованості найкраще підходять ARIMA та лінійна регресія. Для ринків з високою волатильністю і ризиками, таким як прогнози для ринкової турбулентності, GARCH є найбільш ефективним методом. Для складних, великих і нелінійних наборів даних, де необхідно враховувати численні взаємозв'язки, нейронні мережі та метод випадкових лісів демонструють найкращі результати.

Прогнозування динаміки фінансових ринків є однією з найскладніших сфер сучасної аналітики. Порівняння сучасних методів та підходів дає змогу визначити не лише їхні переваги, але й обмеження. Жоден з існуючих методів не забезпечує абсолютно точного прогнозування через складність, нелінійність та динамічність фінансових ринків. Навіть найточніші моделі не можуть врахувати всі випадкові події або аномалії, що впливають на ринок.

Комплексне використання різних підходів може значно підвищити якість прогнозів. Наприклад, інтеграція фундаментального аналізу з методами машинного навчання дозволяє врахувати різні аспекти ринкової поведінки. Зокрема, поєднання ARIMA для короткострокового прогнозування з нейронними мережами для обробки складних нелінійних залежностей може забезпечити кращі результати.

Машинне навчання має значний потенціал для розвитку методів прогнозування фінансових ринків, завдяки здатності працювати з великими обсягами даних і моделювати складні взаємозв'язки. Використання методів глибокого навчання відкриває нові можливості для аналізу патернів ринкової

поведінки та виявлення складних взаємозв'язків. Ключовим фактором для ефективного прогнозування є якість та репрезентативність даних, оскільки недоліки в даних можуть суттєво вплинути на точність прогнозів.

РОЗДІЛ 2 ЗАСТОСУВАННЯ ГЛИБОКОГО НАВЧАННЯ У ФІНАНСОВОМУ ПРОГНОЗУВАННІ

Особливості фінансових часових рядів

Фінансові часові ряди є надзвичайно складними для моделювання та прогнозування через низку специфічних характеристик, які суттєво відрізняють їх від інших типів даних. Однією з найбільш важливих особливостей фінансових даних є висока волатильність, яка відображає значні коливання цін і ринкових показників протягом короткого проміжку часу. Ці коливання можуть бути зумовлені численними факторами, такими як економічні новини, зміни в регуляції, політичні події або природні катастрофи. Висока волатильність створює додаткові труднощі для прогнозування, оскільки навіть незначні зміни в економічному чи політичному контексті можуть призвести до різких цінових коливань. Це підвищує ризики інвестування і вимагає застосування адаптивних моделей, здатних враховувати швидкі зміни в ринкових умовах.

Ще однією важливою характеристикою фінансових рядів є їхня нестаціонарність. На відміну від багатьох природних чи інженерних процесів, де статистичні характеристики, такі як середнє значення та дисперсія, стабільні або змінюються з певними закономірностями, фінансові ряди часто демонструють нестабільність цих параметрів. Параметри розподілу можуть змінюватися не тільки з часом, але й в залежності від економічних циклів, змін в ринковій ситуації чи глобальних кризових явищ. Такі ряди не відповідають традиційним статистичним вимогам, що ускладнює їхній аналіз. Це ставить перед дослідниками завдання розробки спеціальних методів, здатних враховувати зміни в структурі даних і працювати з ними, наприклад, за допомогою моделей, що дозволяють адаптуватися до зміни середовища, таких як адаптивні фільтри або нелінійні методи машинного навчання.

Особливістю фінансових рядів є також наявність значної кількості шумів. Шум у фінансових даних може бути результатом випадкових флуктуацій ринку, помилок вимірювання, а також дрібних подій, які не мають суттєвого впливу на загальну тенденцію. Ці коливання, хоча й можуть бути характерними для короткострокових періодів, не несуть корисної інформації для прогнозування, однак їхня присутність може значно спотворювати результати аналізу. Завдання аналітика в цьому контексті полягає в ефективному відокремленні шумів від справжніх трендів і закономірностей. Для цього використовуються різні методи, починаючи від фільтрації даних до більш складних підходів, таких як застосування нейронних мереж, які здатні виявляти істинні закономірності серед великої кількості непередбачуваних флуктуацій.

Фінансові ряди також мають складні нелінійні залежності, що виникають між різними ринковими факторами. Наприклад, зміни в цінах активів можуть залежати від численних економічних індикаторів, які взаємодіють між собою у нелінійний спосіб. Традиційні лінійні моделі, такі як лінійна регресія або ARIMA, не завжди здатні адекватно описати ці взаємозв'язки, оскільки вони не можуть враховувати складність і взаємодію багатьох факторів, які впливають на фінансові ринки. Для моделювання таких залежностей необхідні потужніші методи, такі як нейронні мережі, зокрема рекурентні нейромережі, що здатні враховувати часові залежності, а також глибоке навчання, яке дозволяє розпізнавати складні патерни в даних.

Нарешті, фінансові часові ряди знаходяться під значним впливом зовнішніх факторів, таких як політичні події, економічні новини, зміни в регуляторному середовищі або природні катаклізми. Ці фактори можуть радикально змінювати поведінку ринку, що додає додаткову непередбачуваність до моделювання фінансових часових рядів. Вплив таких факторів, як зміни в державній політиці, анонсування нових економічних заходів або глобальні кризи, може істотно змінити напрямок ринку, причому ці зміни часто є важко прогнозованими за допомогою традиційних підходів. Для розв'язання цієї проблеми аналітики можуть використовувати додаткові джерела інформації, такі як новини, соціальні медіа чи інші неструктуровані дані, що дозволяють більш точно враховувати зовнішні впливи

на ринок. Тому моделі, які використовують лише історичні фінансові дані, можуть бути менш ефективними без урахування цих факторів, що підвищує важливість інтеграції таких джерел у прогностичні системи.

Підготовка даних є невід'ємною частиною будь-якого процесу аналізу, особливо у фінансовому прогнозуванні, де коректність та чистота вхідних даних безпосередньо впливає на ефективність моделей машинного навчання. Першочерговим етапом підготовки фінансових даних є очищення даних. Це включає виявлення та видалення технічних артефактів, таких як помилкові записи, дублікати, або значення, що не відповідають логічним та статистичним вимогам. Технічні помилки можуть виникати внаслідок збоїв при збиранні даних або неправильних вимірювань, тому важливо здійснити ретельну перевірку цілісності та логічності послідовностей. Такі аномалії, якщо їх не усунути, можуть викривити результати подальшого аналізу та вплинути на точність прогнозування.

Нормалізація даних є важливим етапом для приведення різномірних фінансових показників до єдиного співставного масштабу. Це дозволяє уникнути ситуацій, коли одні параметри домінують через свої великі абсолютні значення, а інші, зі значно меншими величинами, не отримують належної уваги. Найпоширенішими методами нормалізації є *стандартизація*, яка передбачає перетворення даних до нульового середнього та одиничної дисперсії, мінімаксне масштабування, що відображає дані на інтервал $[0, 1]$, а також логарифмічне перетворення, яке може допомогти зменшити вплив сильно асиметричних розподілів, що часто зустрічаються в фінансових даних.

Окрім цього, важливою частиною підготовки є видалення викидів, тобто екстремальних значень, які суттєво відрізняються від основної маси даних і можуть вплинути на результати аналізу. Для виявлення таких викидів використовуються різні статистичні методи, серед яких метод міжквартильного розмаху, який дозволяє визначити межі нормальних значень, *Z-score*, що оцінює віддаленість значення від середнього, та більш складні алгоритми машинного навчання для детекції аномалій. Ці методи дозволяють фільтрувати нетипові або помилкові значення, які можуть спотворювати модель, і покращують загальну точність прогнозу.

Масштабування також є важливим етапом, коли необхідно привести різні фінансові інструменти до єдиної шкали для коректного порівняння та спільного аналізу. Це важливо, оскільки фінансові активи можуть мати різні одиниці виміру, наприклад, акції можуть бути виміряні в доларах, а обсяги торгових операцій — в кількості одиниць. Масштабування дозволяє усунути домінування змінних з великими абсолютними значеннями, забезпечуючи рівноправність всіх факторів при моделюванні та прогнозуванні.

Підготовка фінансових даних є складним багатоступеневим процесом, що вимагає не лише знань з машинного навчання, а й глибокого розуміння специфіки кожного набору даних. Кожен етап, від очищення до нормалізації та масштабування, має важливе значення для створення ефективної моделі прогнозування, яка може дати точні й надійні результати.

Для забезпечення якісної підготовки даних можна використовувати інструменти програмування, наприклад Python, Colab Notebook, який пропонує широкий набір бібліотек для обробки фінансових даних. У запропонованому коді представлено базовий процес підготовки фінансових даних, що включає нормалізацію, видалення аномалій та формування часових послідовностей. Нормалізація даних, як зазначалося раніше, дозволяє привести всі змінні до однакового масштабу, що сприяє ефективнішому навчанню моделей машинного навчання. Видалення викидів гарантує, що вхідні дані не містять суттєвих аномалій, які можуть негативно вплинути на результати моделювання. Нарешті, дані перетворюються у часові послідовності, які можуть бути використані для навчання рекурентних нейронних мереж або інших моделей, які працюють з часовими залежностями.

Основні поняття нейронних мереж та глибокого навчання

Нейронні мережі є революційною технологією штучного інтелекту, що виникла внаслідок глибокого міждисциплінарного вивчення принципів функціонування біологічних нервових систем. Витоки цієї технології сягають 1940-х років, коли Warren McCulloch та Walter Pitts уперше запропонували математичну модель штучного нейрона, яка імітувала базові принципи роботи біологічних нервових клітин.

Фундаментальна концепція нейронних мереж базується на принципі колективної обробки інформації через складну мережу взаємопов'язаних обчислювальних елементів. На відміну від традиційних алгоритмічних підходів, де кожен крок чітко прописується програмістом, нейронні мережі здатні самостійно навчатися, виявляти складні нелінійні залежності та узагальнювати отриману інформацію.

Нейрон у штучній нейронній мережі є аналогом біологічного нейрона і виступає первинним елементом обробки інформації. Кожен нейрон отримує вхідні сигнали, опрацьовує їх через певні математичні перетворення та генерує вихідний сигнал. Математична модель нейрона включає три ключові компоненти:

- вхідні синаптичні зв'язки з ваговими коефіцієнтами;
- функцію агрегації, що об'єднує вхідні сигнали;
- активаційну функцію для нелінійного перетворення результату.

Принцип роботи нейрона нагадує біологічний аналог: вхідні сигнали підсумовуються з урахуванням їхньої ваги, а потім проходять через активаційну функцію, яка визначає, наскільки "збудженим" буде нейрон.

Синапси в нейронній мережі репрезентують зв'язки між нейронами, які мають власні вагові коефіцієнти. Ці коефіцієнти визначають "силу" зв'язку між нейронами та впливають на передачу інформації. Під час навчання мережі вагові коефіцієнти динамічно коригуються, що дозволяє мережі адаптуватися до вхідних даних.

Процес налаштування вагових коефіцієнтів відбувається через механізми зворотного поширення помилки, де мережа порівнює отриманий результат з очікуваним і поступово мінімізує розбіжності.

Активаційні функції виконують критичну роль у нейронних мережах, забезпечуючи нелінійне перетворення вхідних сигналів. Вони дозволяють мережі моделювати складні нелінійні залежності та уникати обмежень лінійних моделей.

Найпоширеніші активаційні функції включають:

- Sigmoid (логістична функція)
- Гіперболічний тангенс (tanh)

Кожна з цих функцій має унікальні властивості та сфери застосування, що робить їх придатними для різних архітектур нейронних мереж.

Класична архітектура нейронної мережі базується на трьох основних типах шарів, кожен з яких виконує специфічну роль у процесі обробки даних. Вхідний шар отримує первинні дані, які передаються на наступні шари для подальшої обробки. Цей шар відповідає за прийом сигналів, які надходять від зовнішнього середовища, і їх подачу в мережу. Приховані шари знаходяться між вхідним і вихідним шарами. Вони виконують складні обчислення та трансформації даних, пропускаючи їх через нейрони, кожен з яких має певні ваги і функції активації. Чим більше прихованих шарів, тим складніші обчислення можуть бути виконані, що дозволяє мережі розв'язувати більш складні завдання. Останнім етапом є вихідний шар, який формує кінцевий результат мережі, наприклад, клас, до якого належить об'єкт, або конкретне значення в регресійній задачі. Кількість прихованих шарів та їхня конфігурація є важливими для здатності мережі до навчання та узагальнення, оскільки це визначає, наскільки глибоко і точно мережа може вивчити взаємозв'язки в даних. На рис. 2.1 зображена базова архітектура нейронної мережі.

Вхідний шар → Приховані шари → Вихідний шар

Рисунок 2.1 – Архітектура класичної нейронної мережі.

Джерело: розроблено автором на основі [43]

Принципи навчання нейронних мереж зводяться до ітеративного процесу, в якому мережа постійно коригує свої параметри для мінімізації похибки між передбаченими та реальними результатами. Похибка визначає, наскільки правильно мережа передбачає результат, і навчання полягає в тому, щоб ця похибка ставала все меншою з кожним кроком. Навчання нейронних мереж здійснюється через три основні парадигми: навчання з учителем, навчання без учителя та навчання з підкріпленням.

Навчання з учителем є найбільш поширеним методом, при якому нейронна мережа навчається на основі вхідних даних та відповідних правильних відповідей. Мережа отримує приклади, які містять і вхідні дані, і правильні результати, і намагається оптимізувати свої параметри так, щоб помилка між її передбаченням і реальним результатом була мінімальною. Навчання з учителем є ефективним для задач класифікації та регресії, де є чітко визначені правильні відповіді.

Навчання без учителя відрізняється тим, що нейронна мережа не має доступу до правильних відповідей, і її завдання полягає в тому, щоб знайти структуру чи шаблони в даних без попередньої інформації про те, які результати є правильними. Цей метод використовується в задачах, де неможливо отримати мітки для всіх даних, наприклад, для кластеризації чи зниження розмірності.

Навчання з підкріпленням полягає в тому, що мережа вчиться на основі взаємодії з середовищем. Мережа отримує винагороду чи покарання залежно від того, наскільки її дії наближають її до цілі. Цей підхід часто використовується в задачах, пов'язаних з оптимізацією поведінки, таких як розв'язання ігрових задач чи управління складними системами. Всі три підходи мають свої механізми корекції

параметрів мережі і досягнення оптимальної продуктивності залежно від типу задачі та наявних даних.

Концептуальна еволюція машинного навчання. Глибоке навчання є революційним підходом у штучному інтелекті, який виник внаслідок еволюції машинного навчання і спричинив значні зміни в способах обробки та аналізу інформації. Цей метод став можливим завдяки розвитку обчислювальних технологій та накопиченню великих обсягів даних, що дозволило створювати потужні моделі, які можуть самостійно навчатися і виконувати складні завдання з високою точністю. Глибоке навчання відрізняється від традиційних статистичних методів та класичних алгоритмів машинного навчання тим, що воно спирається на багатошарові нейронні мережі, здатні навчатися за допомогою великої кількості прихованих шарів. Кожен із цих шарів обробляє дані на різних рівнях абстракції, що дає можливість створювати моделі, які можуть виявляти і моделювати складні закономірності в даних.

Історичні витоки глибокого навчання сягають 1940-х років, коли вчені почали розробляти перші нейронні мережі. Однак справжній прорив у цій галузі стався лише на початку 2010-х років. Це стало можливим завдяки суттєвому збільшенню обчислювальних потужностей комп'ютерів, що дозволило працювати з великими обсягами даних і реалізувати більш складні моделі. Такий розвиток технологій сприяв застосуванню глибокого навчання в різних сферах, таких як комп'ютерне бачення, обробка природної мови, рекомендаційні системи та багато інших.

Однією з основних характеристик глибокого навчання є багатошарова архітектура нейронних мереж. На відміну від традиційних моделей машинного навчання, які зазвичай використовують обмежену кількість шарів, глибокі нейронні мережі можуть мати десятки та навіть сотні шарів. Кожен шар здійснює дедалі складніші перетворення вхідних даних. У початкових шарах мережі виявляються прості ознаки, такі як контури та базові геометричні форми, тоді як глибші шари здатні комбінувати ці елементи в більш складні та абстрактні концепції. Наприклад, в задачах комп'ютерного зору перші шари можуть розпізнавати прості геометричні фігури, а наступні — обличчя, об'єкти або навіть цілі сцени. Це дозволяє мережам

розв'язувати складні завдання, які неможливо було б вирішити за допомогою традиційних методів.

Ще однією з вражаючих характеристик глибокого навчання є здатність до автоматичного виділення ознак. В класичних підходах машинного навчання ознаки зазвичай визначаються вручну експертами, що може бути дуже трудомістким та обмеженим у своїй здатності до виявлення складних залежностей. У глибоких нейронних мережах цей процес автоматизовано. Мережі самостійно виявляють найбільш інформативні ознаки з великого обсягу даних, що дозволяє значно покращити точність прогнозів та розпізнавання. Цей процес схожий на роботу людського мозку, де низчі рівні нервової системи сприймають прості сигнали, а вищі рівні інтегрують їх у складні концепції та абстракції. Завдяки автоматичному виділенню ознак, глибокі мережі можуть ефективно вирішувати завдання розпізнавання, класифікації та прогнозування, не потребуючи додаткової ручної обробки даних.

Глибоке навчання також відоме своєю здатністю до самонавчання та адаптації. Це відбувається завдяки використанню складних алгоритмів зворотного поширення помилки та різноманітних стратегій оптимізації, таких як градієнтний спуск, Adam або RMSprop. Завдяки таким методам, мережа безперервно коригує свої параметри, порівнюючи отримані результати з очікуваними. В такий спосіб мережа поступово зменшує похибки і підвищує свою точність. Цей процес можна порівняти з еволюційним: з кожним циклом мережа покращує свої результати: розвивається та підвищує свою точність. Здатність до самонавчання дозволяє створювати високоякісні моделі, які можуть адаптуватися до нових даних і змінюватися в процесі їх отримання.

Ієрархічна репрезентація даних є ще одним важливим принципом глибокого навчання. Кожен наступний шар мережі поступово трансформує вхідні дані, створюючи дедалі абстрактніші та узагальненіші представлення. Наприклад, у задачах обробки природної мови перші шари мережі можуть аналізувати окремі літери або слова, наступні шари працюють зі складами та синтаксисом, а глибші рівні здатні обробляти цілі речення та виявляти семантичні зв'язки. Така ієрархічна

структура дозволяє глибоким нейронним мережам працювати з даними на різних рівнях абстракції, що суттєво покращує їхню здатність до аналізу складних взаємозв'язків.

Отже, глибоке навчання пропонує новий підхід до обробки і аналізу даних, який має великий потенціал для вирішення найскладніших завдань, де традиційні методи не можуть дати задовільних результатів. Розвиток цієї технології відкриває нові горизонти для штучного інтелекту, зокрема в таких сферах, як обробка природної мови, комп'ютерне бачення, біоінформатика та багато інших.

Класифікація нейронних мереж охоплює різноманітні архітектури, кожна з яких має свої особливості та оптимізована для вирішення специфічних типів завдань. Вибір архітектури залежить від типу даних і поставлених цілей, оскільки різні мережі показують різну ефективність у обробці тексту, зображень, відео або інших форм даних.

є однією з найпростіших архітектур нейронних мереж, де інформація передається в одному напрямку — від вхідного шару до вихідного. У таких мережах немає зворотних зв'язків, і кожен нейрон в одному шарі зв'язаний з усіма нейронами наступного шару. Ця архітектура зазвичай застосовується для завдань класифікації та регресії, де дані не мають внутрішньої структури, такої як послідовність чи просторові відносини. FFNN є хорошим варіантом для простих задач, однак її здатність обробляти більш складні структури даних є обмеженою.

спеціалізуються на обробці даних, що мають просторову структуру, зокрема зображень. CNN використовують згорткові шари, які здійснюють операції згортки для виділення локальних ознак, таких як краї, текстури, форми, а потім пулінгові шари, які знижують розмірність і узагальнюють виявлені ознаки. Завдяки такій архітектурі, CNN є надзвичайно ефективними у задачах комп'ютерного зору, таких як розпізнавання об'єктів, класифікація зображень та сегментація. Мережі на основі CNN здатні автоматично навчатися виділяти важливі ознаки на різних рівнях абстракції, що дозволяє їм досягати високих результатів на великих наборах даних.

— це тип нейронних мереж, призначений для обробки послідовних даних, таких як текст, мовні сигнали або часові ряди. Основною характеристикою RNN є наявність

зворотних зв'язків між нейронами, що дозволяє мережі зберігати інформацію про попередні стани. Це надає RNN здатність "пам'ятати" попередні входи і використовувати цю інформацію для прогнозування наступних елементів у послідовності. RNN активно використовуються в задачах обробки природної мови, перекладу текстів, прогнозування часових рядів та багатьох інших областях, де важливо враховувати залежності між елементами послідовності.

є однією з найсучасніших архітектур, яка базується на механізмі уваги (attention). Це дозволяє моделі зосереджуватися на різних частинах послідовності і враховувати їх взаємозв'язок без необхідності обробляти дані послідовно, як це робиться в RNN. Механізм уваги дозволяє значно зменшити час обробки і забезпечити паралельну обробку даних. Завдяки своїй здатності працювати з довгими залежностями в текстах, Transformer мережі стали основою для таких моделей, як BERT і GPT, які досягли значних успіхів у задачах обробки природної мови, таких як машинний переклад, генерація текстів, питання-відповіді та інші. Архітектура Transformer показала, що механізм уваги може бути набагато ефективнішим за традиційні методи, особливо в задачах, де важливе виявлення глобальних залежностей у даних.

Кожен тип нейронної мережі має свої переваги та обмеження, і вибір між ними залежить від характеру задачі та специфіки оброблюваних даних, а також від якості та обсягу даних. Тому вибір архітектури нейронної мережі є критичним етапом у проєктуванні моделі, що вимагає врахування особливостей задачі, типу даних, ресурсів для навчання, а також можливості масштабування та інтерпретації отриманих результатів.

Особливості рекурентних нейронних мереж для прогнозування часових рядів (RNN(CNN), LSTM, Hybrid)

Рекурентні нейронні мережі (RNN) є одним із найбільш ефективних підходів для прогнозування часових рядів завдяки своїй здатності обробляти послідовні дані

та враховувати часові залежності. У традиційних нейронних мережах кожен вхід обробляється незалежно від інших, тоді як у RNN кожен вихід залежить від попереднього, що дає змогу ефективно моделювати часові залежності. Це робить RNN особливо корисними для прогнозування фінансових часових рядів, де важливим аспектом є врахування зміни тенденцій та взаємозв'язків між різними часовими точками. Проте, класичні RNN мають суттєву обмеженість, пов'язану зі "зникненням" або "експлозією" градієнтів під час навчання, що ускладнює їхнє використання для довгих послідовностей даних. Для базової архітектури RNN, де інформація передається через попередні стани, формула для обчислення активації на кожному етапі часу виглядає

Для базової архітектури RNN, де інформація передається через попередні стани, обчислення активації на кожному етапі часу виконується за формулою (2.1) [52, с.46]:

$$h_t = \tanh(W_{hx} x_t + W_{hh} h_{t-1} + b_h) \quad (2.1)$$

де: h_t – вектор активації на поточному етапі часу t
 x_t – вхід на поточному етапі часу
 W_{hx} – матриця ваг між попереднім станом прихованого шару та поточним вхідом
 W_{hh} – матриця ваг між попереднім станом прихованого шару та поточним станом
 b_h – зсув
 $\tanh$ – функція активації (гіперболічний тангенс).

Для розв'язання цієї проблеми застосовують розроблені спеціалізовані архітектури, такі як Long Short-Term Memory (LSTM) і Gated Recurrent Units (GRU), які успішно подолали недоліки стандартних RNN, забезпечивши ефективну обробку довгострокових залежностей у часових рядах. LSTM мережі використовують спеціальні одиниці пам'яті, що дозволяють їм зберігати інформацію протягом тривалих періодів часу, завдяки чому ці мережі набагато краще підходять для задач, де важливими є залежності на великих проміжках часу. Особливості архітектури LSTM включають використання механізмів забуття та оновлення пам'яті, що дозволяє моделі адаптуватися до зміни трендів у даних, що є важливим аспектом при прогнозуванні фінансових ринків, де різні фактори можуть сильно змінюватися з часом. LSTM є розширенням стандартних RNN і включає механізм пам'яті для

збереження важливої інформації на тривалий час. Базові рівняння для LSTM описують роботу основних компонентів, які контролюють потік інформації через осередок. Шлюз забуття визначає, яку частину попереднього стану необхідно зберегти або відкинути, регулюючи довготривалу пам'ять. Шлюз оновлення відповідає за додавання нової інформації до клітинного стану, визначаючи, які елементи вхідного сигналу мають значення. Оновлення клітинного стану виконується через комбінацію модульованого попереднього стану та нових кандидатів на оновлення. Шлюз виходу керує тим, яка частина клітинного стану вплине на поточний вихідний сигнал, формуючи прихований стан. Завдяки такій структурі LSTM здатна ефективно моделювати як довготривалі, так і короткотривалі залежності в часових рядах.

Однією з важливих переваг LSTM є здатність моделювати не лише лінійні, а й нелінійні залежності, що характерно для багатьох фінансових процесів. Наприклад, рухи на фондових ринках можуть бути зумовлені складними та нелінійними взаємозв'язками між економічними, політичними та психологічними чинниками. Для такої складної природи даних LSTM забезпечують високу точність прогнозів, оскільки здатні адаптуватися до змінюваних умов і враховувати складні патерни.

Основною особливістю LSTM є наявність спеціальних елементів, які виконують функцію контролю за інформацією, що потрапляє в модель, а також за її забуванням чи оновленням. Ці елементи називаються воротами, і кожне з них має свою специфічну функцію у процесі навчання.

Першим важливим елементом є вхідна брама (input gate). Вона визначає, яка частина нової інформації зберігається в клітинному стані мережі. Вхідна брама приймає на вхід поточне значення та прихований стан попереднього кроку, використовуючи їх для прийняття рішення про те, які нові дані вписати в клітинний стан. Цей процес фактично дозволяє моделі фільтрувати значення, що є найбільш релевантними для подальшого аналізу. Вхідна брама працює через сигмоїдну функцію, яка видає значення між 0 і 1, де 0 означає, що інформація відкинута, а 1 — повністю врахована.

Брама забування (forget gate) є важливим елементом архітектури LSTM. Вона контролює, яку частину існуючого клітинного стану потрібно "забути" або видалити. Це дозволяє мережі позбавлятися від старих або непотрібних даних, які більше не мають значення для поточного контексту. Така можливість важлива для коректної роботи з тривалими часовими рядами, де деякі дані, що поступово стають застарілими, можуть лише створювати шум, якщо їх не видаляти. Забування дозволяє зберігати пам'ять про важливі етапи послідовності, одночасно усуваючи ті, що більше не впливають на поточне прогнозування.

Вихідна брама (output gate) є останнім ключовим елементом. Вона визначає, яка частина клітинного стану буде виведена як вихідний сигнал на поточному етапі часу. Вихідна брама фільтрує клітинний стан за допомогою сигмоїдної функції, яка визначає, яка інформація з клітинного стану має бути передана далі в прихований стан мережі для наступних обчислень. Таким чином, вихідна брама контролює, яку інформацію мережа передає на наступний етап, забезпечуючи адаптивність моделі до поточного контексту.

Клітинний стан (cell state) є серцем архітектури LSTM і виконує роль своєрідного «каналу», через який проходить вся інформація в мережі. Він зберігає важливу інформацію протягом довгих послідовностей, не піддаючись зниженню її значення з часом, як це буває в традиційних рекурентних мережах. Клітинний стан ефективно передає дані через мережу, обробляючи і трансформуючи їх у відповідно до функції кожної з брам. Така здатність до збереження і передачі інформації є особливо важливою для задач, де контекст, накопичений на ранніх етапах послідовності, залишається релевантним на пізніших етапах.

Переваги архітектури LSTM є незаперечними. Однією з головних серед них є здатність долати проблему затухання градієнтів, яка є типовою для класичних рекурентних нейронних мереж. У традиційних RNN, коли градієнти через багато етапів зворотного поширення зменшуються до дуже малих значень, це може призвести до того, що модель не буде ефективно навчатися. LSTM вирішує цю проблему за допомогою клітинного стану та воріт, що дозволяє зберігати значення градієнтів протягом довгих послідовностей.

Іншою важливою перевагою є ефективне опрацювання довгих послідовностей. LSTM здатні "пам'ятати" важливу інформацію навіть через багато кроків часу, що робить їх ідеальними для задач, де необхідно враховувати далекі залежності між даними. Така здатність до селективного запам'ятовування та забування дозволяє моделі концентрувати увагу тільки на актуальних даних, покращуючи її здатність до прогнозування в умовах складних, багатофакторних процесів, таких як фінансові ринки.

Одним із найновіших досягнень в області нейронних мереж є гібридні моделі, що поєднують рекурентні нейронні мережі, такі як LSTM, з іншими типами нейронних мереж, зокрема згортковими (CNN). Комбінація RNN та CNN виявилася надзвичайно ефективною для прогнозування складних фінансових часових рядів, оскільки кожен тип мережі додає свої сильні сторони. CNN, у свою чергу, використовуються для виділення локальних ознак та патернів у даних, що є важливим для точного прогнозування короткострокових змін, таких як коливання ринку. Згорткові шари здатні виявляти основні характеристики, що мають важливість для подальшого аналізу, у той час як LSTM дозволяють мережі зберігати та обробляти ці ознаки протягом довгого часу, підвищуючи точність прогнозів на основі минулих тенденцій.

Гібридні архітектури дозволяють комбінувати переваги обох типів нейронних мереж, забезпечуючи як високу здатність до аналізу локальних патернів, так і довгострокову пам'ять для роботи з послідовними даними. Такі моделі є надзвичайно потужними для прогнозування фінансових ринків, де необхідно враховувати як короткострокові зміни, так і довгострокові тренди. Завдяки здатності обробляти складні залежності, такі моделі демонструють значну перевагу у порівнянні з класичними підходами, що використовують лише одну з архітектур. У формулі 2.2 зображена формула для обробки просторових (локальних) ознак з використанням CNN, а потім передача результатів в LSTM для врахування часових залежностей. Після застосування CNN результати передаються в LSTM, де обробка продовжується за допомогою механізмів пам'яті, як це відображено формулою (2.2) [42, с. 72]:

де: x_t – вхідні дані на етапі часу t ;

$CNN(x_t)$ – вектор просторових ознак, обчислений згортковою мережею;

– прихований стан LSTM на етапі часу t , який враховує як просторові, так і часові залежності.

Отже, рекурентні нейронні мережі, зокрема LSTM, а також гібридні моделі, що поєднують RNN з CNN, є одними з найефективніших підходів для прогнозування фінансових часових рядів. Їхня здатність моделювати складні залежності та враховувати як короткострокові, так і довгострокові патерни робить ці моделі надзвичайно корисними для задач фінансового прогнозування, де традиційні методи часто виявляються недостатньо точними та гнучкими.

2.4 Метрики оцінювання якості прогнозування

Оцінювання якості моделей прогнозування є важливим етапом при застосуванні глибоких нейронних мереж, особливо в таких сферах, як фінансове прогнозування, де точність результатів може безпосередньо впливати на економічні рішення. Для цього застосовуються різноманітні метрики.

Однією з найпростіших і найбільш інтуїтивно зрозумілих метрик є *Mean* (MAE), яка вимірює середню абсолютну різницю між фактичними та передбаченими значеннями. MAE є хорошим показником для оцінювання загальної точності прогнозу, оскільки легко інтерпретується: чим менше значення MAE, тим точніше модель. Оскільки метрика використовує абсолютні величини похибки, вона стійка до екстремальних значень або викидів, що може бути важливим при роботі з фінансовими часовими рядами, де викиди часто є результатом випадкових або незначних подій. Однак MAE не враховує величину цих помилок у контексті їхньої значущості, тобто важливість великих помилок для моделі може бути недооцінена.

(MSE) є іншою популярною метрикою, яка в порівнянні з MAE надає більше значення великим похибкам, оскільки різниця між фактичними і передбаченими значеннями підноситься до квадрату. Це означає, що значні відхилення мають сильніший вплив на загальний результат, що робить MSE корисною, коли необхідно

мінімізувати великі похибки. Метрика часто використовується в задачах, де критичні великі помилки (наприклад, різкі фінансові втрати), і їхня мінімізація є пріоритетом. Проте MSE чутливий до викидів, оскільки квадрати великих відхилень можуть значно спотворити загальну картину, особливо коли в наборі даних є екстремальні значення. Це може робити її менш стабільною порівняно з іншими метриками, коли дані мають значну варіабельність або шум.

(RMSE) є квадратним коренем з MSE і повертає метрику до оригінальних одиниць вимірювання, що робить її більш інтерпретованою. Завдяки цьому, RMSE дає більш наочне уявлення про середню величину похибки, оскільки величини прогнозованих і фактичних значень вимірюються в тих самих одиницях. Як і MSE, RMSE робить акцент на значних помилках, що дозволяє ефективно оцінювати моделі в контексті задач, де важливо уникати великих відхилень. Зниження RMSE свідчить про підвищення точності прогнозу. Однак вона також залишається чутливою до викидів, хоча і меншою мірою, ніж MSE, через наявність квадратного кореня. Загалом, RMSE є однією з найбільш поширених метрик для оцінювання якості моделей у задачах регресії, зокрема в фінансових прогнозах.

(MAPE) є метрикою, яка оцінює середню абсолютну похибку у відсотках. Це дозволяє порівнювати точність прогнозів незалежно від масштабу змінних, що робить MAPE корисною для моделей, які працюють з різними наборами даних або мають різні одиниці вимірювання. Зазвичай метрика застосовується в задачах, де важлива відносна точність прогнозів, наприклад, в економічному прогнозуванні або при оцінюванні ефективності інвестицій. Метрика зручно інтерпретується, оскільки дає змогу побачити похибку в відсотковому вираженні. Однак MAPE має певні обмеження, особливо при роботі з даними, близькими до нуля, оскільки похибка у таких випадках може набувати дуже великих значень, що призводить до неточностей. Це може бути серйозною проблемою, коли фінансові дані містять значення, які наближаються до нуля, що є типовим для певних активів чи ринків.

РОЗДІЛ 3 ПРАКТИЧНА РЕАЛІЗАЦІЯ ПРОГНОЗНОЇ МОДЕЛІ ДЛЯ ЦІН НА АКЦІЇ НА ОСНОВІ ГЛИБОКОГО НАВЧАННЯ

Попередня обробка статистичних даних

Прогнозування цін на акції є одним з найскладніших та найактуальніших завдань у фінансовій аналітиці. Використання методів глибокого навчання для передбачення динаміки фінансових ринків становить особливий інтерес через їхню здатність виявляти складні нелінійні залежності в масивах фінансових даних. Метою даного дослідження є розробка інтелектуальної системи, яка здатна передбачати майбутню вартість акцій на основі аналізу попередніх часових рядів з використанням сучасних нейромережевих архітектур.

Для вирішення задачі прогнозування цін на акції у роботі використано набір інструментів, що дозволяє ефективно обробляти великі обсяги фінансових даних та реалізувати моделі машинного навчання. Основна частина роботи проведена у середовищі *Google Colab*, що забезпечує зручний доступ до потужних обчислювальних ресурсів та сприяє швидкій розробці і тестуванню моделей. Для реалізації та аналізу застосовано мову програмування *Python*, яка на сьогодні є одним з найпопулярніших інструментів для наукових досліджень і обробки даних завдяки своїй простоті, широкому спектру бібліотек та підтримці наукових і статистичних методів.

Для завантаження та підготовки фінансових даних використано бібліотеку *yfinance*. Ця бібліотека забезпечує доступ до історичних даних фінансових інструментів з різних біржових платформ, таких, як Yahoo Finance. Зокрема функція `get_stock_data()` відповідає за вивантаження котирувань обраного фінансового інструменту за визначений проміжок часу.

Завантаження даних здійснене за допомогою простих запитів, що дозволило ефективно отримувати необхідні фінансові часові ряди для подальшого аналізу.

Для побудови і тренування моделей використано бібліотеку *tensorflow*, яка є потужним фреймворком для розробки нейронних мереж та інших моделей машинного навчання. Вона дозволяє працювати з різноманітними архітектурами нейронних мереж, включаючи глибинні нейронні мережі, рекурентні нейронні мережі та LSTM, що є особливо важливими для прогнозування часових рядів, зі складною нелінійною структурою.

Для попередньої обробки даних, нормалізації та розділення наборів даних на тренувальні та тестові вибірки використано бібліотеку *scikit-learn*, яка є стандартом для чисельних методів машинного навчання та статистичних алгоритмів. Вона надає зручні інструменти для попереднього аналізу та трансформації даних, таких як масштабування, нормалізація та крос-валідація моделей.

Принципово важливим моментом для подальшого прогнозування є коректність та повнота вхідних даних. Система перевіряє наявність необхідних параметрів, зокрема стовпця закриття біржової сесії, який слугуватиме основним джерелом інформації для прогнозування. Трансформація вхідного масиву в двовимірний формат дозволяє оптимізувати подальшу обробку нейронною мережею.

Функція `prepare_data()` реалізує технічно складний механізм перетворення часового ряду на послідовність, придатну для навчання нейронної мережі. Її ключова мета - створення навчальної вибірки, де кожен попередній часовий інтервал розглядається як вхідні дані, а наступне значення - як цільова змінна.

Алгоритм послідовно проходить через весь масив даних, формуючи навчальні послідовності фіксованої довжини. Наприклад, якщо обрано інтервал у 10 днів, то мережа навчатиметься передбачати наступне значення, базуючись на динаміці попередніх десяти днів. Такий підхід дозволяє змодельовати внутрішні закономірності та тренди фінансового ринку.

Використання бібліотек `numpy` та `ufinance` забезпечує високу ефективність та гнучкість підготовки даних. Перетворення одновимірного масиву на багатовимірний формат (`reshape`) є критичним для коректної роботи нейронних мереж, особливо архітектур, призначених для роботи з послідовностями, як-от LSTM або GRU.

Важливо розуміти, що прогнозування цін на акції - це надзвичайно складне завдання. Фінансові ринки характеризуються високою волатильністю та впливом численних зовнішніх факторів, які важко формалізувати. Жодна математична модель не може гарантувати абсолютно точне передбачення, проте глибоке навчання надає потужний інструментарій для аналізу складних нелінійних залежностей.

.2 Розробка моделі прогнозування: вибір архітектури та налаштування гіперпараметрів

Розробка ефективної моделі машинного навчання для передбачення цін на акції є складним інженерним завданням, яке вимагає глибокого розуміння як математичних принципів, так і особливостей архітектури нейронних мереж. У контексті фінансового прогнозування ключову роль відіграє вибір архітектури, здатної адекватно інтерпретувати часові послідовності з притаманними їм складними нелінійними залежностями.

Модель ARIMA є однією з найпоширеніших у прогнозуванні часових рядів завдяки своїй універсальності та здатності моделювати широкий спектр даних. Її ефективність значною мірою залежить від правильного налаштування ключових гіперпараметрів: порядку авторегресії (p), ступеня диференціювання (d) та порядку ковзного середнього (q). Параметр p визначає кількість попередніх значень, які враховуються під час прогнозування, тоді як d використовується для забезпечення стаціонарності часового ряду шляхом диференціювання. Параметр q враховує вплив попередніх похибок на поточний прогноз.

Процес налаштування гіперпараметрів включає тестування різних їхніх комбінацій для визначення оптимальних значень. Для цього застосовуються статистичні критерії, такі як AIC (інформаційний критерій Акаїке) та BIC (байєсівський інформаційний критерій), які дозволяють обрати модель із найкращою збалансованістю між складністю та якістю прогнозу. Крім того, використання крос-

валідації дає змогу перевірити узгодженість моделі з новими даними, що є важливим етапом для підвищення її прогностичної спроможності.

Експоненційне згладжування є простим і водночас потужним методом для прогнозування часових рядів із трендами або сезонними коливаннями. Основним параметром цих моделей є коефіцієнт згладжування α , який визначає вагу останніх спостережень у прогнозі. Значення α впливає на чутливість моделі: нижчі значення забезпечують сильніше згладжування, роблячи модель менш чутливою до короткострокових коливань, тоді як вищі значення дозволяють швидше реагувати на зміни у даних.

Налаштування α зазвичай виконується шляхом оптимізації, спрямованої на мінімізацію середньоквадратичної похибки прогнозу. Ефективність вибраного параметра перевіряється на тестових вибірках, що дозволяє оцінити здатність моделі до узагальнення та її стійкість у прогнозуванні нових даних.

нейронні мережі становлять особливий клас глибоких архітектур, спеціально призначених для роботи з послідовностями. На відміну від традиційних нейронних мереж, LSTM мають унікальну здатність зберігати та передавати контекстну інформацію через тривалі часові інтервали, що робить їх надзвичайно ефективними для аналізу фінансових часових рядів.

Архітектура LSTM-моделі, представлена функцією `create_lstm_model()`, демонструє витончений підхід до моделювання складних залежностей. Перший шар LSTM з 64 нейронами та налаштуванням `return_sequences=True` дозволяє мережі опрацьовувати повну послідовність вхідних даних, зберігаючи проміжні стани. Подальший шар Dropout із коефіцієнтом 0.2 слугує механізмом регуляризації, який запобігає перенавчанню шляхом тимчасового "вимикання" частини нейронних зв'язків.

Конволюційні нейронні мережі (CNN), традиційно асоційовані з обробкою зображень, демонструють несподівану ефективність у аналізі часових послідовностей. Функція `create_cnn_model()` реалізує одновимірну конволюційну архітектуру, яка здатна виявляти локальні патерни та закономірності в динаміці фінансових показників.

Багатошарова структура з конволюційними та пулінговими шарами послідовно зменшує розмірність даних, витягуючи найбільш інформативні ознаки. Шари MaxPooling дозволяють мережі концентруватися на найбільш значущих характеристиках, ігноруючи несуттєві флуктуації. Завершальні повнозв'язні шари інтегрують extracted features у фінальне прогностичне передбачення.

Найбільш просунутим підходом є *створення гібридної моделі*, яка поєднує переваги CNN та LSTM. Функція `create_hybrid_model()` демонструє інноваційний підхід до архітектурної композиції, де конволюційні шари спочатку виділяють просторові характеристики, а потім LSTM-шари опрацьовують темпоральні залежності.

Така архітектура дозволяє мережі одночасно аналізувати локальні патерни та глобальні тренди, що особливо важливо при прогнозуванні фінансових показників. Багаторівнева структура з dropout-шарами забезпечує додаткову стійкість моделі до перенавчання.

Всі представлені моделі компілюються з використанням *оптимізатора Adam* - сучасного алгоритму градієнтного спуску, який демонструє високу ефективність у задачах глибокого навчання. Функція втрат Mean Squared Error (MSE) забезпечує мінімізацію квадратичної похибки передбачення, а метрика Mean Absolute Error (MAE) надає додаткову інформацію про точність моделі.

Важливо розуміти, що жодна математична модель не може гарантувати абсолютно точне передбачення динаміки фінансових ринків. Представлені архітектури слугують потужним інструментом аналізу, але вимагають ретельного налаштування, валідації та постійного вдосконалення з урахуванням специфіки конкретних фінансових інструментів.

Кожна з презентованих архітектур – ARIMA, Exponential model, LSTM, CNN та гібридна модель - має власні сильні та слабкі сторони. Вибір остаточної архітектури залежить від результатів експериментальних досліджень, специфіки вхідних даних та конкретних дослідницьких цілей.

Представлений підхід є фундаментом для створення інтелектуальної системи прогнозування, яка поєднує математичну строгість класичних статистичних методів та потужність сучасних технологій штучного інтелекту.

.3 Навчання, оптимізація моделі, підбір кращих параметрів та оцінка точності і якості моделей

Представлена реалізація демонструє комплексний підхід до розв'язання надзвичайно складної задачі прогнозування цін на акції з використанням сучасних технологій машинного навчання. Фундаментальна мета дослідження полягає у створенні інтелектуальної системи, здатної виявляти приховані закономірності в динаміці фінансових ринків та генерувати статистично обґрунтовані передбачення.

Методологія дослідження базується на синергетичному підході, що передбачає одночасне навчання та порівняння декількох нейромережових архітектур. Обрання акцій Apple Inc. (AAPL) як об'єкта дослідження зумовлено високою ліквідністю та репрезентативністю цінового ряду. Часовий горизонт з 1 січня 2018 року до поточної дати забезпечує охоплення різноманітних ринкових циклів та економічних трансформацій.

Принципово важливим етапом є коректна *підготовка вхідних даних*. Застосування MinMaxScaler дозволяє здійснити нормалізацію значень до уніфікованого діапазону від 0 до 1, що забезпечує більш стабільне навчання нейронних мереж. Параметр `time_steps`, встановлений на позначці 60 днів, визначає глибину часового контексту для прогнозування, створюючи оптимальний баланс між повнотою історичної інформації та чутливістю до актуальних тенденцій.

Дослідження реалізує п'ять різних архітектурних підходів: ARIMA, 50 епох з валідаційною вибіркою 10% та batch-розміром 32, для LSTM, batch-розміром 32 для CNN та гібридної моделі. Такий підхід дозволяє не лише навчити моделі, але й здійснити їх порівняльну оцінку за різними метриками точності.

Для комплексної оцінки ефективності моделей використовуються три статистичні метрики:

- середньоквадратична похибка (RMSE);
- середня абсолютна похибка (MAE);
- середньопроцентна абсолютна похибка (MAPE)є

Такий підхід забезпечує всебічне розуміння прогностичних властивостей розроблених моделей.

Завершальним етапом дослідження є графічна візуалізація результатів, що дозволяє здійснити наочне порівняння прогностичних значень різних моделей з реальною ціновою динамікою. Графічне представлення слугує потужним інструментом інтерпретації отриманих результатів.

Важливо усвідомлювати, що представлена реалізація є перевіркою працездатності моделі, побудови прогностичної системи. Подальші дослідження вимагатимуть поглибленої оптимізації архітектур, експериментів з гіперпараметрами та розширення інформаційної бази.

Представлений підхід демонструє потужність сучасних технологій машинного навчання у розв'язанні складних задач фінансового прогнозування, відкриваючи нові горизонти для досліджень у сфері штучного інтелекту та фінансової аналітики.

Базова методологія дослідження передбачає використання п'яти різних нейромережевих архітектур (ARIMA, Exponential model, LSTM, CNN, гібридної) для прогнозування цін на акції компанії Apple. Незначні коригування стосуються переважно технічних параметрів навчання та оцінки результатів.

Параметри налаштування моделей після проведення першого навчання та аналізу результатів було змінено. Порівняно з попередньою версією, відбулася одна принципова зміна - зменшення розміру batch з 32 до 16 для LSTM та збільшено розміру batch з 16 до 32 для CNN та гібридної моделі. Причини цієї зміни пов'язані з кількома важливими аспектами машинного навчання:

енший batch size може допомогти моделі краще узагальнювати patterns, зменшуючи ризик переналаштування;

енші batch дозволяють частіше оновлювати вагові коефіцієнти, що може призвести до більш стабільної збіжності;

астіші оновлення ваг дають моделі можливість швидше адаптуватися до тонких особливостей даних.

Інші параметри залишилися незмінними:

- Часовий діапазон з 1 січня 2018 року
- Кількість днів для прогнозування (60 днів)
- Кількість епох навчання (50)
- Частка валідаційної вибірки (0.1 або 10%)

В процесі нормалізації даних застосовано MinMaxScaler з налаштуванням діапазону від 0 до 1, що забезпечує уніфікацію вхідних даних та покращення процесу навчання нейронних мереж.

Основні зміни полягають у незначних модифікаціях процесу оцінювання та представлення результатів:

одано розширену обробку винятків при завантаженні даних;

бережено механізм оцінювання моделей через метрики RMSE, MAE та MAPE
инесено функцію оцінки моделей в окрему локальну функцію `evaluate_model()`

бережено механізм графічної візуалізації результатів

Друга спроба є логічним уточненням та незначною оптимізацією попередньої версії, що демонструє ітеративний підхід до дослідження. Зменшення batch size є свідомим експериментальним кроком, спрямованим на потенційне покращення якості прогнозування.

Для підвищення якості прогнозування можна запропонувати такі кроки:

кспериментування з різною кількістю нейронів у шарах

астосування технік регуляризації

икористання більшої кількості технічних індикаторів

провадження механізмів крос-валідації

озширення часового діапазону вхідних даних

Код ілюструє послідовний та методологічно обґрунтований підхід до застосування машинного навчання в задачі прогнозування цін на акції. Інтеграція

різномісних нейромережових архітектур сприяє отриманню більш точних та репрезентативних результатів, оскільки кожна з моделей вносить свій внесок у покращення прогностичної здатності. Однак необхідно усвідомлювати, що жодна математична модель не здатна з абсолютно високою точністю передбачити майбутні ціни акцій, оскільки фінансові ринки підлягають численним випадковим та складним факторам впливу. Представлена реалізація виступає дослідницьким інструментом, що демонструє потенціал машинного навчання в контексті фінансової аналітики та прогнозування.

Загальний висновок: модифікована версія реалізації є логічним етапом розвитку попереднього дослідження, що підтверджує поступальний характер наукового прогресу в області прогнозування фінансових часових рядів.

У контексті дослідження прогнозування цін на акції компанії Apple застосовано комплексний підхід до тестування та валідації нейромережових моделей. Процес оцінки ефективності ґрунтується на використанні трьох ключових статистичних метрик, які надають всебічну характеристику якості прогнозування: середньоквадратична похибка (Root Mean Square Error, RMSE), середня абсолютна похибка (Mean Absolute Error, MAE) та середньопроцентна абсолютна похибка

Вибір саме цих метрик обумовлений їхньою здатністю надавати різномісноту інформацію про точність прогнозування. RMSE особливо чутлива до великих відхилень, MAE забезпечує лінійну оцінку похибки, а MAPE дозволяє оцінити відносну похибку у відсотковому вираженні, що особливо важливо при аналізі фінансових часових рядів.

У першій спробі проведено тестування трьох різних нейромережових архітектур: LSTM (Long Short-Term Memory), CNN (Convolutional Neural Network) та Ексоненціальне згладжування. Отримані результати, представлені в табл. 3.1, демонструють диференційовану ефективність кожної архітектури.

Проведений ряд додаткових тестів з класичними статистичними моделями, такими як ARIMA та Метод експоненційного згладжування призвів до результатів, зображених на рис. 3.1 та рис. 3.2.

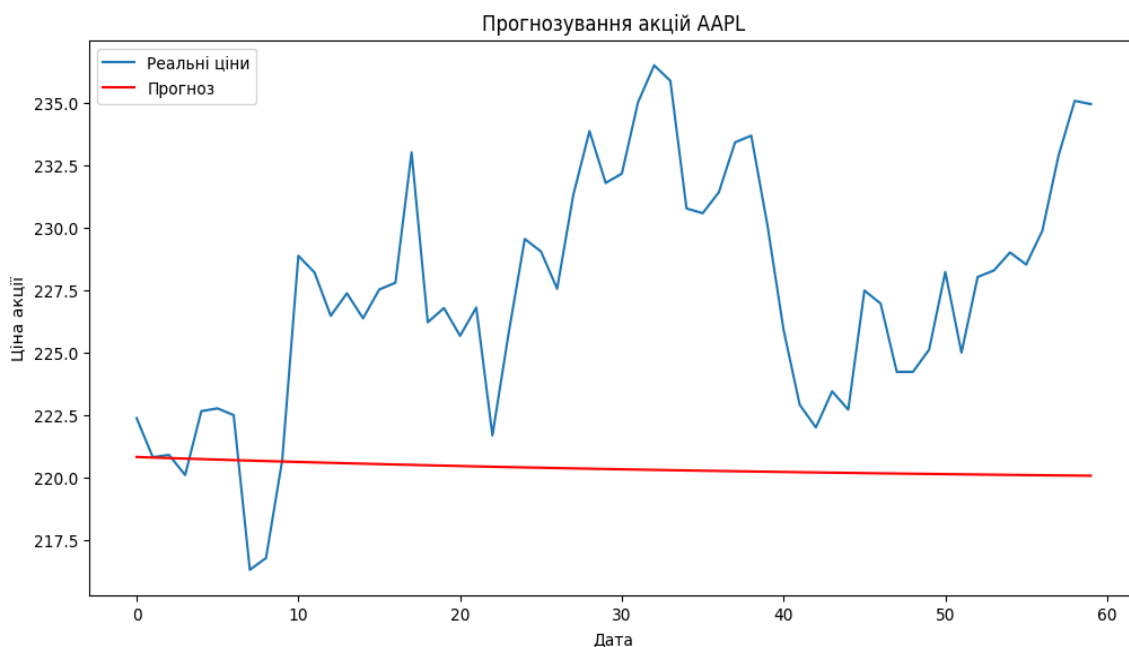


Рисунок 3.1 – Результати прогнозування ціни акцій за методом ARIMA

Джерело: розроблено автором

Згідно з графіком, модель прогнозування показує пряму лінію для прогнозу ціни акцій AAPL (Apple Inc.). Це означає, що модель передбачає, що ціна акцій буде залишатися постійною протягом наступного періоду спостереження. Така пряма лінія для прогнозу вказує, що використана модель ARIMA (Autoregressive Integrated Moving Average) не змогла виявити складні динамічні та сезонні закономірності в ціні акцій Apple. Зазвичай акції компаній технологічного сектору, таких як Apple, можуть мати складні тренди та сезонні коливання, які простіша модель ARIMA не здатна ефективно спрогнозувати.

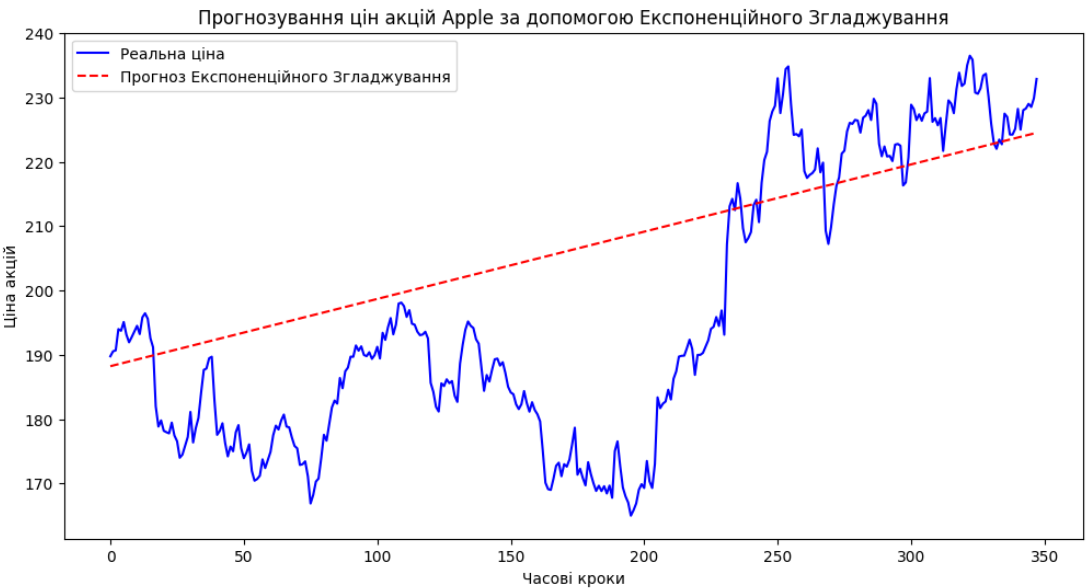


Рисунок 3.2 – Результати експоненціального згладжування часового ряду ціни акцій

Джерело: розроблено автором

Експоненціальна модель прогнозування виявляється недостатньо ефективною для аналізу волатильних ринків через свою обмежену здатність адаптуватися до різких змін у даних. Її проста структура дозволяє враховувати лише базові тренди або середні значення, що робить її малопридатною для моделювання складних, нелінійних залежностей, характерних для фінансових часових рядів. Волатильність, яка проявляється у вигляді частих та значних коливань цін, значно ускладнює завдання прогнозування, оскільки проста експоненціальна модель не має механізмів для обробки таких динамічних змін.

Таблиця 3.1 – Результати метрик першого етапу дослідження

		Гібридна		
--	--	----------	--	--

Джерело: розроблено автором

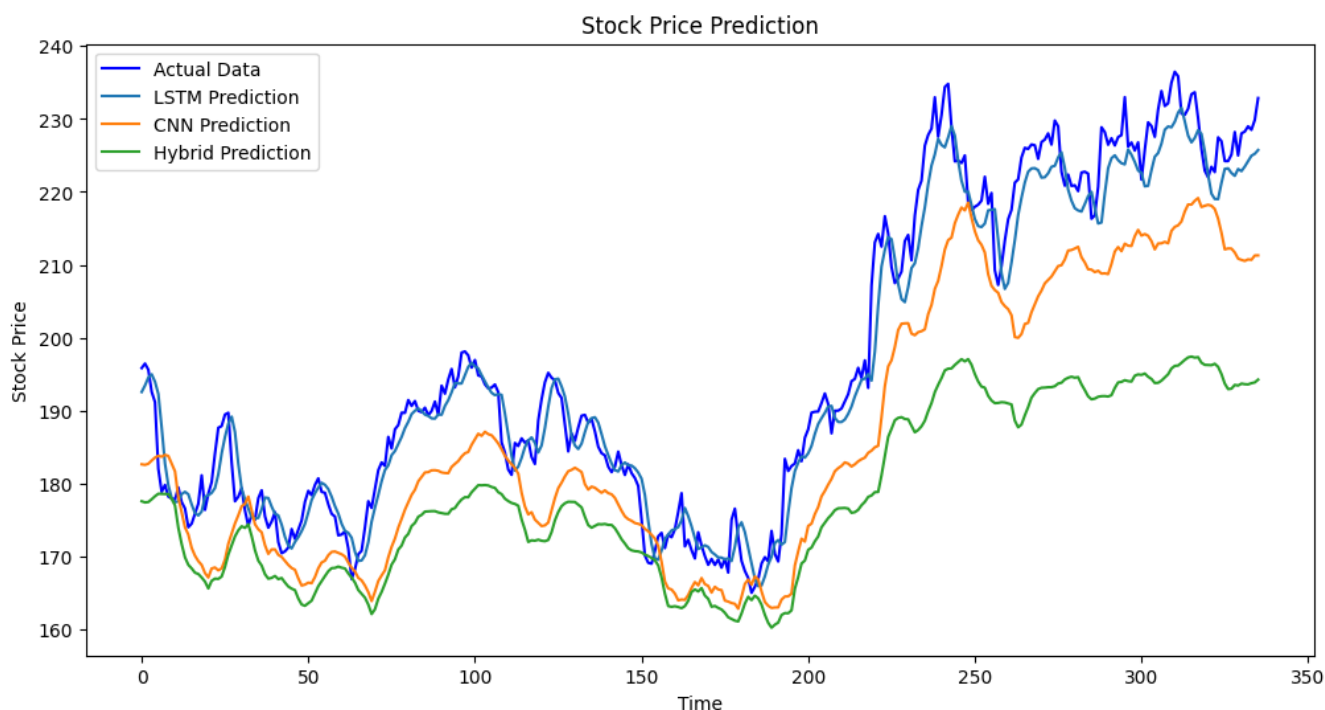


Рисунок 3.3— Результат навчання на першому етапі дослідження без покращених гіперпараметрів

Джерело: розроблено автором

Порівняльний аналіз результатів навчання нейромереж показує, що LSTM модель демонструє найкращу точність прогнозування серед трьох архітектур. Найнижчі значення RMSE, MAE та MAPE свідчать про її вищу здатність виявляти часові залежності та закономірності в динаміці цін на акції Apple.

На другому етапі було запроваджено низку методологічних вдосконалень, спрямованих на підвищення якості та надійності прогнозування:

проваджено складніший механізм обробки потенційних помилок та некоректних даних. Це дозволило підвищити стійкість алгоритму до неочікуваних варіацій вхідних даних та знизити ризик некоректної інтерпретації фінансової інформації.

бережено попередній підхід до оцінювання ефективності моделей через метрики RMSE, MAE та MAPE. Це забезпечує можливість коректного порівняння результатів між різними експериментальними спробами.

инесення функції оцінки моделей `evaluate_model()` в окрему локальну функцію дозволило:

- підвищити модульність коду;
- спростити процес налагодження та тестування;
- забезпечити кращу читабельність та супроводжуваність програмного забезпечення.

родовжено практику графічного представлення результатів, що надає додаткові можливості для інтерпретації та аналізу отриманих прогнозів.

На рис. 3.3 можна побачити результати другого тесту. Всі моделі покращили результати. Результати за основними метрикам наведені в табл. 3.2

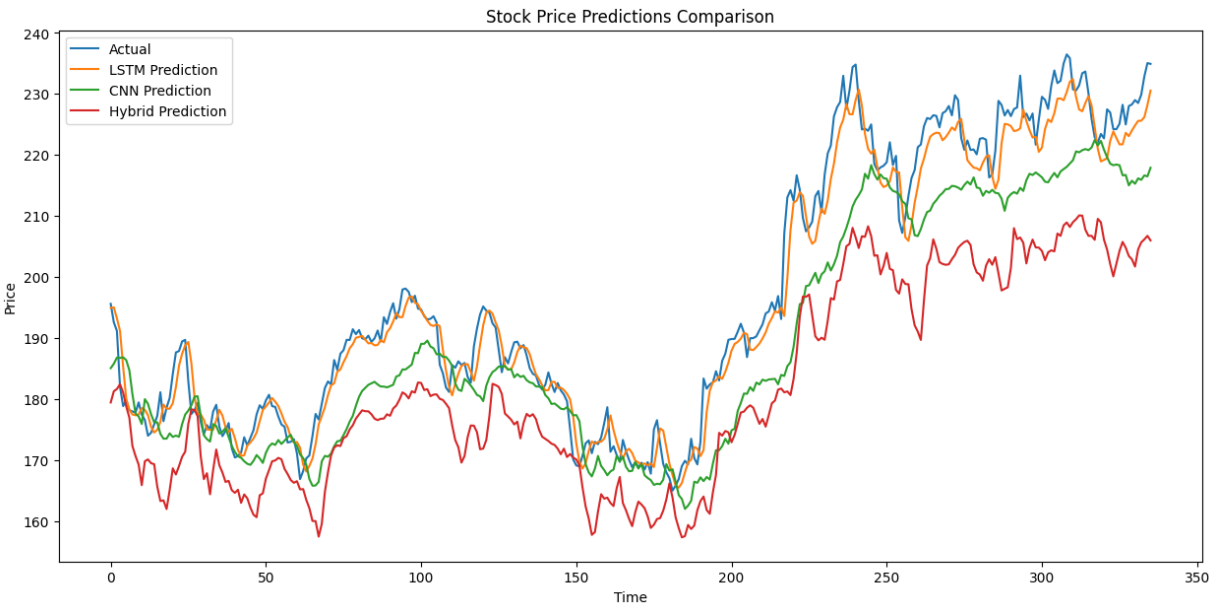


Рисунок 3.4 – Графічне зображення прогнозування після зміни гіперпараметрів
Джерело: розроблено автором

Таблиця 3.2 – Результати по основним метрикам після зміни гіперпараметрів

LSTM модель	CNN модель	Гібридна модель

Джерело: розроблено автором

Важливо усвідомлювати, що жоден статистичний метод не може забезпечити абсолютно точне передбачення динаміки фінансових ринків. Представлена

методологія є дослідницьким інструментом, який демонструє потенціал машинного навчання у прогнозуванні часових рядів.

Серед перспективних напрямків подальших досліджень можна виділити:

- Впровадження додаткових технік регуляризації
- Експериментування з архітектурами нейронних мереж
- Розширення набору вхідних технічних індикаторів
- Застосування більш складних механізмів крос-валідації
- Включення зовнішніх факторів впливу на ціну акцій

Кожна ітерація дослідження дозволяє поступово вдосконалювати методологію та наближатися до більш точних прогностичних моделей у складній та динамічній сфері фінансового аналізу.

У цьому розділі проводиться глибокий порівняльний аналіз результатів прогнозування, отриманих за допомогою різних моделей машинного навчання, зокрема рекурентної нейронної мережі довгої короткочасної пам'яті (LSTM), згорткової нейронної мережі (CNN), гібридної моделі, а також класичних статистичних методів, таких як ARIMA та метод експоненційного згладжування. Порівняння здійснено з фокусуванням на оцінювання точності прогнозів і на виявлення переваг і обмежень різних підходів до аналізу часових рядів.

В результаті першої серії експериментів, в яких використано стандартні параметри моделей, при порівнянні ефективності трьох різних архітектур нейронних мереж: LSTM, CNN і гібридної моделі, яка поєднувала характеристики обох попередніх архітектур виявилися значні відмінності у точності прогнозів для кожної з моделей.

Модель LSTM (Long Short-Term Memory) показала найкращі результати за всіма основними метриками: $RMSE = 3.9792$, $MAE = 3.1032$, $MAPE = 1.55\%$. Це свідчить про здатність LSTM моделі ефективно вивчати та моделювати складні, нелінійні залежності у часових рядах фінансових даних, таких як ціни акцій. Модель LSTM продемонструвала відмінні результати завдяки своїй архітектурі, здатній враховувати залежності на великих відрізках часу та враховувати значні зміни в даних. Це особливо важливо при прогнозуванні волатильних ринків, де важливо

адаптуватися до постійно змінюваних умов. Модель LSTM має унікальну здатність зберігати інформацію про попередні стани, що дозволяє їй краще прогнозувати майбутні тенденції на основі історичних даних.

Модель CNN (Convolutional Neural Network) продемонструвала гірші результати порівняно з LSTM: $RMSE = 12.1664$, $MAE = 10.4066$, $MAPE = 5.14\%$. Хоча CNN в основному застосовуються для обробки зображень, її здатність працювати з часовими рядами також була перевірена в контексті фінансового прогнозування. Однак, CNN має більшу вразливість до короткострокових коливань і змін у даних, що може призводити до зниження точності прогнозу на волатильних ринках. Конволюційні шари у CNN здатні ефективно виділяти важливі ознаки з даних, але можуть мати труднощі з врахуванням тривалих залежностей у часових рядах, що є важливим для прогнозування складних фінансових патернів.

Гібридна модель, яка поєднувала LSTM та CNN, показала ще гірші результати: $RMSE = 21.2556$, $MAE = 18.1674$, $MAPE = 8.79\%$. Така модель намагалася комбінувати переваги двох різних підходів, але результати вказують на те, що поєднання цих двох методів не завжди дає очікуваний результат. У даному випадку, гібридна модель не змогла ефективно інтегрувати сильні сторони кожної з моделей, що призвело до зниження точності прогнозування. Це може бути зумовлено складністю налаштування і збільшеною кількістю параметрів, що ускладнює навчання моделі та її адаптацію до даних.

Модель ARIMA (AutoRegressive Integrated Moving Average) показала кращі результати за CNN і гібридну модель, з $RMSE = 8.458$, $MAE = 7.2789$, $MAPE = 3.1522\%$. ARIMA є класичним методом для аналізу часових рядів і зазвичай використовується для прогнозування короткострокових трендів. Однак, в умовах волатильних ринків, де важливо враховувати швидкі зміни та коливання, ARIMA не була такою ефективною, як LSTM. Хоча ARIMA продемонструвала стабільні результати, вона все ж поступалася більш сучасним методам, таким як LSTM, які мають більший потенціал для адаптації до змінних умов ринку.

Експоненціальна модель, яка є одним з найпростіших підходів до прогнозування, показала найгірші результати серед усіх моделей: $RMSE = 18.0079$,

$MAE = 14.4024$, $MAPE = 11.72\%$. Цей метод має обмежену здатність працювати з волатильними ринками, оскільки він не враховує складних патернів та динамік, характерних для фінансових часових рядів. Експоненціальна модель покладається на прості припущення про тренди та рівновагу в даних, що робить її неефективною для прогнозування на ринках, що характеризуються високою волатильністю.

Загалом, результати першого експерименту показують, що LSTM є найефективнішою моделлю для прогнозування фінансових ринків, оскільки вона здатна враховувати довгострокові залежності та адаптуватися до волатильних змін. Інші моделі, хоча й мають свої переваги в окремих умовах, показали значно гірші результати в контексті фінансового прогнозування на волатильних ринках.

Після коригування стандартних параметрів проведено другий експерименті з прогнозування ціни акцій. Порівняння результатів прогнозування, проведених на основі двох експериментів, дає змогу виявити ефективність різних моделей при використанні стандартних і покращених гіперпараметрів. У першому експерименті розглянуто п'ять моделей: LSTM, CNN, гібридну модель, ARIMA та експоненціальну модель. Результати першого експерименту дозволяють провести детальний аналіз ефективності кожної моделі з окремими висновками.

У другому експерименті здійснено оптимізацію гіперпараметрів для трьох моделей: LSTM, CNN та гібридної моделі. Важливо зазначити, що моделі ARIMA та експоненціальна модель не були включені до цього експерименту через їхню низьку ефективність у попередньому аналізі. Оптимізація гіперпараметрів мала значний позитивний вплив на результати, підвищивши точність усіх трьох моделей.

Модель LSTM, після налаштування гіперпараметрів, досягла ще кращих результатів. Значення RMSE знизилося до 3.6401, MAE досягло 2.8126, а MAPE зменшилося до 1.40%. Ці результати свідчать про високу здатність LSTM адаптуватися до даних після оптимізації, що підтверджує ефективність цього підходу в контексті прогнозування на фінансових ринках. Параметри моделі, зокрема кількість шарів і розмір партії, були налаштовані таким чином, що дозволили знизити похибки прогнози. Крім того, LSTM модель показала значне поліпшення в здатності

враховувати довгострокові залежності та відображати складні тренди ринку, що є основною перевагою цієї технології у прогнозуванні фінансових даних.

Модель CNN також показала покращення в результатах прогнозування після оптимізації. Величина RMSE зменшилася з 12.1664 до 10.1205, MAE знизилася до 8.2054, а MAPE зменшилася до 4.03%. Хоча ці показники не досягли рівня LSTM, все ж спостерігається значне поліпшення точності. Оптимізація гіперпараметрів дозволила CNN більш ефективно обробляти часові ряди, зменшуючи вплив короткострокових коливань і зміни в даних. CNN стала більш чутливою до ключових характеристик фінансових даних, що дозволило знизити похибки прогнозу. Однак, навіть з покращеннями, CNN не змогла досягти рівня LSTM, оскільки не має такої ж здатності до роботи з довгостроковими залежностями в часових рядах, що важливо для прогнозування фінансових ринків.

Гібридна модель, що поєднувала LSTM та CNN, також показала покращення після оптимізації гіперпараметрів. Зокрема, значення RMSE знизилася до 16.7227, E до 14.9688, а MAPE зменшилося до 7.37%. Це свідчить про певні переваги комбінованого підходу, однак її точність все ж залишалася меншою порівняно з окремими моделями LSTM та CNN. Поясненням цього може бути складність інтеграції двох різних моделей, що може не завжди призводити до кращих результатів, особливо якщо базові моделі не демонструють виняткових результатів. Гібридна модель могла не повною мірою використати сильні сторони обох підходів, що обмежило її здатність досягти високої точності прогнозування.

Загалом, оптимізація гіперпараметрів призвела до помітних покращень у результатах всіх трьох моделей, зокрема LSTM, яка продовжує демонструвати найкращі результати з усіх моделей, показуючи свою ефективність у прогнозуванні фінансових ринків. CNN також значно покращила свої результати, хоча її точність все ще не досягала рівня LSTM. Гібридна модель, хоча і показала покращення, залишалася менш ефективною, ніж окремі моделі LSTM та CNN.

Рекомендації щодо використання моделі в реальних умовах

На основі результатів експериментів, проведених з різними моделями прогнозування, можна зробити кілька рекомендацій щодо їх використання в реальних умовах фінансових ринків.

Модель LSTM для прогнозування волатильних ринків. Модель LSTM продемонструвала найкращі результати серед усіх розглянутих моделей, як до, так і після оптимізації гіперпараметрів. Вона виявилася найбільш ефективною при роботі з складними, нелінійними залежностями в даних, що є характерним для фінансових часових рядів. Її здатність моделювати довгострокові залежності робить її ідеальним вибором для прогнозування волатильних ринків, де важливі не лише короткострокові зміни, а й довгострокові тренди. Рекомендується використовувати LSTM для побудови прогнозів на основі великих обсягів історичних даних, таких як ціни акцій, обсяги торгів, індекси та інші фінансові показники. Крім того, для досягнення ще вищої точності, слід регулярно проводити оптимізацію гіперпараметрів і адаптувати модель до нових даних.

CNN для короткострокових прогнозів та інтеграції з іншими методами. Модель CNN показала значні покращення після налаштування гіперпараметрів, що свідчить про її потенціал для використання в практичних умовах. Однак, вона все ж поступається LSTM у прогностичних здатностях на довгострокових даних. CNN може бути ефективною для короткострокових прогнозів, де важливі більш короткі залежності та коливання на ринку. Крім того, її можна використовувати як частину комбінованих підходів, зокрема у гібридних моделях, де CNN може відповідати за виявлення більш локальних паттернів у даних, а LSTM за моделювання довгострокових трендів. Тому CNN може бути корисною в поєднанні з іншими моделями для забезпечення більш комплексного підходу до прогнозування.

Гібридні моделі як додатковий інструмент. Гібридна модель, яка поєднує LSTM і CNN, показала покращення після оптимізації гіперпараметрів, але залишалася

менш точною, ніж LSTM. Хоча комбінування двох моделей може бути корисним для більш комплексного аналізу, слід враховувати, що цей підхід не завжди дає значно кращі результати, особливо коли базові моделі не є достатньо сильними. Рекомендується використовувати гібридні моделі лише в тих випадках, коли є необхідність поєднувати переваги різних методів або в умовах, коли необхідно враховувати кілька типів даних або різні часові горизонти прогнозування. Для оптимального використання гібридних моделей важливо проводити ретельну налаштування параметрів кожної окремої моделі, щоб уникнути негативного впливу інтеграції.

Обмеження використання ARIMA та експоненціальних моделей. У даному дослідженні моделі ARIMA та експоненціальна модель продемонстрували значно нижчу ефективність порівняно з LSTM і CNN. Ці моделі мають обмежену здатність до прогнозування волатильних і нелінійних ринків, де важливі складні залежності між даними. ARIMA може бути корисною у випадках, коли ринок має стабільний тренд і не піддається значним коливанням. Експоненціальні моделі можуть використовуватись для простих прогнозів, коли ринок стабільний, і короткострокові зміни є незначними. Проте для більш точних прогнозів на волатильних ринках рекомендується використовувати більш складні підходи, такі як LSTM або CNN.

Необхідність постійної адаптації моделей до нових умов. Оскільки фінансові ринки є динамічними і піддаються швидким змінам, необхідно регулярно оновлювати моделі, проводити їх повторну оптимізацію та налаштування гіперпараметрів. Технології глибокого навчання, такі як LSTM і CNN, дозволяють адаптувати моделі до нових умов, але це вимагає постійної роботи з даними та регулярної перевірки результатів. Це особливо важливо під час великих економічних змін або криз, коли моделі, що працювали добре в попередні періоди, можуть перестати давати точні прогнози.

Використання цих моделей в реальних умовах дозволяє ефективно прогнозувати фінансові ринки, проте важливо враховувати специфіку ринку та постійно адаптувати моделі до нових умов.

ВИСНОВКИ

Актуальність прогнозування цін на акції обумовлена потребами ефективного функціонування фінансових ринків, забезпечення стабільності інвестиційних рішень і сприяння економічному розвитку. Розробка нових методів прогнозування дозволяє знизити невизначеність і приймати обґрунтовані рішення, що є важливим у швидкозмінному економічному середовищі.

У межах даного дослідження проведено глибокий аналіз наукових джерел щодо прогнозування фінансових ринків та висвітлених в них підходів та порівняння різних методів прогнозування цін на акції з використанням як класичних методів прогнозування часових рядів, так і технологій глибокого навчання. Зокрема, досліджено ефективність рекурентних нейронних мереж (RNN), зокрема їх розширення у вигляді моделей LSTM (довгої короткочасної пам'яті), а також згорткових нейронних мереж (CNN) та гібридних моделей, що поєднують властивості цих двох архітектур.

Основною метою дослідження було визначити, яка з цих моделей може найбільш ефективно передбачати ціни на акції, зокрема на фінансові часові ряди, що є складними для прогнозування через свою високу волатильність та залежність від багатьох зовнішніх факторів, таких як економічні показники, політичні зміни та соціальні тренди. Для досягнення цієї мети, в рамках дослідження було реалізовано кілька етапів: вибір і підготовка даних, розробка та налаштування моделей, а також оцінка ефективності прогнозів за допомогою класичних статистичних та сучасних метрик якості.

В рамках даного дослідження перевірена ефективність кожної з моделей у контексті прогнозування фінансових даних, зокрема цін на акції. У процесі експериментів використано три основні архітектури: LSTM, CNN та гібридна модель, яка поєднує елементи обох типів нейронних мереж.

Проведене дослідження порівняння різних моделей прогнозування на основі двох експериментів демонструє значний потенціал використання методів глибокого навчання для прогнозування фінансових ринків. Результати експериментів свідчать, що модель LSTM є найефективнішою серед розглянутих підходів для аналізу часових рядів, зокрема для прогнозування волатильних ринків, де важливими є складні, нелінійні залежності. Високі показники точності моделі LSTM ($RMSE = 3.9792$, $MAE = 3.1032$, $MAPE = 1.55\%$) підтверджують її здатність адаптуватися до даних і забезпечувати точні прогнози навіть у складних умовах ринку.

Моделі CNN, хоча й поступаються LSTM за точністю, показали потенціал до покращення після оптимізації гіперпараметрів. Їхні результати ($RMSE = 10.1205$, $MAE = 8.2054$, $MAPE = 4.03\%$) вказують на можливість використання CNN для короткострокових прогнозів або як частини комбінованих моделей. Гібридна модель, яка поєднує LSTM і CNN, показала покращення, але її точність все ж залишалася нижчою за LSTM, що вказує на складність інтеграції двох моделей із різними характеристиками та підходами до прогнозування.

Моделі ARIMA та експоненціальна модель продемонстрували значно гірші результати порівняно з нейронними мережами і підкресли обмеженість традиційних статистичних методів для роботи з волатильними фінансовими ринками, при прогнозуванні яких потрібно враховувати складні та нелінійні залежності в даних.

У результаті оптимізації гіперпараметрів усі розглянуті моделі глибокого навчання показали значні покращення. Проте модель LSTM залишалася лідером за точністю, що свідчить про її високу здатність до навчання та адаптації до різноманітних умов фінансового ринку. Оптимізація моделей, зокрема CNN та гібридних моделей, також дозволила досягти значних покращень, хоча вони все ще поступаються LSTM за ефективністю.

Загалом, дослідження підтвердило ефективність моделей глибокого навчання, зокрема LSTM, для прогнозування фінансових ринків. Однак для досягнення максимальної точності прогнозів необхідно проводити регулярну оптимізацію моделей, а також враховувати специфіку ринку та динаміку економічних змін. Для

короткострокових прогнозів або у разі обмежених обсягів даних доцільно використовувати CNN, а для комбінованих підходів — гібридні моделі.

Прогнозування цін на акції за допомогою методів глибокого навчання має величезний потенціал у сфері фінансів. Однією з основних можливостей є розробка інвестиційних стратегій. Моделі глибокого навчання, зокрема LSTM, здатні аналізувати великі обсяги історичних даних і виявляти приховані патерни в поведінці ринку. Це дозволяє інвесторам приймати більш обґрунтовані рішення, знижуючи ризики і підвищуючи ефективність інвестицій.

Ще однією важливою можливістю є створення автоматизованих торгових систем, які можуть здійснювати покупки та продажі акцій на основі прогнозів ринку. Завдяки таким системам трейдери можуть зменшити вплив людського фактору та оптимізувати процеси торгівлі. Моделі прогнозування також можуть бути інтегровані в аналітичні платформи, що надають інвесторам точну інформацію про майбутні зміни цін на акції, що сприяє ефективному управлінню портфелем.

Застосування методів глибокого навчання в прогнозуванні дозволяє також краще управляти фінансовими ризиками. Оскільки моделі можуть передбачати негативні коливання ринку, інвестори можуть розробляти стратегії хеджування для мінімізації потенційних втрат. Включення технічних індикаторів, таких як ковзні середні або індекс відносної сили, дозволяє доповнити прогнози важливими фінансовими факторами, що робить їх точнішими і надійнішими.

Використання різних архітектур глибокого навчання, таких як CNN та LSTM, дозволяє здійснювати прогнози на різних часових горизонтах, від короткострокових до довгострокових, що відповідає потребам різних типів інвесторів. Саме завдяки можливості роботи з великими обсягами фінансових даних такий підхід забезпечує детальний аналіз фінансового ринку і дає змогу його учасникам приймати обґрунтовані економічні рішення.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

- ама, Ю. Гіпотеза ефективного ринку: теоретичні основи та практичне застосування / Юджин Фама. – Київ: Наукова думка, 2020. – 256 с
- андельброт, Б. Фрактальний підхід до аналізу фінансових ринків / Бенуа Мандельброт. – Львів: Світ, 2019. – 312 с.
- андельброт, Б. Фрактали та фінансові ринки: небезпечний світ фінансової теорії / Бенуа Мандельброт, Річард Л. Гудман. – Київ: Альтерпрес, 2018. – 432 с.
- амільтон, Дж. Таймсерії в економіці / Джеймс Гамільтон. – Київ: Економічна думка, 2019. – 480 с.
- амільтон, Дж. Таймсерії в економіці / Джеймс Гамільтон. – Київ: Економічна думка, 2019. – 480 с.
- нгл, Р. Моделі умовної гетероскедастичності у фінансовій економіці / Роберт Енгл. – Київ: Видавництво економічної літератури, 2018. – 420 с.
- окс, Дж., Дженкінс, Г. Аналіз часових рядів: прогнозування та контроль / Джордж Бокс, Гвіл'ям Дженкінс. – Львів: Центр наукових досліджень, 2017. – 350
- анеман, Д. Мислення швидко і повільно / Деніел Канеман. – Київ: Наш формат, 2021. – 512 с.
- иллер, Р. Економіка поведінки: як психологія впливає на фінансові рішення / Роберт Шиллер. – Львів: Літопис, 2020. – 368 с.
- Шиллер, Р. Економіка поведінки: як психологія впливає на фінансові рішення / Роберт Шиллер. – Львів: Літопис, 2020. – 368 с.
- Тверські, А., Канеман, Д. Теорія перспектив: аналіз рішень за умов ризику / Амос Тверські, Деніел Канеман. – Київ: Інститут економічних досліджень, 2018. – 290 с.
- ей, К. М. Машинне навчання та аналіз великих даних у фінансових ринках / К. М. Лей. – Київ: Дело, 2021. – 420 с.
- етров, А. Б. Квантова економіка: нові виміри фінансового аналізу / А. Б. Петров. – Харків: Основа, 2022. – 356 с.

- ваненко, С. П. Сучасні технології прогнозування фінансових ринків / С. П. Іваненко. – Львів: Видавництво ЛПУ, 2020. – 284 с.
- оваленко, О. М. Штучний інтелект у фінансовій аналітиці / О. М. Коваленко. – Київ: Наукова думка, 2023. – 412 с.
- оваленко, О. М. Квантові методи в економічному прогнозуванні / О. М. Коваленко. – Харків: ХНУ, 2022. – 312 с.
- етров, С. І. Еволюція методів фінансового прогнозування / С. І. Петров. – Київ: Наукова думка, 2023. – 278 с.
- ваненко, А. Д. Концептуальні підходи до аналізу фінансових ринків / А. Д. Іваненко. – Львів: ЛПУ, 2021. – 256 с.
- етров, С. І. Технічний аналіз фінансових ринків / С. І. Петров. – Київ: Економіка, 2022. – 312 с.
- жерела: 20. Ковальчук, О. М. Поведінковий фінансовий аналіз: психологічні чинники та економічні рішення / О. М. Ковальчук. – Київ: Економіка, 2020. – 350 с.
- Іванов, С. В. Статистичні методи прогнозування у фінансах / С. В. Іванов. – Харків: Інфо-Пак, 2021. – 280 с.
- Павленко, О. М. Моделі прогнозування для фінансових ринків / О. М. Павленко. – Київ: Наукова думка, 2020. – 310 с.
- Ковальчук, В. І. Фінансові моделі та їх застосування в аналізі ризиків / В. І. Ковальчук. – Харків: Права, 2019. – 280 с.
- Левченко, С. П. Методи прогнозування волатильності на фінансових ринках / С. П. Левченко. – Київ: Наукова думка, 2020. – 320 с.
- Іванов, М. В. Основи лінійної регресії в фінансових ринках / М. В. Іванов. – Київ: Книга, 2019. – 285 с.
- Петров, О. І. Лінійна регресія в фінансовому аналізі / О. І. Петров. – Львів: Інститут економіки, 2020. – 312 с.
- Іванов, А. О. Машинне навчання у фінансовому прогнозуванні / А. О. Іванов. – Київ: Фінансовий інститут, 2022. – 350 с.

- етрова, М. В. Великі дані та машинне навчання в фінансах / М. В. Петрова. — Одеса: Економічний університет, 2021. — 270 с.
- Коваленко, І. Л. Автоматизація фінансового прогнозування за допомогою машинного навчання / І. Л. Коваленко. — Харків: Науковий центр фінансових технологій, 2023. — 330 с.
- Андрєєв, С. М. Використання LSTM для прогнозування фінансових ринків / С. М. Андрєєв. — Київ: Економічний науковий центр, 2022. — 325 с.
- Белкін, М. І. Машинне навчання для виявлення шахрайства в фінансових операціях / М. І. Белкін. — Харків: Фінансові технології, 2023. — 290 с.
- Ковальчук, О. М. Проблеми та перспективи застосування глибоких нейронних мереж у фінансовому прогнозуванні / О. М. Ковальчук. — Київ: Фінансові дослідження, 2024. — 312 с.
- Яковлєв, І. Г. Проблеми перенавчання в моделях глибокого навчання для фінансового прогнозування / І. Г. Яковлєв. — Львів: Технології майбутнього, 2023. — 280 с.
- Іванов, О. С. Машинне навчання для фінансового прогнозування / О. С. Іванов. — Київ: Наукова думка, 2022. — 320 с.
- Петров, В. М. Алгоритми машинного навчання для фінансових прогнозів: теорія та практика / В. М. Петров. — Харків: Видавництво університету, 2021. — 250 с.
- Сидоренко, О. В. Методики прогнозування на фінансових ринках за допомогою машинного навчання / О. В. Сидоренко. — Київ: Наукова думка, 2022. — 210 с.
- Коваленко, С. Л. Інтелектуальні методи аналізу фінансових ринків / С. Л. Коваленко. — Одеса: Меридіан, 2020. — 180 с.
- Григор'єв, І. О. Глибоке навчання та його застосування в економічних прогнозах / І. О. Григор'єв. — Львів: Інститут економіки, 2023. — 245 с.
- Іванова, О. В. Алгоритми машинного навчання в аналізі ринкових трендів / О. В. Іванова. — Дніпро: Університет економіки, 2021. — 230 с.
- Чорна, Н. Г. Використання нейронних мереж для прогнозування валютних курсів / Н. Г. Чорна. — Харків: Видавничий центр, 2022. — 198 с.

- Макаренко, П. О. Штучний інтелект в фінансах: алгоритми та моделі / П. О. Макаренко. — Київ: Технополіс, 2021. — 250 с.
- Ковальчук, Р. І. Прогнозування ціни акцій за допомогою рекурентних нейронних мереж / Р. І. Ковальчук. — Київ: Науковий світ, 2020. — 270 с.
- Левченко, А. М. Стратегії машинного навчання для фінансових ринків / А. М. Левченко. — Запоріжжя: ТехноМед, 2021. — 220 с.
- Дмитренко, С. В. Основи фінансового прогнозування за допомогою глибинного навчання / С. В. Дмитренко. — Черкаси: Вища школа, 2020. — 210 с.
- Попов, В. В. Аналіз великих даних для прогнозування на фондових ринках / В. В. Попов. — Київ: Академія наук, 2022. — 195 с.
- Бондаренко, Л. І. Машинне навчання в економічних моделях / Л. І. Бондаренко. — Харків: Парадигма, 2020. — 265 с.
- Василенко, К. П. Поглиблені методи аналізу фінансових ринків за допомогою машинного навчання / К. П. Василенко. — Одеса: Інтерпрес, 2022. — 240 с.
- Ребров, О. С. Застосування алгоритмів градієнтного бустингу в прогнозуванні фондових індексів / О. С. Ребров. — Львів: Науковий центр, 2021. — 230 с.
- Шевченко, О. М. Моделювання фінансових ринків за допомогою нейронних мереж / О. М. Шевченко. — Харків: Глобус, 2020. — 250 с.
- Кравченко, М. О. Теорія та практичне застосування машинного навчання для фінансового прогнозування / М. О. Кравченко. — Київ: Науково-освітній центр, 2023. — 260 с.
- Бондар, С. В. Машинне навчання у фінансових ринках: Теорія та практика / С. В. Бондар. — Харків: Фінансова академія, 2023. — 320 с.
- Литвиненко, А. А. Алгоритми машинного навчання для прогнозування фінансових показників / А. А. Литвиненко. — Одеса: Наукове видання, 2022. — 280 с.
- Руденко, І. В. Методи глибокого навчання в аналізі фінансових даних / І. В. Руденко. — Львів: Академічне видавництво, 2024. — 340 с.
- Бенджіо, Й., Лекус, Ю., & Хінтон, Г. (2015). *Глибоке навчання: нейронні мережі для складних завдань* (Deep Learning: Neural Networks for Complex Tasks). Науковий журнал з машинного навчання

Іванченко В. І., Бабенко О. М. Штучні нейронні мережі та їх застосування. – Київ: Наукова думка, 2018. – 240 с.

ДОДАТКИ

Додаток 1.1

Код виконання дослідження експеримент №1

```

# Завантажуємо дані
def get_stock_data(symbol, start_date, end_date):
    df = yf.download(symbol, start=start_date, end=end_date, progress=False)
    if 'Close' not in df:
        raise ValueError("No 'Close' data available for the given symbol.")
    return df['Close'].values.reshape(-1, 1)

# Підготовка даних
def prepare_data(data, time_steps):
    X, y = [], []
    for i in range(len(data) - time_steps):
        X.append(data[i:(i + time_steps)])
        y.append(data[i + time_steps])
    return np.array(X), np.array(y)

# Створення LSTM моделі
def create_lstm_model(input_shape):
    model = Sequential([
        LSTM(64, return_sequences=True, input_shape=input_shape),
        Dropout(0.2),
        LSTM(32, return_sequences=False),
        Dense(16, activation='relu'),
        Dense(1)
    ])
    model.compile(optimizer='adam', loss='mse', metrics=['mae'])
    return model

# Створення CNN моделі
def create_cnn_model(input_shape):
    model = Sequential([
        Conv1D(filters=64, kernel_size=3, activation='relu', input_shape=input_shape),
        MaxPooling1D(pool_size=2),
        Dropout(0.2),
        Conv1D(filters=32, kernel_size=3, activation='relu'),
        MaxPooling1D(pool_size=2),
        Flatten(),
        Dense(50, activation='relu'),
        Dense(1)
    ])
    model.compile(optimizer='adam', loss='mse', metrics=['mae'])
    return model

```

Продовження Додатку 1.1

```

# Створення гібридної моделі (CNN + LSTM)
def create_hybrid_model(input_shape):
    model = Sequential([
        Conv1D(filters=64, kernel_size=3, activation='relu', input_shape=input_shape),
        MaxPooling1D(pool_size=2),
        Dropout(0.2),
        Conv1D(filters=32, kernel_size=3, activation='relu'),
        MaxPooling1D(pool_size=2),
        LSTM(32, return_sequences=True),
        Dropout(0.2),
        LSTM(32, return_sequences=False),
        Dense(16, activation='relu'),
        Dense(1)
    ])
    model.compile(optimizer='adam', loss='mse', metrics=['mae'])
    return model

def main():
    # Параметри
    symbol = 'AAPL'
    start_date = '2018-01-01'
    end_date = datetime.now().strftime('%Y-%m-%d')
    time_steps = 60

    # Завантаження даних
    try:
        data = get_stock_data(symbol, start_date, end_date)
    except ValueError as e:
        print(e)
        return

    # Нормалізація даних
    scaler = MinMaxScaler(feature_range=(0, 1))
    scaled_data = scaler.fit_transform(data)

    # Підготовка даних
    X, y = prepare_data(scaled_data, time_steps)

    # Розділення на тренувальні та тестові дані
    train_size = int(len(X) * 0.8)
    X_train, X_test = X[:train_size], X[train_size:]
    y_train, y_test = y[:train_size], y[train_size:]

```

Продовження Додатку 1.1

```
if __name__ == "__main__":  
    main()
```

Код виконання дослідження експеримент №2

```
# Завантажуємо дані
def get_stock_data(symbol, start_date, end_date):
    df = yf.download(symbol, start=start_date, end=end_date)
    return df['Close'].values.reshape(-1, 1)

# Підготовка даних
def prepare_data(data, time_steps):
    X, y = [], []
    for i in range(len(data) - time_steps):
        X.append(data[i:(i + time_steps)])
        y.append(data[i + time_steps])
    return np.array(X), np.array(y)

# Створення LSTM моделі
def create_lstm_model(input_shape):
    model = Sequential([
        LSTM(50, return_sequences=True, input_shape=input_shape),
        LSTM(50, return_sequences=False),
        Dense(25),
        Dense(1)
    ])
    model.compile(optimizer='adam', loss='mse')
    return model

# Створення CNN моделі
def create_cnn_model(input_shape):
    model = Sequential([
        Conv1D(filters=64, kernel_size=3, activation='relu', input_shape=input_shape),
        MaxPooling1D(pool_size=2),
        Conv1D(filters=32, kernel_size=3, activation='relu'),
        MaxPooling1D(pool_size=2),
        Flatten(),
        Dense(50, activation='relu'),
        Dense(1)
    ])
    model.compile(optimizer='adam', loss='mse')
    return model
```

Продовження Додатку 1.2

```
# Основна функція
def main():
    # Параметри
    symbol = 'AAPL' # Apple Inc.
    start_date = '2018-01-01'
    end_date = datetime.now().strftime('%Y-%m-%d')
    time_steps = 60 # 60 днів для прогнозування

    # Завантаження даних
    data = get_stock_data(symbol, start_date, end_date)

    # Нормалізація даних
    scaler = MinMaxScaler()
    scaled_data = scaler.fit_transform(data)

    # Підготовка даних
    X, y = prepare_data(scaled_data, time_steps)

    # Розділення на тренувальні та тестові дані
    train_size = int(len(X) * 0.8)
    X_train, X_test = X[:train_size], X[train_size:]
    y_train, y_test = y[:train_size], y[train_size:]

    # Створення та навчання моделей
    lstm_model = create_lstm_model((time_steps, 1))
    cnn_model = create_cnn_model((time_steps, 1))
    hybrid_model = create_hybrid_model((time_steps, 1))

    # Навчання моделей
    lstm_history = lstm_model.fit(X_train, y_train, epochs=50, batch_size=32,
                                  validation_split=0.1, verbose=1)
    cnn_model.fit(X_train, y_train, epochs=50, batch_size=16,
                  validation_split=0.1, verbose=1)
    hybrid_model.fit(X_train, y_train, epochs=50, batch_size=16,
                     validation_split=0.1, verbose=1)
```

Продовження Додатку 1.2

```

# Зворотня трансформація для отримання реальних цін
predictions = {
    'LSTM': scaler.inverse_transform(lstm_pred),
    'CNN': scaler.inverse_transform(cnn_pred),
    'Hybrid': scaler.inverse_transform(hybrid_pred),
    # 'Ensemble': scaler.inverse_transform(ensemble_pred.reshape(-1, 1))
}

# Оцінка результатів
y_test_actual = scaler.inverse_transform(y_test)

for name, pred in predictions.items():
    rmse = np.sqrt(mean_squared_error(y_test_actual, pred))
    mae = mean_absolute_error(y_test_actual, pred)
    mape = np.mean(np.abs((y_test_actual - pred) / y_test_actual)) * 100
    print(f'\n{name} Results:')
    print(f'RMSE: {rmse:.4f}')
    print(f'MAE: {mae:.4f}')
    print(f'MAPE: {mape:.2f}%')

# Візуалізація результатів
plt.figure(figsize=(15, 7))
plt.plot(y_test_actual, label='Actual')
for name, pred in predictions.items():
    plt.plot(pred, label=f'{name} Prediction')
plt.title('Stock Price Predictions Comparison')
plt.xlabel('Time')
plt.ylabel('Price')
plt.legend()
plt.show()

if __name__ == "__main__":
    main()

```

Модель ARIMA

```
import numpy as np
import pandas as pd
import yfinance as yf
from datetime import datetime
from statsmodels.tsa.arima.model import ARIMA
from sklearn.metrics import mean_squared_error, mean_absolute_error
import matplotlib.pyplot as plt

def load_stock_data(symbol, start_date, end_date):
    """
    Завантаження історичних даних про акції
    """
    df = yf.download(symbol, start=start_date, end=end_date)
    return df['Close']

def prepare_data(data):
    """
    Підготовка даних для моделі ARIMA
    """
    return data.values

def train_arima_model(train_data, order=(1,1,1)):
    """
    Тренування моделі ARIMA
    """
    model = ARIMA(train_data, order=order)
    model_fit = model.fit()
    return model_fit

def forecast_stock_prices(model_fit, test_data, time_steps):
    """
    Прогнозування цін на акції
    """
    forecast = model_fit.forecast(steps=time_steps)
    return forecast

def calculate_metrics(actual, forecast):
    """
    Розрахунок метрик точності прогнозування
    """
    rmse = np.sqrt(mean_squared_error(actual, forecast))
    mae = mean_absolute_error(actual, forecast)
    mape = np.mean(np.abs((actual - forecast) / actual)) * 100

    return {
        'RMSE': rmse,
        'MAE': mae,
        'MAPE': mape
    }
```

Продовження Додатку 1.3

```

def plot_results(actual, forecast, symbol):
    """
    Графічна візуалізація результатів
    """
    plt.figure(figsize=(12, 6))
    plt.plot(actual, label='Реальні ціни')
    plt.plot(forecast, label='Прогноз', color='red')
    plt.title(f'Прогнозування акцій {symbol}')
    plt.xlabel('Дата')
    plt.ylabel('Ціна акції')
    plt.legend()
    plt.show()

def main():
    # Параметри
    symbol = 'AAPL' # Apple Inc.
    start_date = '2018-01-01'
    end_date = datetime.now().strftime('%Y-%m-%d')
    time_steps = 60 # 60 днів для прогнозування

    # Завантаження даних
    stock_data = load_stock_data(symbol, start_date, end_date)

    # Поділ на навчальну та тестову вибірки
    train_size = len(stock_data) - time_steps
    train_data, test_data = stock_data[:train_size], stock_data[train_size:]

    # Підготовка даних
    train_values = prepare_data(train_data)
    test_values = prepare_data(test_data)

    # Тренування моделі
    model_fit = train_arima_model(train_values)

    # Прогнозування
    forecast = forecast_stock_prices(model_fit, test_values, time_steps)

    # Розрахунок метрик
    metrics = calculate_metrics(test_values, forecast)

    # Друк результатів
    print("Метрики прогнозування:")
    for metric, value in metrics.items():
        print(f"{metric}: {value}")

    # Графічна візуалізація
    plot_results(test_values, forecast, symbol)

if __name__ == "__main__":
    main()

```

Експоненційне згладжування

```
def main():  
    # Параметри  
    symbol = 'AAPL'  
    start_date = '2018-01-01'  
    end_date = datetime.now().strftime('%Y-%m-%d')  
  
    # Завантаження даних  
    try:  
        data = get_stock_data(symbol, start_date, end_date)  
    except ValueError as e:  
        print(e)  
        return  
  
    # Прогнозування  
    exponential_smoothing_forecast(data)  
  
if __name__ == "__main__":  
    main()
```

Короткий звіт подібності



Дата звіту11/30/2024

Дата редагування —



Звіт не був оцінений.

Звіт подібності

метадані

Заголовок

Зубова_Юнькова_плагіат

Автор

Науковий керівник / Експерт

Зубова

Юнькова

підрозділ

кафедра математичного моделювання та статистики

Тривога

У цьому розділі ви знайдете інформацію щодо текстових спотворень. Ці спотворення в тексті можуть говорити про **МОЖЛИВІ** маніпуляції в тексті. Спотворення в тексті можуть мати навмисний характер, але частіше характер технічних помилок при конвертації документа та його збереженні, тому ми рекомендуємо вам підходити до аналізу цього модуля відповідально. У разі виникнення запитань, просимо звертатися до нашої служби підтримки.

Заміна букв		18
Інтервали		0
Мікропробіли		80
Білі знаки		0
Парафрази (SmartMarks)		10

Обсяг знайдених подібностей

Коефіцієнт подібності визначає, який відсоток тексту по відношенню до загального обсягу тексту було знайдено в різних джерелах. Зверніть увагу, що високі значення коефіцієнта не автоматично означають плагіат. Звіт має аналізувати компетентна / уповноважена особа.

1.34%

1.34%

КП 1

25

Довжина фрази для коефіцієнта подібності 2

0.24%

0.24%

КП 2

18104

Кількість слів

0.24%

0.24%

КЦ

140369

Кількість символів

Подібності за списком джерел

Нижче наведений список джерел. В цьому списку є джерела із різних баз даних. Колір тексту означає в якому джерелі він був знайдений. Ці джерела і значення Коефіцієнту Подібності не відображають прямого плагіату. Необхідно відкрити кожне джерело і проаналізувати зміст і правильність оформлення джерела.

10 найдовших фраз		Копію тексту	
ПОРЯДКОВИЙ НОМЕР	НАЗВА ТА АДРЕСА ДЖЕРЕЛА URL (НАЗВА БАЗИ)	КІЛЬКІСТЬ ІДЕНТИЧНИХ СЛІВ (ФРАГМЕНТІВ)	
1	https://ir.kneu.edu.ua/server/api/core/bitstreams/d83820f4-7feb-40c7-ae6d-b4388a256d2d/content	44	0.24 %
2	Формування та управління портфелем цінних паперів 6/8/2020 Kyiv National Economic University named after Vadym Hetman KNEU (кафедра економіко-математичного моделювання)	14	0.08 %
3	https://ela.kpi.ua/bitstreams/18e3b753-9611-4083-8354-9c4f176b892b/download	14	0.08 %

Копія наукових тез

Зубова М.С.

Електронна пошта: *ukr.marianna7@gmail.com*

Науковий керівник : к.ф.-м.н., доцент **Юнькова О.О.**

Електронна пошта: *olenayunkova@gmail.com*

Київський Національний Економічний Університет імені Вадима Гетьмана

**ПІДВИЩЕННЯ ТОЧНОСТІ ПРОГНОЗУВАННЯ ЦІН АКЦІЙ
ШЛЯХОМ КОМБІНУВАННЯ РІЗНИХ АРХІТЕКТУР ГЛИБОКОГО
НАВЧАННЯ**

**IMPROVING STOCK PRICE PREDICTION ACCURACY THROUGH
COMBINATION OF DIFFERENT DEEP LEARNING ARCHITECTURES**

В сучасних умовах точне прогнозування цін акцій є критично важливим завданням для всіх суб'єктів фінансового ринку: інвесторів, трейдерів та фінансових установ. Традиційні методи прогнозування в багатьох випадках не забезпечують достатню точність прогнозування через високу волатильність ринку та складність врахування численних факторів впливу. Альтернативою їм стали методи машинного навчання, а також комбінування різних архітектур глибокого навчання, які значно розширили можливості дослідників і утворили перспективний напрямок для підвищення точності прогнозування.

Мета дослідження полягає у розробці та оцінюванні ефективності гібридного підходу до прогнозування цін акцій, що базується на комбінуванні

Продовження Додатку 3

різних архітектур глибокого навчання, зокрема LSTM (Long Short-Term Memory), CNN (Convolutional Neural Networks), ансамблів та гібридної моделі (LSTM +

Методи та результати дослідження

Методологія:

- Використано історичні дані акцій APPLE за 2018-2023 роки
- Розроблено та протестовано чотири архітектури:

STM для аналізу часових рядів

NN для виявлення патернів

нсамблева модель (комбінація прогнозів)

ібридна модель (інтеграція LSTM та CNN)

Результати: Порівняльні метрики моделей:

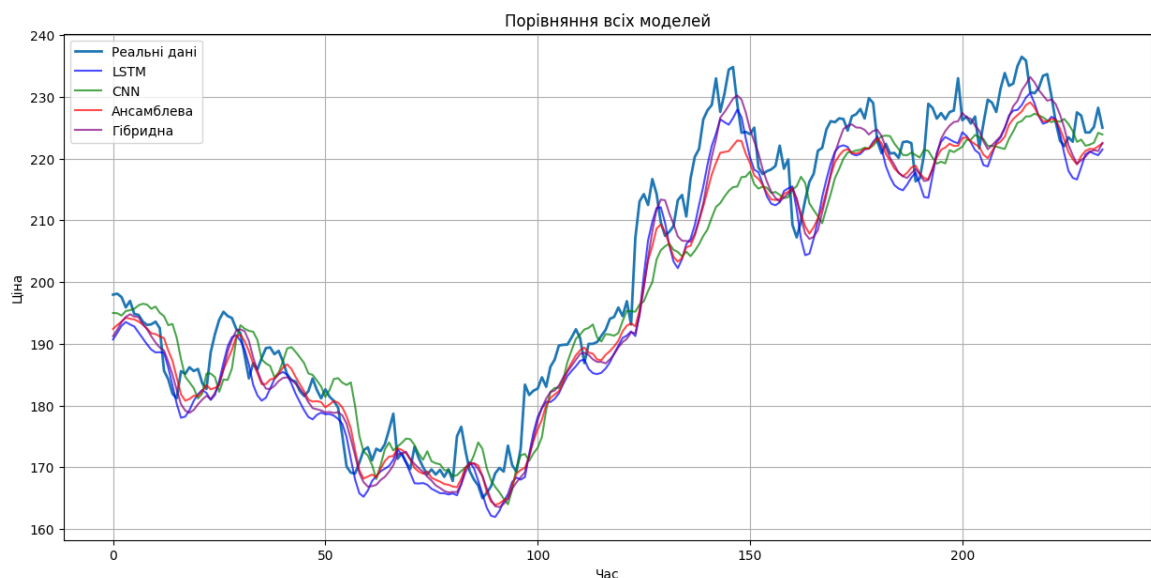


Рис 1. Порівняння моделей

- Гібридна модель продемонструвала найкращу точність за всіма метриками оцінки (MAPE, MAE, RMSE)

показала другий найкращий результат з невеликим відривом від гібридної моделі

- Ансамблева модель зайняла третє місце за точністю прогнозування

Продовження Додатку 3

- CNN виявила найнижчу ефективність серед усіх тестованих архітектур

Порівняльна таблиця метрик:			
Модель	MAPE (%)	MAE	RMSE
LSTM	2.46	5.01	5.94
CNN	2.66	5.47	7.06
Ансамблева	2.32	4.79	5.91
Гібридна	2.18	4.42	5.50

Рис 2. Метрики кожної з моделей

Висновок: Гібридна модель продемонструвала найвищу точність прогнозування серед усіх протестованих архітектур, що робить її оптимальним вибором для прогнозування цін акцій APPLE.

Ключові слова: глибоке навчання, LSTM, CNN, ансамблеві методи, прогнозування часових рядів, машинне навчання, метрики точності, RMSE, MAE,

Література:

1. Петренко В.А., Ковальчук С.М. Застосування LSTM-мереж для прогнозування фінансових часових рядів. Вісник Київського національного університету технологій та дизайну. 2023. Вип. 4(153). С. 82-91.

2. Іванченко О.Д., Мельник Р.В. Порівняльний аналіз архітектур нейронних мереж у задачах економічного прогнозування. Науковий вісник Чернівецького університету. Серія: Комп'ютерні системи та компоненти. 2024. Том 13(1). С. 24-33.

3. Коваленко Т.П., Шевченко М.О. Ансамблеві методи машинного навчання в прогнозуванні економічних показників. Економіка і прогнозування. 2023. №2. С.