

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
Західноукраїнський національний університет
Факультет комп'ютерних інформаційних технологій
Кафедра інформаційно-обчислювальних систем і управління

ЛУЧИШИН Богдан Андрійович

**Метод розробки Discord-бота на основі машинного навчання
для аналізу фондового ринку / Method for Developing a
Discord Bot Based on Machine Learning for Stock Market
Analysis**

спеціальність: 122 - Комп'ютерні науки
освітньо-професійна програма - Комп'ютерні науки

Кваліфікаційна робота

Виконав студент групи КНм-21
Б.А. Лучишин

Науковий керівник:
к.т.н., доцент П.Є. Биковий

Кваліфікаційну роботу
допущено до захисту:

«___» _____ 20___ р.

В.о. завідувача кафедри

_____ Н.М. Васильків

ТЕРНОПІЛЬ - 2024

Факультет комп'ютерних інформаційних технологій
Кафедра інформаційно-обчислювальних систем і управління
Освітній ступінь «магістр»
спеціальність: 122 – Комп'ютерні науки
освітньо-професійна програма – Комп'ютерні науки

ЗАТВЕРДЖУЮ
Завідувач кафедри
_____ М.П. Комар
«_____» _____ 20__ р.

**ЗАВДАННЯ
НА КВАЛІФІКАЦІЙНУ РОБОТУ СТУДЕНТУ**
Лучишин Богдан Андрійович

(прізвище, ім'я, по батькові)

1. Тема кваліфікаційної роботи

Метод розробки Discord-бота на основі машинного навчання для аналізу фондового ринку / Method for Developing a Discord Bot Based on Machine Learning for Stock Market Analysis

керівник роботи к.т.н., доцент П.Є. Биковий,
затверджені наказом по університету від 12 грудня 2023 року № 753.

2. Строк подання студентом закінченої кваліфікаційної роботи 4 грудня 2024 р.

3. Вихідні дані до кваліфікаційної роботи: завдання на кваліфікаційну роботу студента, наукові статті, технічна література.

4. Основні питання, які потрібно розробити

- огляд предметної області фондового ринку;
- аналіз методів машинного навчання для аналізу фондового ринку;
- огляд існуючих рішень для автоматизації аналізу фондового ринку;
- постановка завдання дослідження;
- розробка методології створення Discord-бота для аналізу фондового ринку;
- опис використаних даних та методів їхньої обробки;
- реалізація порівняльного аналізу методів прогнозування фондового ринку;
- тестування Discord-бота та оцінка результатів його роботи;
- аналіз результатів експериментальних досліджень та порівняння з існуючими рішеннями.

5. Перелік графічного матеріалу у роботі

- алгоритм роботи Discord-бота для аналізу фондового ринку;
- графіки результатів порівняння методів аналізу фондового ринку.

6. Консультанти розділів кваліфікаційної роботи

Розділ	Прізвище, ініціали та посада консультанта	Підпис, дата	
		Завдання видав	Завдання прийняв

7. Дата видачі завдання 12 грудня 2023 р.

КАЛЕНДАРНИЙ ПЛАН

№ з/п	Назва етапів кваліфікаційної роботи	Строк виконання етапів кваліфікаційної роботи	Примітка
1	Затвердження теми кваліфікаційної роботи, ознайомлення з літературними джерелами та складання плану роботи.	до 01.01. 2024 р.	
2	Написання 1 розділу кваліфікаційної роботи	до 01.03. 2024 р.	
3	Написання 2 розділу кваліфікаційної роботи	до 20.05.2024 р.	
4	Написання 3 розділу кваліфікаційної роботи	до 28.10. 2024 р.	
5	Представлення попереднього варіанту кваліфікаційної роботи, перевірка та внесення змін керівником	до 11.11.2024 р.	
6	Опрацювання зауважень та представлення завершеного варіанту кваліфікаційної роботи. Підготовка супроводжуючих документів.	до 25.11.2024 р.	
7	Перевірка кваліфікаційної роботи на оригінальність тексту.	до 30.11.2024 р.	
8	Оформлення кваліфікаційної роботи та отримання допуску до захисту	до 04.12.2024 р.	
9	Подання кваліфікаційної роботи до захисту на засіданні атестаційної комісії.	до 14.12. 2024 р.	

Студент _____ Б.А. Лучишин
підпис

Керівник роботи _____ к.т.н., доцент П.С. Биковий
підпис

РЕЗЮМЕ

Кваліфікаційна робота на тему «Метод розробки Discord-бота на основі машинного навчання для аналізу фондового ринку» виконана для здобуття освітнього ступеня «Магістр» за спеціальністю 122 «Комп'ютерні науки» в межах освітньо-професійної програми «Комп'ютерні науки» написана обсягом в 79 сторінок і містить 15 ілюстрації, 2 додатки та 32 використаних джерел.

Метою роботи є створення автоматизованого інструменту у вигляді Discord-бота, здатного здійснювати аналіз і прогнозування фондового ринку за допомогою методів машинного навчання.

Методи дослідження, використані у роботі, методи машинного навчання (Random Forest Regressor, Gradient Boosting Regressor, Support Vector Regressor) та алгоритми інтеграції даних через API.

Результати дослідження: Кваліфікаційна робота включає розробку Discord-бота, який забезпечує інтеграцію отримання фінансових даних через API, їх попередню обробку, аналіз і прогнозування. Порівняльний аналіз точності моделей показав, що ансамблеві методи мають найвищу точність прогнозування з коефіцієнтом детермінації $R^2 = 0.995$.

Практичне значення роботи полягає в здатності бота автоматизувати фінансовий аналіз. Це рішення є корисним для трейдерів, інвесторів та аналітиків для здійснення прогнозів та виявлення трендів на фондовому ринку в реальному часі.

Рекомендації для подальших досліджень включають розширення функціональності бота шляхом впровадження нових методів машинного навчання, оптимізації обробки активів з високою волатильністю та адаптації бота до аналізу інших фінансових ринків.

Ключові слова: DISCORD-БОТ, АНАЛІЗ ДАНИХ, ПРОГНОЗУВАННЯ, МАШИННЕ НАВЧАННЯ, АНСАМБЛЕВІ МЕТОДИ, ФОНДОВИЙ РИНОК, АВТОМАТИЗАЦІЯ АНАЛІЗУ.

ABSTRACT

The qualification work on the topic «Method for Developing a Discord Bot Based on Machine Learning for Stock Market Analysis» for Master's degree on speciality 122 «Computer Science» educational and professional program «Computer Science» is written on 79 pages and it contains 15 figures, 2 annexes and 32 sources.

The aim of the qualification work is to create an automated tool in the form of a Discord bot capable of performing stock market analysis and forecasting using machine learning methods.

Research methods used in the work include systems analysis, machine learning methods (Random Forest Regressor, Gradient Boosting Regressor, Support Vector Regressor), and data integration algorithms via API.

Research results: The qualification work includes the development of a Discord bot that integrates financial data acquisition via API, preprocessing, analysis, and forecasting. A comparative analysis of the model accuracy revealed that ensemble methods showed the highest prediction accuracy, with an R^2 coefficient of 0.995.

Practical significance: The practical significance of the research lies in the bot's ability to automate financial analysis. This solution is valuable for traders, investors, and analysts for forecasting and detecting stock market trends in real-time.

Recommendations for future research include extending the bot's functionality by incorporating new machine learning methods, optimizing the handling of highly volatile assets, and adapting the bot to analyze other financial markets.

Keywords: DISCORD BOT, DATA ANALYSIS, FORECASTING, MACHINE LEARNING, ENSEMBLE METHODS, STOCK MARKET, AUTOMATION OF ANALYSIS.

ЗМІСТ

Вступ.....	7
1 Теоретико-методичні засади аналізу фондового ринку	9
1.1 Аналіз прикладної області.....	9
1.2 Методи машинного навчання для аналізу фондового ринку	14
1.3 Аналіз існуючих рішень для автоматизації аналізу фондового ринку	16
1.4 Постановка завдання.....	18
Висновки до розділу 1	20
2 Методологія розробки Discord-бота та аналіз методів прогнозування фондового ринку	21
2.1 Опис використаних даних	21
2.2 Розробка методології для створення Discord-бота.....	23
2.3 Огляд методів машинного навчання та часових рядів.....	27
2.4 Реалізація порівняльного аналізу	31
2.5 Результати дослідження методів.....	33
Висновки до розділу 2	38
3 Реалізація та тестування Discord-бота для аналізу фондового ринку	40
3.1 Опис архітектури та функціональних можливостей бота	40
3.2 Експериментальне тестування та результати роботи бота	47
3.3 Аналіз та оцінка результатів реалізації.....	55
Висновки до розділу 3	58
Висновки.....	60
Список використаних джерел.....	62
Додаток А Копії публікацій.....	66
Додаток Б Текст програми бота для аналізу фондового ринку	76

ВСТУП

Фондовий ринок є важливим елементом економічної системи, що впливає на розвиток бізнесу, інвестиційні можливості та фінансову стабільність країн. Прогнозування коливань фондового ринку є надзвичайно складним завданням через численні фактори, які можуть впливати на ціноутворення, та високу непередбачуваність фінансових інструментів. Традиційні методи аналізу, такі як технічний і фундаментальний аналіз, здебільшого базуються на історичних даних і вимагають значних витрат часу та ресурсів для досягнення оптимальних результатів. У зв'язку з цим, виникає необхідність в нових підходах до аналізу ринку, здатних автоматизувати цей процес і підвищити точність прогнозів.

Актуальність теми дослідження зумовлена зростаючою важливістю ефективного аналізу фондового ринку та прогнозування його тенденцій, що є ключовим завданням для інвесторів та фінансових аналітиків. Традиційні методи прогнозування, такі як технічний аналіз і фундаментальний аналіз, вже давно використовуються для прийняття рішень на фондових ринках. Однак, із розвитком технологій машинного навчання, з'явилися нові можливості для покращення точності прогнозів, автоматизації процесів і зниження впливу людського фактору. Зокрема, використання Discord-бота, здатного в реальному часі аналізувати дані та надавати прогнози, є перспективним напрямом для досягнення більш ефективних результатів у цій галузі.

Метою даного дослідження є розробка та тестування Discord-бота для аналізу фондового ринку за допомогою методів машинного навчання. Завданнями дослідження є: розробка методології створення Discord-бота, реалізація алгоритмів для аналізу даних у реальному часі, порівняння різних методів прогнозування фондового ринку, а також проведення експериментальних тестувань для оцінки ефективності роботи бота в умовах реальних фінансових даних.

Об'єктом дослідження є процес аналізу фондового ринку за допомогою методів машинного навчання. Предметом дослідження є алгоритми та методи, застосовувані для прогнозування фінансових даних, а також програмне середовище для створення та тестування Discord-бота.

Для досягнення поставленої мети були використані такі методи дослідження, як аналіз існуючих рішень для автоматизації аналізу фондового ринку, методи машинного навчання для прогнозування фінансових даних, а також методи тестування програмного забезпечення.

Наукова новизна дослідження полягає в розробці нового інструменту для автоматизованого аналізу фондового ринку на базі Discord-бота, здатного працювати в реальному часі, що дозволяє значно зменшити час на аналіз та прийняття рішень.

Практичне значення одержаних результатів полягає у можливості застосування розробленого Discord-бота в реальних умовах фінансових ринків для аналізу тенденцій і прогнозування змін, що може стати корисним інструментом для трейдерів, інвесторів та аналітиків.

Апробація результатів дослідження здійснювалася шляхом представлення основних теоретичних положень і практичних результатів на X Всеукраїнській науково-практичній конференції «Інтелектуальні комп'ютерні системи та мережі», що відбулася 5 листопада 2024 року на кафедрі комп'ютерної інженерії Західноукраїнського національного університету (м. Тернопіль, вул. О. Теліги, 8), а також на Інтернет-конференції МНІК «Інформаційне суспільство: технологічні, економічні та технічні аспекти становлення» (випуск 93), яка проходила 12–13 листопада 2024 року за підтримки оргкомітету МНІК (а/с 797, м. Тернопіль, 46005).

1 ТЕОРЕТИКО-МЕТОДИЧНІ ЗАСАДИ АНАЛІЗУ ФОНДОВОГО РИНКУ

1.1 Аналіз прикладної області

Фондовий ринок є складною та динамічною системою, що включає торгівлю акціями, облігаціями та іншими фінансовими інструментами. Він відіграє ключову роль у функціонуванні економіки, забезпечуючи механізм перерозподілу капіталу та залучення інвестицій. Аналітика фондового ринку є важливою складовою для прийняття рішень як індивідуальними інвесторами, так і великими фінансовими установами.

Традиційні методи аналізу фондового ринку включають фундаментальний та технічний аналіз. Фундаментальний аналіз базується на оцінці фінансових показників компаній, таких як прибуток, виручка, рентабельність тощо. Технічний аналіз зосереджується на вивченні історичних цінових даних та обсягах торгів для прогнозування майбутніх цінових рухів [1].

Сучасні методи аналізу фондового ринку з використанням машинного навчання. З розвитком технологій та доступністю великих обсягів даних методи машинного навчання стали невід'ємною частиною сучасного аналізу фондового ринку. Ці методи дозволяють автоматизувати процес аналізу, виявляти приховані закономірності та робити точніші прогнози.

Методи машинного навчання, які використовуються для аналізу фондового ринку, охоплюють декілька основних категорій, кожна з яких застосовується для вирішення специфічних завдань.

Перша категорія — це регресійні моделі. Вони слугують для прогнозування значень, наприклад, ціни акцій, на основі історичних даних. Серед найпоширеніших моделей можна виокремити лінійну регресію, яка дозволяє оцінювати взаємозв'язок між незалежними змінними та залежною змінною; поліноміальну регресію, яка ефективна для нелінійних залежностей; а також регресію з використанням дерев рішень, що може враховувати взаємозв'язки між змінними у складних системах. У дослідженні Ashfaq N., Nawaz Z., Piyas M. (2021) дев'ять різних алгоритмів регресії були оцінені для прогнозування цін акцій за

допомогою таких метрик, як середньоквадратична похибка (MSE) та коефіцієнт детермінації (R^2), що підкреслює їхню ефективність для завдань прогнозування [2].

Друга категорія методів машинного навчання, яка широко застосовується в аналізі фондового ринку, охоплює класифікаційні моделі. Вони використовуються для групування акцій або ринкових умов за різними ознаками, такими як цінові рівні, обсяги торгів, технічні показники тощо. Класифікація є важливою для розпізнавання трендів ринку, зокрема в умовах складності та волатильності фінансових даних. Серед найбільш популярних класифікаційних моделей можна виділити логістичну регресію, яка добре підходить для задач, де є дві категорії (наприклад, прогнозування зростання або падіння ринку). Іншою важливою моделлю є підтримка векторних машин (SVM), яка дозволяє досягати високої точності навіть для даних, що не є лінійно роздільними. SVM допомагає виявляти межі між різними класами (наприклад, між умовами зростання та спадання ринку), навіть коли ці межі складні і не очевидні. Ще однією популярною класифікаційною моделлю є нейронні мережі, які здатні обробляти великі обсяги даних та виявляти складні патерни в них. Нейронні мережі можуть обробляти як лінійні, так і нелінійні залежності, що робить їх надзвичайно ефективними для прогнозування складних економічних та фінансових показників [2].

До третьої категорії методів машинного навчання для прогнозування фондового ринку належать нейронні мережі та глибоке навчання. Глибоке навчання включає такі складні моделі, як рекурентні нейронні мережі (RNN) та згорткові нейронні мережі (CNN), які виявляють складні залежності у великих наборах даних. Рекурентні нейронні мережі особливо корисні для аналізу часових рядів, оскільки вони враховують попередні стани даних, що дозволяє прогнозувати майбутні тенденції на основі історичних даних. Згорткові нейронні мережі (CNN), які традиційно використовуються в обробці зображень, також показали свою ефективність у фінансових додатках завдяки їх здатності виділяти важливі ознаки з великих і складних наборів даних. Ці моделі дозволяють знаходити глибші патерни в даних, що є важливим при роботі з великими обсягами ринкової інформації та допомагають створювати точніші прогнози щодо майбутнього руху акцій та інших фінансових інструментів [2].

Такі підходи дозволяють створювати більш точні моделі прогнозування, які здатні враховувати багатогранність ринкових умов, що робить машинне навчання та глибоке навчання надзвичайно потужними інструментами для фінансового аналізу.

Автоматизація аналізу фондового ринку вимагає використання різноманітних технологій для ефективного збору, обробки та аналізу великих обсягів фінансових даних. Одним із основних підходів до автоматизації є використання інтерфейсів програмування додатків (API), які дозволяють автоматично отримувати актуальні та історичні дані з різних джерел без необхідності ручного збору інформації [3].

API для фінансових даних надають можливість інтеграції з іншими системами та базами даних, що дає змогу отримувати інформацію в реальному часі. Наприклад, за допомогою API можна автоматично запитувати котирування акцій, валютні курси, фінансові звіти компаній, дані з ринків товарів або індексів фондового ринку. Це дозволяє оперативно оновлювати дані та використовувати їх для подальшого аналізу та прогнозування. Автоматизація збору таких даних знижує час, необхідний для їх отримання, що є критичним для ринків, які працюють в реальному часі, таких як фондовий [3].

Що стосується обробки даних, то для автоматизації аналізу часто застосовуються методи машинного навчання, які здатні працювати з великими обсягами інформації. Для цього можуть використовуватися як прості моделі, наприклад, лінійна або поліноміальна регресія, так і складніші моделі, такі як нейронні мережі, які включають рекурентні нейронні мережі (RNN) для роботи з часовими рядами, або згорткові нейронні мережі (CNN), які можуть виявляти складні закономірності в даних. Така автоматизація допомагає не лише зменшити людський фактор у процесі аналізу, але й значно прискорити обробку даних і прийняття рішень.

Крім того, для забезпечення ефективності автоматизації важливим аспектом є інтеграція таких систем з іншими програмними засобами, що дозволяє створювати більш гнучкі та адаптивні рішення для аналізу фондового ринку. Наприклад, можна автоматично генерувати аналітичні звіти, здійснювати прогнози

щодо змін вартості акцій, а також надсилати ці звіти користувачам або іншим системам у реальному часі. Відсутність необхідності вручну здійснювати пошук або обробку даних значно підвищує ефективність роботи аналітиків та інвесторів, знижує час на прийняття рішень і допомагає покращити результати інвестицій.

Завдяки інтеграції з API, можливість отримувати дані в реальному часі дозволяє зменшити час на прийняття рішень у фінансовому середовищі та збільшити точність прогнозів. Інтерфейси API, зокрема, можуть забезпечити високу частоту оновлення даних, що дозволяє миттєво реагувати на зміни на ринку. Така автоматизація має величезний потенціал для розробки складних систем прогнозування та аналітики, які будуть ефективно працювати на великих наборах даних та враховувати мінливі умови ринку.

Автоматизація аналізу фондового ринку за допомогою різноманітних технологій включає не лише використання API для збору та обробки даних, але й інтеграцію з ботами, які здатні виконувати складні завдання в режимі реального часу [4]. Використання ботів для автоматизації є надзвичайно популярним методом, особливо в контексті систем моніторингу фінансових ринків, адже вони можуть інтегруватися з API, працювати з величезними масивами даних і виконувати різноманітні завдання без втручання людини.

Завдяки своїй здатності автоматизувати рутинні завдання, боти можуть значно зменшити час, необхідний для збору і обробки фінансової інформації. Вони здатні здійснювати пошук і аналіз даних, виконувати перевірку актуальності інформації, отримувати оновлення у реальному часі та здійснювати прогнози на основі попередніх даних або алгоритмів машинного навчання. Наприклад, бот може автоматично стежити за змінами на фондовому ринку, порівнювати показники акцій і, на основі заданих критеріїв, робити прогнози про можливі зміни вартості активів [4].

Автоматизація через бота надає можливість миттєвого реагування на події на ринку. Наприклад, якщо ринок змінюється внаслідок певних факторів (новини, корпоративні звіти, політичні події тощо), бот може автоматично збирати необхідну інформацію, аналізувати її за допомогою вбудованих алгоритмів машинного навчання та надавати результат або рекомендації користувачеві в

реальному часі. Всі ці дії здійснюються без втручання людини, що значно збільшує ефективність аналізу та дозволяє інвесторам приймати рішення без затримок.

Особливістю таких ботів є їх висока доступність і здатність працювати на різних платформах. Боти можуть бути розроблені для роботи через месенджери, електронну пошту або інші інтерфейси, що дає змогу зручно взаємодіяти з користувачами. Вони можуть надавати повідомлення та звіти безпосередньо в месенджерах, таких як Telegram або Slack, або через інші канали зв'язку. Це дозволяє інвесторам постійно залишатися в курсі подій на ринку, отримувати рекомендації або прогнози щодо потенційних змін у цінах активів і оперативно реагувати на зміни ситуації.

Однією з найбільших переваг використання ботів для автоматизації аналізу фондового ринку є їх здатність до інтеграції з іншими системами та платформами. Наприклад, бот може бути налаштований для роботи з кількома різними API, що надають фінансові дані, або для роботи з іншими аналітичними інструментами, такими як моделі машинного навчання або алгоритми для обробки великих даних. Така гнучкість дозволяє ботам бути дуже адаптивними і виконувати складні завдання, що потрібні для проведення комплексного аналізу ринку.

Ще однією суттєвою перевагою є їх здатність до автоматичного навчання. Використовуючи алгоритми машинного навчання, боти можуть покращувати свої прогнози та точність аналізу на основі нових даних [4]. Наприклад, якщо бот спочатку використовує базову модель для прогнозування зміни вартості акцій, він може вчитися на історичних даних та оптимізувати свою стратегію відповідно до результатів, що були досягнуті раніше.

Автоматизація через бота також дозволяє виконувати багато задач паралельно, що є дуже важливим для ринку з високим рівнем волатильності, де необхідно постійно стежити за змінами. У цьому контексті боти стають важливим інструментом для розробки автоматизованих торгових стратегій, що можуть працювати в автоматичному режимі, здійснювати торги, базуючись на заздалегідь заданих алгоритмах.

У результаті, автоматизація за допомогою ботів має великий потенціал для розвитку на ринку фінансових послуг, особливо в умовах зростаючої потреби в

швидкому та точному аналізу фінансових даних. Вони значно полегшують завдання збору, обробки та аналізу фінансової інформації, забезпечують високу доступність, дозволяють інтегруватися з іншими системами та підвищують ефективність прийняття рішень на основі отриманих даних.

1.2 Методи машинного навчання для аналізу фондового ринку

Методи машинного навчання (ML) для аналізу фондового ринку є надзвичайно актуальними, оскільки вони допомагають автоматизувати аналіз великих обсягів даних, виявляти складні закономірності і робити точні прогнози про рух ринку. Застосування ML у фінансах охоплює різні підходи, від простих методів регресії до складних нейронних мереж, кожен з яких має свої переваги в залежності від типу задачі.

Лінійна регресія є найпростішим методом, що використовується для прогнозування числових значень, таких як ціна акцій або індекси. Цей метод дозволяє знайти лінійну залежність між вхідними змінними та залежною змінною. Однак на фондовому ринку такі прості моделі часто не дають точних прогнозів, оскільки ринок піддається численним зовнішнім впливам, що створюють нелінійні залежності. Тому, для підвищення точності прогнозів, використовують поліноміальну регресію, яка дозволяє моделювати складніші взаємозв'язки між факторами [5].

У задачах, де необхідно передбачити певну категорію (наприклад, чи зросте чи знизиться ціна акцій), використовують методи класифікації. Логістична регресія є одним з найпоширеніших методів класифікації, де результат прогнозу виводиться в межах певного діапазону ймовірностей. Ще одним потужним методом є метод підтримки векторних машин (SVM), який дозволяє знаходити оптимальні розподіли між різними класами вхідних даних, зменшуючи помилки класифікації. SVM часто застосовуються для прогнозування рухів на фондових ринках, оскільки вони добре працюють з великими і складними наборами даних [6].

Застосування рекурентних нейронних мереж (RNN) є дуже популярним для прогнозування часових рядів, таких як ціни акцій або індекси. Основною

перевагою RNN є здатність моделювати часову залежність між значеннями, тобто враховувати попередні дані для прогнозування майбутнього. Однак, стандартні RNN мають обмеження, оскільки можуть не зберігати інформацію про далекі часові залежності, що є критичним для довгострокового прогнозування. Тому для подолання цих обмежень використовують LSTM (Long Short-Term Memory), який є варіантом RNN, здатним ефективно працювати з довгостроковими залежностями. LSTM-моделі особливо корисні в аналізі фондового ринку, оскільки вони здатні зберігати важливу інформацію протягом тривалого періоду часу, що дозволяє прогнозувати тренди на основі минулих коливань на ринку [7].

Для ще більш складних задач прогнозування застосовують глибинне навчання (Deep Learning), яке включає в себе різноманітні архітектури, зокрема глибокі нейронні мережі. Такі методи дозволяють моделювати складні багаторівневі взаємозв'язки між різними ознаками ринку, такими як економічні показники, новини, соціальні мережі і навіть глобальні тренди. Завдяки здатності автоматично витягувати ознаки з великих обсягів даних, глибинне навчання значно підвищує точність прогнозів. Наприклад, використання Convolutional Neural Networks (CNN) для аналізу графічних даних або Autoencoders для зменшення розмірності і полегшення класифікації може дати важливі результати для ринку цінних паперів [8].

Ансамблеві методи, такі як Random Forest або Gradient Boosting, можуть також використовуватися для покращення прогнозів. Вони комбінують результати декількох моделей для досягнення більш стабільних і точних прогнозів. Random Forest складається з великої кількості рішень дерев, кожне з яких приймає рішення незалежно, а потім результати об'єднуються для отримання фінального прогнозу. Метод Gradient Boosting, зокрема, дозволяє коригувати помилки попередніх моделей, що робить його потужним інструментом для прогнозування фондового ринку [9].

Машинне навчання забезпечує важливі інструменти для аналізу фондового ринку, дозволяючи значно покращити точність прогнозів, знижувати ризики і автоматизувати багато процесів. Поєднання різних методів, таких як регресія,

класифікація, нейронні мережі та ансамблеві моделі, може бути дуже ефективним для створення надійних моделей прогнозування.

1.3 Аналіз існуючих рішень для автоматизації аналізу фондового ринку

Автоматизація аналізу фондового ринку залишається одним із ключових напрямків досліджень у фінансовій галузі. Це пов'язано зі зростаючим обсягом даних, які потрібно обробляти для прийняття ефективних рішень, та потребою у зменшенні людського фактора у процесах трейдингу. В останнє десятиліття з'явилося багато рішень, які використовують інструменти машинного навчання, обробки великих даних та глибокого навчання, що підтверджується результатами низки сучасних досліджень [10, 11].

Одним із популярних підходів є використання алгоритмів машинного навчання (ML), таких як метод опорних векторів (SVM), ансамблеві методи (Random Forest, Gradient Boosting), перцепція тощо. Наприклад, у дослідженні [12] було встановлено, що алгоритми ансамблю демонструють високу стійкість до шумів у даних і можуть використовуватись для прогнозування фінансових показників. У той же час, метод опорних векторів (SVM) показав високу ефективність при аналізі часових рядів, оскільки враховує як лінійні, так і нелінійні залежності у даних [12].

Глибоке навчання (DL) також є перспективним напрямком автоматизації. Особливої уваги заслуговують рекурентні нейронні мережі (RNN) та їхні вдосконалені варіанти, такі як Long Short-Term Memory (LSTM) та Gated Recurrent Unit (GRU). Вони успішно застосовуються для аналізу часових рядів та прогнозування тенденцій фондового ринку, оскільки здатні враховувати як короткострокові, так і довгострокові залежності [13, 14]. Як показано у дослідженні [15], використання LSTM для аналізу фінансових даних дозволило досягти точності понад 80% у прогнозуванні ринкових трендів.

Для аналізу текстових даних, таких як новини та публікації у соціальних мережах, використовуються алгоритми обробки природної мови (NLP). Наприклад, у роботі [16] було проаналізовано використання трансформерів, таких

як BERT, для оцінки ринкових настроїв. Використання цих інструментів дозволяє визначати вплив новин на фондовий ринок у режимі реального часу.

Крім того, існує низка програмних платформ для автоматизації аналізу фондового ринку, серед яких QuantConnect, Alpaca, MetaTrader 5, Bloomberg Terminal, FinnHub тощо.

QuantConnect – це платформа для алгоритмічної торгівлі, що надає користувачам можливість створювати, тестувати та запускати свої стратегії на основі машинного навчання. Вона підтримує різні мовні інтерфейси, такі як C#, Python і F#, та інтеграцію з брокерами для реальної торгівлі. QuantConnect використовує машинне навчання для аналізу ринкових даних, побудови моделей та прогнозування цінових рухів [17, 18].

Alpaca - це API для алгоритмічної торгівлі, який дозволяє трейдерам розробляти та автоматизувати свої торгові стратегії. Alpaca пропонує безкоштовний доступ до реальних ринкових даних та історичних даних, що є необхідним для навчання та тестування моделей машинного навчання. Вона підтримує різні програмні інтерфейси для інтеграції з іншими інструментами [19, 20].

MetaTrader 5 (MT5) є популярною платформою для алгоритмічної торгівлі, що підтримує розробку торгових роботів за допомогою мови програмування MQL5. MT5 забезпечує доступ до широкого спектру ринкових індикаторів та інструментів технічного аналізу, а також можливість інтеграції з Python для розробки складних моделей машинного навчання [21].

Bloomberg Terminal використовується професійними трейдерами для аналізу ринку в режимі реального часу, хоча його висока вартість обмежує доступ до цього інструменту [22].

Використання ботів в Discord стає все більш популярним для автоматизації різних завдань, включаючи аналіз фондового ринку. Discord-боти можуть надавати в реальному часі інформацію про ринок, виконувати автоматичні торгові стратегії та аналізувати дані на основі моделей машинного навчання.

Приклади Discord-ботів:

- Stock Bot – популярний бот для Discord, що надає інформацію про ринкові дані, новини та аналітику [23];
- FinnHub – бот, що використовує машинне навчання для прогнозування цін на акції та надання рекомендацій [24].

1.4 Постановка завдання

Основною метою кваліфікаційної роботи є розробка Discord-бота, який використовує моделі машинного навчання для аналізу та прогнозування тенденцій фондового ринку. Для реалізації цієї мети необхідно розглянути кілька ключових аспектів.

По-перше, необхідно провести аналіз області застосування, щоб зрозуміти важливість автоматизації аналізу фондового ринку у фінансовій галузі. Цей аналіз висвітлить проблеми та складнощі, пов'язані з точним прогнозуванням ринкових трендів.

По-друге, потрібно провести комплексний аналіз методів машинного навчання для аналізу фондового ринку. Різні моделі, такі як регресор випадкового лісу (Random Forest Regressor), регресор посилення градієнта (Gradient Booster Regressor) та регресор опорного вектора (Support Vector Regressor, SVR), розглядаються з точки зору їх теоретичних принципів і методології. Сильні сторони та обмеження кожної моделі оцінені для визначення їхньої придатності для прогнозування фондового ринку.

Завдання також передбачає аналіз існуючих програмних засобів для автоматизації аналізу фондового ринку. Переглядаючи та оцінюючи характеристики, функціональні можливості та зручність використання цих інструментів, можна визначити будь-які прогалини чи обмеження, котрі запропонований Discord-бот має на меті усунути.

На основі цих аналізів сформульована постановка проблеми, яка полягає в розробці Discord-бота, який дасть користувачам змогу отримувати аналітичну інформацію та прогнози фондового ринку з підвищеною точністю та надійністю.

Бот має забезпечити зручний інтерфейс і включати розширені моделі прогнозування та показники оцінки.

Необхідно описати загальну структуру системи з окресленням її основних підсистем і функцій. Особливу увагу варто приділити використаним моделям машинного навчання, оскільки вони складають ядро системи. Крім того, розглянуто інформаційну підтримку для аналізу фондового ринку, щоб забезпечити користувачів відповідними даними та ресурсами.

Завдання полягає в розробці структури Discord-бота, забезпеченні його надійності, ефективності та масштабованості. Інтерфейс користувача буде спроектований таким чином, щоб полегшити використання та забезпечити інтуїтивно зрозумілий доступ до функцій аналізу.

При впровадженні Discord-бота необхідно провести експерименти з реальними даними про фондовий ринок для прогнозування ринкових тенденцій. Результати будуть оцінюватися за допомогою різних показників оцінювання якості, таких як середня квадратична помилка (MSE), середня абсолютна помилка (MAE), оцінка R-квадрат (R^2) та інші. Вони визначатимуть ефективність і точність моделей прогнозування, які використовуються в боті.

Підсумовуючи, необхідно розробити комплексний Discord-бот для аналізу фондового ринку, вирішуючи проблеми та обмеження існуючих інструментів. Бот надасть користувачам розширені можливості прогнозування та цінну інформацію про ринок.

Враховуючи сказане вище, можна зробити наступні висновки:

1. Задача аналізу фондового ринку є досить новою, але вона має суттєві відмінності від задачі аналізу звичайних ринкових даних. Існуючі рішення зазвичай представляють комерційні сервіси, а не деталі реалізації.

2. Для вирішення задачі аналізу використовуються різні методи та моделі, які мають свої переваги та недоліки в залежності від кількості залучених змінних, кількості часових кроків, тривалості прогнозування, вибору методу прогнозування, вибору параметра, який потрібно передбачити.

3. Обрано методи регресії випадкового лісу, регресії посилення градієнта та регресії опорного вектора для аналізу фондового ринку, оскільки вони є широко

визнаними та використовуваними моделями в області машинного навчання та аналізу даних.

Висновки до розділу 1

1. Фондовий ринок є складною динамічною системою, яка потребує різноманітних підходів для аналізу, включаючи фундаментальний, технічний аналіз і сучасні методи машинного навчання.

2. Сучасні технології дозволяють ефективно застосовувати машинне навчання для аналізу фондового ринку, що забезпечує високу точність прогнозування на основі великих обсягів даних.

3. Методи машинного навчання, такі як регресійні моделі, класифікаційні алгоритми та нейронні мережі, демонструють різний рівень ефективності залежно від задач: від прогнозування цін акцій до виявлення ринкових трендів.

4. Автоматизація аналізу фондового ринку із використанням API та ботів значно підвищує швидкість і точність обробки даних, сприяючи прийняттю оперативних та обґрунтованих інвестиційних рішень.

5. Інтеграція технологій машинного навчання та автоматизації створює можливості для розробки гнучких рішень, здатних враховувати складність і мінливість ринкових умов.

2 МЕТОДОЛОГІЯ РОЗРОБКИ DISCORD-БОТА ТА АНАЛІЗ МЕТОДІВ ПРОГНОЗУВАННЯ ФОНДОВОГО РИНКУ

2.1 Опис використаних даних

У межах цього дослідження було використано історичні дані про ціни акцій компанії Apple Inc. (AAPL) за період з 13 грудня 2010 року по 20 травня 2024 року [25]. Дані були отримані з фондового ринку та збережені у форматі CSV (Comma-Separated Values), що є стандартом для зберігання табличних даних. Формат CSV забезпечує зручність у зберіганні та обробці великих обсягів інформації, особливо в задачах аналізу фінансових даних.

Датасет складався з 3381 запису, які містили необхідну інформацію для аналізу динаміки цін акцій протягом зазначеного періоду. У кожному записі було відображено кілька ключових показників, що характеризували торгові операції. Зокрема, показник дати дозволяв відстежувати хронологічну послідовність подій, охоплюючи всі торгові дні протягом періоду дослідження. Ціна відкриття відображала вартість акцій на початку кожного торгового дня, тоді як максимальна та мінімальна ціни визначали цінові межі, досягнуті акціями протягом дня.

Ціна закриття показувала вартість акцій на момент завершення торгового дня, а скоригована ціна закриття враховувала такі корпоративні події, як розділення акцій чи виплата дивідендів, забезпечуючи точніше уявлення про реальну ринкову вартість. Окрім цього, показник обсягу відображав кількість акцій, що були продані або куплені протягом кожного торгового дня, що дозволяло оцінити рівень активності на ринку.

Такий багатокомпонентний набір даних забезпечив комплексну основу для аналізу, дозволяючи не лише вивчати загальні тенденції, але й оцінювати динаміку цін у різних аспектах. Завдяки цьому дослідження отримало можливість провести детальний аналіз, необхідний для побудови моделей прогнозування та оцінки ефективності застосованих методів машинного навчання.

Для аналізу та прогнозування використовувалася ціна закриття (Close) акцій Apple Inc., оскільки вона є основним індикатором, що відображає кінцеву оцінку

ринку на певну дату. Ці дані є важливим ресурсом для фінансового аналізу, трейдингу та інвестицій.

Для реалізації Discord-бота в реальному часі були використані актуальні дані, отримані через Alpha Vantage API. Дані в реальному часі, що включали показники, такі як ціна відкриття, закриття, максимальна та мінімальна ціна акцій, обсяг торгів тощо, були отримані безпосередньо з ринку і використовувалися для прогнозування цін за допомогою ансамблю методів, зокрема Random Forest Regressor, Gradient Boosting Regressor і Support Vector Regressor. Це дозволило боту адаптуватися до змін на ринку та забезпечити оперативне прогнозування на основі реальних даних [25].

Через обмеження безкоштовної підписки на API, дані, які можна отримати, мають ліміт на останні 100 торгових днів. З цієї причини для прогнозування використовувалися лише дані за 100 останніх днів, що також стало обмеженням для точності та обсягу прогнозів.

Процес підготовки даних включав кілька етапів. Спочатку були отримані історичні дані за допомогою HTTP-запитів до Alpha Vantage API. Також використовувалися дані з CSV файлу, що містили інформацію про 3381 торговий день. Проте для тестування Discord-бота було обрано лише останні 100 торгових днів. У разі порівняльного дослідження процес підготовки даних охоплював усі етапи, і дані включали такі показники, як дата, ціна відкриття, ціна закриття, максимальна та мінімальна ціна, а також обсяг торгів.

Далі отримані дані були організовані у формат DataFrame за допомогою бібліотеки pandas, що дозволило зручно маніпулювати ними та проводити подальшу обробку. Наступним кроком стало масштабування даних для нормалізації та підготовки їх до машинного навчання. Всі ознаки були масштабовані, що забезпечило однаковий масштаб для всіх параметрів і допомогло правильно працювати моделям, особливо тим, які використовують методи регресії.

Після цього дані були розділені на навчальний набір (80% всіх даних) та тестовий набір (20% даних). Такий розподіл дозволив оцінити точність моделі на нових, раніше не використаних даних, що є важливим для тестування ефективності моделей і запобігання перенавчанню.

Для обробки даних та побудови моделей було використано середовище програмування Python, з бібліотеками pandas для обробки та маніпулювання даними, matplotlib для візуалізації результатів і scikit-learn для створення та тестування моделей машинного навчання. Ці інструменти дозволили ефективно аналізувати та моделювати динаміку цін акцій з метою прогнозування майбутніх цінових тенденцій [25].

Результати дослідження були опубліковані в статті "Порівняльний аналіз методів машинного навчання та часових рядів для прогнозування фондового ринку: підхід бота Discord" [25]. У статті проведено порівняння моделей машинного навчання та методів часових рядів для прогнозування фондового ринку, включаючи такі моделі, як RandomForestRegressor, GradientBoosting, Support Vector Regressor (SVR) і Long Short-Term Memory (LSTM) [25]. Оцінка ефективності цих моделей була виконана на даних акцій AAPL, і результати показали значну перевагу ансамблевих методів перед LSTM в контексті прогнозування фондового ринку [25].

2.2 Розробка методології для створення Discord-бота

Метою даного дослідження є розробка Discord-бота, який може автоматизовано отримувати та аналізувати дані про ціни акцій, використовуючи методи машинного навчання. Це дозволить трейдерам та інвесторам отримувати актуальну інформацію та прогнозні дані в зручному форматі, що сприятиме прийняттю обґрунтованих рішень на фондовому ринку.

Discord – це цифрова платформа для спілкування, яка надає користувачам можливість обмінюватися текстовими повідомленнями, голосовими та відеодзвінками, а також створювати сервери для спільнот з різноманітними тематичними каналами [26]. Discord широко використовується як у сфері геймінгу, так і в професійних чи освітніх спільнотах завдяки своїй зручності, багатофункціональності та можливості інтеграції з API. Через API Discord можна створювати ботів, які автоматизують різні процеси, включаючи обробку

повідомлень, інтеграцію з іншими платформами та надання користувачам доступу до зовнішніх сервісів [26].

Розробка бота розпочалася зі створення нового бота на платформі Discord та отримання токена доступу, необхідного для автентифікації бота під час взаємодії з Discord API. Токен доступу є унікальним ідентифікатором, який дозволяє боту виконувати дії на платформі Discord від імені користувача.

Ініціалізація бота виконувалася за допомогою бібліотеки discord.py, яка забезпечує зручні інтерфейси для роботи з API Discord. Вказання префіксу команд та інтенів (наборів дозволів) дозволило боту отримувати події з сервера, такі як повідомлення та реакції, що є необхідними для його функціонування.

Для отримання даних про ціни акцій використовувалося фондове API, наприклад, Alpha Vantage.

Alpha Vantage – це широко використовуваний сервіс, що надає доступ до фінансових даних через API (Application Programming Interface) [27]. Він дозволяє отримувати як історичні, так і актуальні дані про акції, валютні курси, індекси та криптовалюти [27]. Це робить його надзвичайно корисним для трейдерів, фінансових аналітиків і розробників додатків, які працюють з фінансовою інформацією.

Основною перевагою Alpha Vantage є широкий спектр доступних даних. Сервіс пропонує інформацію про ціни акцій, включаючи дані про ціни відкриття, закриття, максимальні та мінімальні значення, а також обсяг торгів [27]. Крім того, API підтримує роботу з валютними курсами, криптовалютами та технічними індикаторами, такими як RSI, SMA, EMA та Bollinger Bands [27]. Також сервіс надає фундаментальні показники, які дозволяють аналізувати фінансову звітність компаній [27].

Щоб почати користуватися Alpha Vantage, необхідно зареєструватися на їхній платформі та отримати унікальний API-ключ, який потрібен для авторизації запитів [27]. Цей ключ використовується для інтеграції даних у програми чи аналітичні скрипти, які можуть бути написані мовами Python, JavaScript, R або іншими [27]. У безкоштовній версії доступ до API має певні обмеження: не більше п'яти запитів на хвилину та 500 запитів на день [27]. Крім того, історичні дані для

безкоштовних підписників обмежені останніми 100 торговими днями, що може бути враховано при виборі сервісу для аналізу [27].

Для роботи з API Alpha Vantage надається детальна документація, де пояснюється, як правильно формувати запити та працювати з даними у форматах JSON або CSV [27]. Також сервіс пропонує готові бібліотеки для Python та інших мов, що значно спрощує інтеграцію.

За допомогою HTTP-запитів до відповідного API здійснювався збір даних про ціни акцій для подальшого аналізу. Використання API дозволило отримувати актуальні та історичні дані про ринкові ціни, що є важливими для точного прогнозування.

Отримані з фондового API дані були підготовлені для аналізу. Підготовка даних включала створення структури даних (DataFrame), їх очищення та нормалізацію. Важливим етапом було масштабування даних за допомогою методів нормалізації, що забезпечило однаковий масштаб для всіх ознак, полегшуючи навчання моделей.

Дані були розділені на навчальний та тестовий набори для підготовки моделей машинного навчання. Розподіл на навчальний і тестовий набори забезпечував можливість оцінки точності моделей на нових, невидимих раніше даних, що є важливим для запобігання перенавчанню моделей.

Для прогнозування цін акцій у цьому дослідженні було застосовано кілька моделей машинного навчання, кожна з яких має свої особливості та переваги [28]. Зокрема, Random Forest Regressor використовується як метод ансамблевого навчання, що залучає множину дерев рішень для підвищення точності прогнозів завдяки їх комбінованій роботі. Gradient Boosting Regressor базується на поступовому навчанні слабких моделей, де кожна наступна модель намагається зменшити помилки попередньої, що забезпечує більш точний результат. Support Vector Regressor (SVR), своєю чергою, є методом опорних векторів, який демонструє високу ефективність, особливо при роботі з малими вибірками даних. Використання цих моделей дало змогу оцінити їхню здатність прогнозувати зміни цін акцій та порівняти результати, отримані за допомогою різних підходів.

Моделі були навчені на навчальних даних, після чого здійснювалося прогнозування цін акцій на основі тестових даних. Тренування моделей включало налаштування параметрів та вибір оптимальних гіперпараметрів для підвищення точності прогнозів [28].

Після навчання моделей були здійснені прогнози для тестового набору даних. Результати прогнозування оцінювалися за допомогою метрик точності, таких як середньоквадратична помилка (MSE), середня абсолютна помилка (MAE) та коефіцієнт детермінації (R^2) [28]. Це дозволило об'єктивно оцінити точність та ефективність кожної з моделей.

Результати аналізу та прогнозування були відправлені до Discord-каналу за допомогою розробленого бота. Це забезпечило автоматизацію процесу збору, обробки та аналізу даних, а також надання користувачам актуальної інформації у зручному форматі.

Одним з недоліків даного дослідження є використання лише даних про ціни акцій Apple Inc. для навчання та тестування моделей. Майбутні дослідження можуть бути спрямовані на розширення обсягу даних, включаючи акції інших компаній та інші фінансові індикатори, що може підвищити загальну точність прогнозів.

Крім того, є потенціал для використання більш складних методів машинного навчання та глибокого навчання, таких як гібридні моделі або архітектури, що враховують часові ряди та зовнішні фактори.

Таким чином, розроблений Discord-бот дозволить значно спростити процеси збору та аналізу фінансових даних, забезпечуючи високу точність прогнозування та зручність використання для трейдерів та інвесторів. Подальше покращення та розширення функціональності бота може стати перспективним напрямком для майбутніх досліджень та розробок.

2.3 Огляд методів машинного навчання та часових рядів

Часові ряди є важливим об'єктом дослідження в сучасній науці про дані. Їх аналіз знаходить широке застосування у фінансових дослідженнях, економічних прогнозах, медичних дослідженнях та різних технічних галузях. Використання методів машинного навчання значно розширює можливості роботи з часовими рядами, оскільки дозволяє моделювати складні нелінійні залежності та враховувати вплив різноманітних факторів.

Класичні статистичні методи для аналізу часових рядів включають авторегресійні моделі (AR), моделі ковзного середнього (MA), їхні комбінації у вигляді моделей авторегресії та ковзного середнього (ARMA), а також розширені версії із інтеграційною складовою – ARIMA [29, 30].

Авторегресійні моделі (AR) використовують попередні значення ряду для прогнозування майбутніх даних. Модель $AR(p)$ враховує p попередніх значень, а її загальне рівняння має вигляд [29]:

$$X_t = \alpha + \sum_{i=1}^p \beta_i X_{t-i} + \epsilon_t, \quad (2.1)$$

де X_t – значення ряду на момент часу t ;

α – константа;

β_i – коефіцієнти моделі;

ϵ_t – випадковий шум.

Моделі ковзного середнього (MA) базуються на аналізі помилок попередніх прогнозів. Модель $MA(q)$ використовує q останніх помилок, а її математичне представлення має вигляд [29]:

$$X_t = \mu + \epsilon_t + \sum_{j=1}^q \theta_j \epsilon_{t-j}, \quad (2.2)$$

де μ – середнє значення;

θ_j – коефіцієнти моделі;

ϵ_t – шум.

Комбіновані моделі $ARMA(p, q)$ поєднують переваги AR та MA, що дозволяє враховувати як значення самого ряду, так і попередні похибки. Формула ARMA має наступний вигляд [29]:

$$X_t = \alpha + \sum_{i=1}^p \beta_i X_{t-i} + \epsilon_t + \sum_{j=1}^q \theta_j \epsilon_{t-j}, \quad (2.3)$$

де α , β_i , та θ_j – параметри моделі.

Для роботи з нестационарними часовими рядами застосовуються моделі ARIMA(p,d,q), які включають інтеграційну складову (d), що дозволяє моделювати тренди чи сезонність [30].

Сучасні методи машинного навчання відіграють важливу роль у розширенні можливостей аналізу часових рядів. Вони дозволяють моделювати складні нелінійні залежності, враховувати вплив численних факторів та працювати з великими обсягами даних. Такі методи знаходять застосування у фінансових прогнозах, аналізі кліматичних змін, виявленні медичних аномалій та інших завданнях, де часові ряди є ключовим джерелом інформації. До найбільш поширених підходів у машинному навчанні, що застосовуються для аналізу часових рядів, належать дерева рішень, ансамблеві методи, метод опорних векторів та нейронні мережі.

Дерева рішень є базовим методом, який побудований на ієрархічній структурі ухвалення рішень. Основним принципом цього підходу є поділ даних на вузли, кожен з яких відповідає за певну умову або характеристику. Попри свою простоту та зручність у використанні, дерева рішень мають суттєвий недолік – схильність до перенавчання, особливо при роботі з великими наборами даних.

Ансамблеві методи, зокрема випадковий ліс (Random Forest), градієнтний бустинг (Gradient Boosting) та Support Vector Regressor (SVR), значно покращують ефективність та точність моделей [9]. Випадковий ліс використовує сукупність дерев рішень, що навчаються на різних підмножинах даних та характеристик, з метою зменшення варіативності прогнозів. Це досягається за рахунок агрегування результатів багатьох дерев, що забезпечує стабільність та стійкість до шуму. Градієнтний бустинг, у свою чергу, оптимізує результати послідовно: кожна наступна модель фокусується на помилках, допущених попередніми. Такий підхід дозволяє враховувати складні взаємозв'язки між характеристиками даних.

Support Vector Regressor (SVR) є різновидом методу опорних векторів, адаптованим для задач регресії. Його основний принцип полягає у пошуку

гіперплощини, яка відображає залежності між змінними, мінімізуючи похибки в межах заданого епсилон-параметра. SVR відзначається здатністю працювати з нелінійними даними за допомогою ядрових функцій, наприклад, радіальної базисної функції. У поєднанні з ансамблевими методами SVR дозволяє покращити точність моделі, забезпечуючи більшу стійкість до перешкод та нерегулярностей у даних [9].

Метод опорних векторів (SVM) також є ефективним інструментом для аналізу часових рядів. Цей підхід працює з багатовимірними даними, використовуючи гіперплощини для розділення чи моделювання залежностей між спостереженнями. Використання ядрових функцій у SVM забезпечує можливість враховувати складні нелінійні залежності. Завдяки цій властивості метод успішно застосовується для вирішення завдань як класифікації, так і регресії.

Нейронні мережі, зокрема рекурентні архітектури (RNN), довготривала короткочасна пам'ять (LSTM) та мережі з використанням гейтованих рецидивних модулів (GRU), стали одним із найбільш перспективних підходів до аналізу часових рядів [14, 30]. Ці моделі мають здатність обробляти послідовні дані, зберігаючи довготривалі залежності між спостереженнями. LSTM, наприклад, використовує спеціальні комірки пам'яті, які дозволяють контролювати інформацію, що зберігається, додається чи видаляється, завдяки чому модель здатна обробляти тривалі залежності у часових рядах. GRU є спрощеною альтернативою LSTM, що зменшує обчислювальні витрати, зберігаючи здатність ефективно працювати з довготривалими залежностями [14].

Для аналізу фондового ринку різні методи машинного навчання мають свої переваги та недоліки.

LSTM (Long Short-Term Memory) – це особливий тип рекурентних нейронних мереж, що ефективно працює з довготривалими залежностями в даних [30]. LSTM має перевагу у здатності моделювати складні нелінійні залежності у часових рядах, що є важливим для прогнозування фондового ринку, де ці залежності можуть бути дуже глибокими. Однак, модель потребує значних обчислювальних ресурсів і часу для навчання, що може бути недоліком при роботі з великими обсягами даних. Недоліком також є схильність до перенавчання на малих наборах даних.

RNN (Recurrent Neural Network) є основою для LSTM, але не має такої потужної пам'яті для збереження довготривалих залежностей, як LSTM [30]. Через це RNN може страждати від проблеми «зниклого градієнта», коли важливі залежності з часом втрачаються. Це обмежує її ефективність для аналізу фондового ринку, де довгострокові тенденції мають велике значення. Однак для короткострокових прогнозів RNN може бути корисною.

GRU (Gated Recurrent Unit) є вдосконаленою версією RNN, зберігаючи здатність обробляти довготривалі залежності, але з меншою складністю, ніж LSTM [14]. GRU дає менше обчислювальних витрат, однак може бути менш точним, ніж LSTM, коли йдеться про обробку дуже складних і глибоких часових рядів [14]. В той же час, для практичних задач, таких як фінансове прогнозування, де важлива точність і швидкість, GRU може бути корисною альтернативою.

ARIMA (AutoRegressive Integrated Moving Average) — це класичний статистичний метод для прогнозування часових рядів, який ефективно працює для лінійних даних [29]. Однак, для фондових ринків, де присутні складні нелінійні залежності, ARIMA може показати обмежену точність [29]. Крім того, ARIMA вимагає стабільності та стаціонарності даних, що рідко спостерігається на фондових ринках.

SVM (Support Vector Machines) є потужним методом для моделювання складних залежностей у високорозмірних просторах. В задачах прогнозування фондового ринку SVM здатен обробляти нелінійні дані завдяки використанню ядрових функцій. Однак, SVM може мати високу обчислювальну складність, особливо на великих наборах даних. Це обмежує його застосування в реальних фінансових прогностичних системах.

Ансамблеві методи, такі як Random Forest, Gradient Boosting та Support Vector Regressor (SVR), зазвичай використовуються для покращення точності прогнозів. Random Forest об'єднує результати багатьох дерев рішень, знижуючи варіативність і підвищуючи стабільність прогнозів. Градієнтний бустинг послідовно покращує результати, фокусуючись на помилках попередніх моделей, що робить цей метод дуже ефективним для складних задач, таких як прогнозування фондових ринків.

SVR дозволяє працювати з нелінійними даними, але має схильність до перенавчання, якщо правильно не налаштовані параметри моделі.

Дослідження [31], проведене Aich та Kim, вказує на те, що машинне навчання, включаючи згадані методи, значно підвищує точність прогнозів на фондовому ринку порівняно з традиційними статистичними методами, такими як ARIMA. Це дослідження також вказує, що для найбільш точного прогнозування на фондових ринках ефективно комбінувати різні моделі, застосовуючи ансамблеві підходи, які зменшують вплив випадкових помилок та підвищують стабільність [31].

Таким чином, для ефективного аналізу фондового ринку варто застосовувати комбінацію моделей, де LSTM і GRU можуть бути корисні для довгострокового прогнозування, а ансамблеві методи та SVM — для покращення точності в умовах складних, нелінійних даних.

Сучасні методи машинного навчання надають широкі можливості для роботи з часовими рядами, значно перевершуючи традиційні статистичні підходи за точністю та адаптивністю. Вибір конкретного методу залежить від характеру даних, складності завдання та вимог до точності прогнозування.

2.4 Реалізація порівняльного аналізу

У цьому підрозділі детально розглянуто процес реалізації порівняльного аналізу методів прогнозування цін акцій компанії Apple Inc. (AAPL) [25]. Як зазначено у підрозділі 2.1, використано історичні дані про ціни акцій за період з 13 грудня 2010 року до 20 травня 2024 року, що охоплюють значний часовий проміжок і містять 3381 запис. Такий обсяг даних дозволяє забезпечити як точність аналізу, так і надійність отриманих результатів. У цьому дослідженні здійснювався аналіз із використанням двох основних підходів: методу Long Short-Term Memory (LSTM) та ансамблю моделей машинного навчання [25]. Кожен із них має свої унікальні особливості, які дозволяють оцінювати ефективність методів прогнозування у різних умовах.

Для забезпечення коректного порівняння методів важливим етапом стала підготовка даних. Первинний набір даних містив такі показники, як дата, відкрита ціна, найвища ціна, найнижча ціна, ціна закриття, скоригована ціна закриття та обсяг торгів. Вибір цих показників був обумовлений їхньою значущістю для прогнозування часових рядів.

Дані було приведено до форматів, які забезпечують їхню сумісність із методами аналізу. Зокрема, стовпець із датами був перетворений до формату `datetime`, що забезпечило коректну індексацію та сортування. Після цього для зменшення впливу масштабів окремих ознак на процес навчання моделей було застосовано нормалізацію за допомогою `MinMaxScaler`, яка перетворює дані у діапазон від 0 до 1. Це сприяло зниженню ризику домінування значень із високими варіаціями над іншими ознаками.

Окрім того, для підготовки даних були використані техніки, спрямовані на розбиття часових рядів на вхідні сегменти (фрагменти даних, які слугували входом для моделі) та вихідні значення (цільові прогнозовані показники). Такий підхід дозволив врахувати вплив попередніх періодів на подальшу динаміку цін.

Метод LSTM базується на рекурентних нейронних мережах, які здатні моделювати залежності між даними на довгих часових відрізках. На першому етапі було сформовано набір даних для навчання, валідації та тестування. При цьому, враховуючи властивості часових рядів, було дотримано хронологічного порядку, що виключає інформаційний витік між підмножинами даних.

Модель LSTM була навчена із використанням двох варіантів кількості епох – 200 та 2000 [25]. Це дало змогу оцінити вплив кількості навчальних циклів на якість прогнозів. Під час навчання застосовувалися функції активації `tanh` та `sigmoid`, а оптимізація виконувалася з використанням алгоритму `Adam`, який забезпечує швидке збіження та ефективність у задачах прогнозування часових рядів.

Для оцінки точності прогнозів були використані метрики середньоквадратичної помилки (MSE), середньої абсолютної помилки (MAE) та коефіцієнта детермінації (R^2) [25, 30]. MSE дозволяє оцінити середнє квадратичне відхилення між прогнозами та реальними значеннями, що важливо для задач, де критичне значення мають великі відхилення [25, 30]. MAE характеризує середню

абсолютну помилку, що є зручним для інтерпретації результатів, а R^2 показує, яку частку варіацій у даних пояснює модель [25, 30].

Другий підхід передбачав використання ансамблевих моделей, які об'єднують прогнози кількох алгоритмів для підвищення точності. У цьому дослідженні були задіяні три моделі: RandomForestRegressor, GradientBoostingRegressor та Support Vector Regressor (SVR) [25].

На першому етапі кожна модель навчалася окремо на навчальному наборі даних. RandomForestRegressor використовував ансамбль дерев рішень, що забезпечувало стійкість до переобучення та високу точність. GradientBoostingRegressor будував моделі послідовно, кожна наступна зосереджувалася на помилках попередніх, що підвищувало загальну точність. Модель SVR, у свою чергу, забезпечувала високу ефективність на малих вибірках завдяки застосуванню ядра RBF для пошуку нелінійних залежностей.

Після цього прогнози окремих моделей були об'єднані за допомогою VotingRegressor, що дозволило зважено врахувати результати кожної моделі. Оцінювання ансамблю виконувалося за допомогою тих самих метрик, що й для LSTM, забезпечуючи коректне порівняння. Аналіз результатів, детально розглянутий у підрозділі 2.5, дозволив визначити, який метод є більш ефективним для прогнозування цін акцій AAPL, враховуючи різні аспекти точності та стабільності прогнозів.

2.5 Результати дослідження методів

У цьому підрозділі представлені результати прогнозування цін акцій компанії Apple Inc. (AAPL) за допомогою двох підходів: Long Short-Term Memory (LSTM, довга короткострокова пам'ять) та ансамблю моделей машинного навчання. Аналіз включає оцінювання ефективності кожного методу за різними метриками, такими як середньоквадратична помилка (Mean Squared Error, MSE), середня абсолютна помилка (Mean Absolute Error, MAE), коефіцієнт детермінації (R-squared, R^2), відсоткова зміна (Percentage Change), тренд (Trend) та волатильність (Volatility). Для більш глибокого розуміння точності прогнозів

результати ілюструються графіками, які демонструють відхилення прогнозованих значень від фактичних.

Метод LSTM був реалізований з різною кількістю епох (навчальних циклів моделі), що дозволило порівняти вплив глибини навчання на точність прогнозів. У таблиці наведено результати для 200 та 2000 епох.

Результати для LSTM моделі (200 епох):

- Stock Symbol (Символ акції): AAPL
- Forecast Period (Період прогнозування): 676 днів
- Mean Squared Error (Середньоквадратична помилка, MSE): 20412.54
- Mean Absolute Error (Середня абсолютна помилка, MAE): 142.79
- R-squared (Коефіцієнт детермінації, R^2): -2371406.73
- Accuracy Description (Опис точності): Модель не є дуже точною.
- Percentage Change (Відсоткова зміна): +4.85%
- Trend (Тренд): Висхідний (Upward)
- Volatility (Волатильність): 4.95

Результати для LSTM моделі (2000 епох):

- Stock Symbol (Символ акції): AAPL
- Forecast Period (Період прогнозування): 676 днів
- Mean Squared Error (Середньоквадратична помилка, MSE): 20705.49
- Mean Absolute Error (Середня абсолютна помилка, MAE): 143.73
- R-squared (Коефіцієнт детермінації, R^2): -2405440.16
- Accuracy Description (Опис точності): Модель не є дуже точною.
- Percentage Change (Відсоткова зміна): -8.00%
- Trend (Тренд): Низхідний (Downward)
- Volatility (Волатильність): 6.85

Результати показують, що метод LSTM у даному дослідженні не продемонстрував високої точності прогнозування. Наприклад, середньоквадратична помилка (MSE) для моделі з 200 епохами склала 20412.54, а середня абсолютна помилка (MAE) — 142.79. Збільшення кількості епох до 2000 призвело до незначного зростання цих показників, що може свідчити про переобучення моделі.

Коефіцієнт детермінації (R^2), який у випадку точності моделі має наближатися до 1, отримав значення -2371406.73 для моделі з 200 епохами та -2405440.16 для моделі з 2000 епохами, що вказує на невідповідність прогнозів реальним даним.

Крім того, показники відсоткової зміни ціни (+4.85% для 200 епох і -8.00% для 2000 епох) та тренд (висхідний для 200 епох і низхідний для 2000 епох) свідчать про значну варіативність у результатах. Волатильність (4.95 для 200 епох і 6.85 для 2000 епох) також демонструє нестабільність прогнозів.

Ансамбль моделей машинного навчання був використаний для об'єднання прогнозів трьох алгоритмів: Random Forest Regressor (регресор на основі випадкового лісу), Gradient Boosting Regressor (регресор градієнтного підсилення) та SVR (підтримуючий векторний регресор). Цей підхід дозволяє ефективно поєднати переваги кожного алгоритму, мінімізувати їх слабкі сторони та отримати стабільніші й точніші прогнози.

У цьому дослідженні використано ансамбль моделей із VotingRegressor (регресор голосування), що комбінує прогнози, зважуючи результати кожної моделі для досягнення оптимальної точності. Аналіз ефективності ансамблю виконано за допомогою тих самих метрик, що й для методу LSTM, що забезпечує коректне порівняння результатів.

Результати прогнозування за допомогою ансамблю моделей:

- Total days of data (Загальна кількість днів даних): 3381
- Stock Symbol (Символ акції): AAPL
- Forecast Period (Період прогнозування): 677 днів
- Mean Squared Error (Середньоквадратична помилка, MSE): 15.02
- Mean Absolute Error (Середня абсолютна помилка, MAE): 2.58
- R-squared (Коефіцієнт детермінації, R^2): 0.995
- Accuracy Description (Опис точності): Модель є надзвичайно точною.
- Percentage Change (Відсоткова зміна): +697.23%
- Trend (Тренд): Висхідний (Upward)
- Volatility (Волатильність): 56.15

Результати ансамблю моделей демонструють значну перевагу над методом LSTM. Середньоквадратична помилка (MSE) склала лише 15.02, тоді як середня абсолютна помилка (MAE) — 2.58, що вказує на високу точність прогнозів.

Коефіцієнт детермінації (R^2), який має наближатися до 1 для точної моделі, становить 0.995. Це свідчить про те, що ансамбль моделей може пояснити майже всю варіативність у даних, тобто його прогнози максимально наближені до реальних значень.

Високий відсоток зміни ціни (+697.23%) та висхідний тренд (Upward) вказують на позитивні прогнози щодо динаміки акцій компанії AAPL. Однак значна волатильність (56.15) демонструє потенційні ризики, пов'язані з непередбачуваними коливаннями цін.

Таким чином, ансамбль моделей, показав себе як високоефективний інструмент для прогнозування цін акцій. Результати чітко демонструють переваги методу ансамблю моделей над методом LSTM. Навіть при збільшенні кількості епох, LSTM не зміг забезпечити точні прогнози, тоді як ансамбль моделей досягнув високої точності та стабільності. Негативні значення R^2 для LSTM вказують на те, що моделі не змогли уловити основні тенденції в даних, тоді як ансамбль моделей, навпаки, продемонстрував здатність адекватно моделювати ці тенденції.

Графіки прогнозів підтверджують ці висновки. На рисунках 2.1, 2.2 та 2.3 відображено результати прогнозування цін акцій Apple Inc. (AAPL) за допомогою двох підходів: методу LSTM та ансамблю моделей машинного навчання.

На рисунку 2.1 наведено результати прогнозування, отримані за допомогою методу LSTM із використанням 200 епох навчання. Як видно з графіка, прогнозовані значення суттєво відхиляються від фактичних, що свідчить про недостатню точність моделі за таких налаштувань. Це дає підстави припустити, що при меншій кількості епох модель не здатна повною мірою врахувати складні закономірності у даних, що може бути наслідком недостатнього навчання. Таким чином, 200 епох виявилось недостатньо для досягнення високої точності прогнозування.

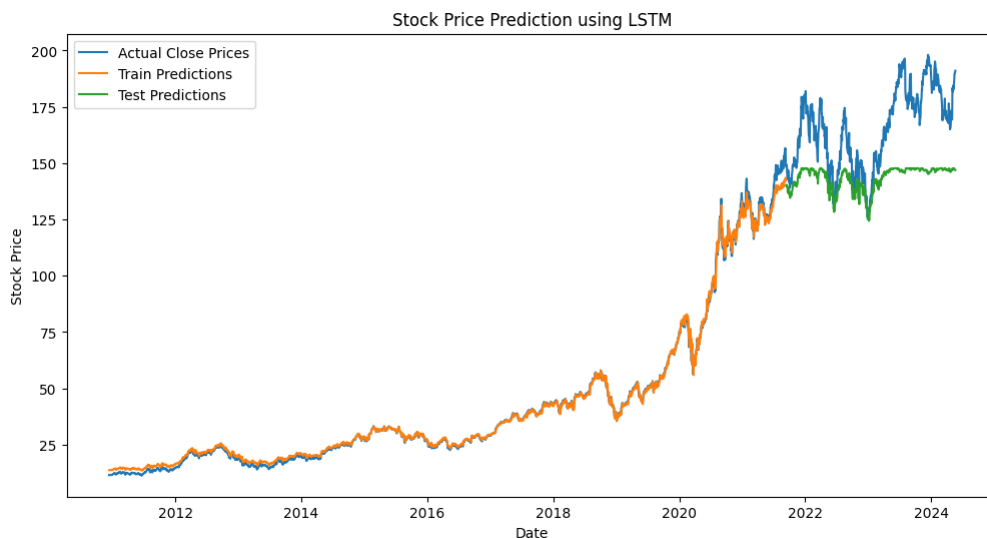


Рисунок 2.1 – Прогнозування цін акцій AAPL за методом LSTM (200 епох)

На рисунку 2.2 показано результати прогнозування тією ж моделлю LSTM, але вже з використанням 2000 епох. Хоча очікувалося, що збільшення кількості епох покращить точність, результати демонструють ще більші відхилення від реальних значень. Така поведінка може бути зумовлена явищем перенавчання, коли модель втрачає здатність узагальнювати отримані дані для нових випадків.

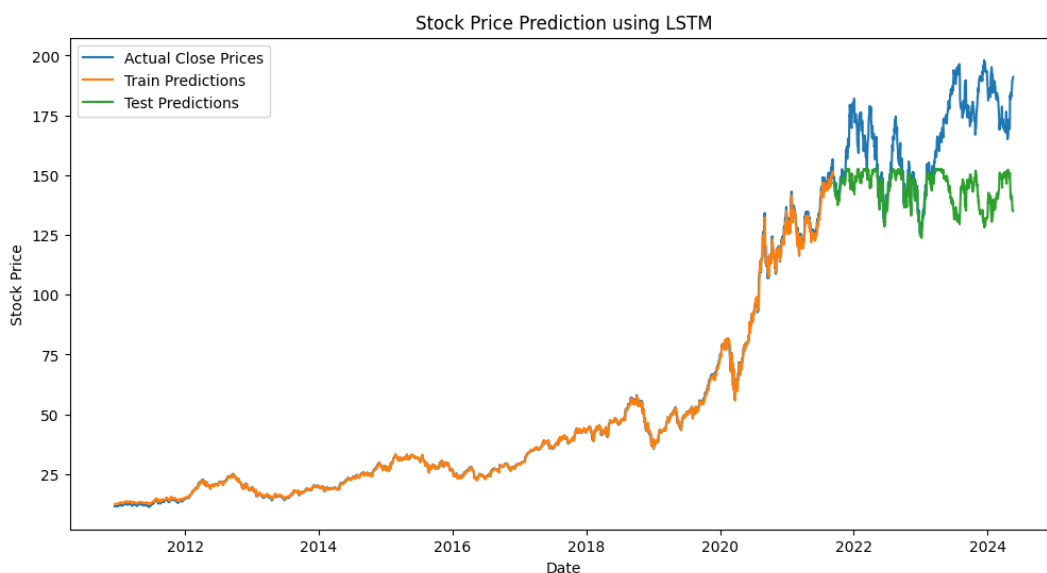


Рисунок 2.2 – Прогнозування цін акцій AAPL за методом LSTM (2000 епох)

На рисунку 2.3 наведено результати прогнозування, отримані за допомогою ансамблю моделей, який поєднує прогнози декількох алгоритмів. Прогнозування

за цим підходом значно краще узгоджується з фактичними значеннями, що свідчить про високу точність та надійність методу. Ансамбль моделей демонструє здатність ефективно враховувати складні взаємозв'язки у даних і суттєво зменшує помилки прогнозування.

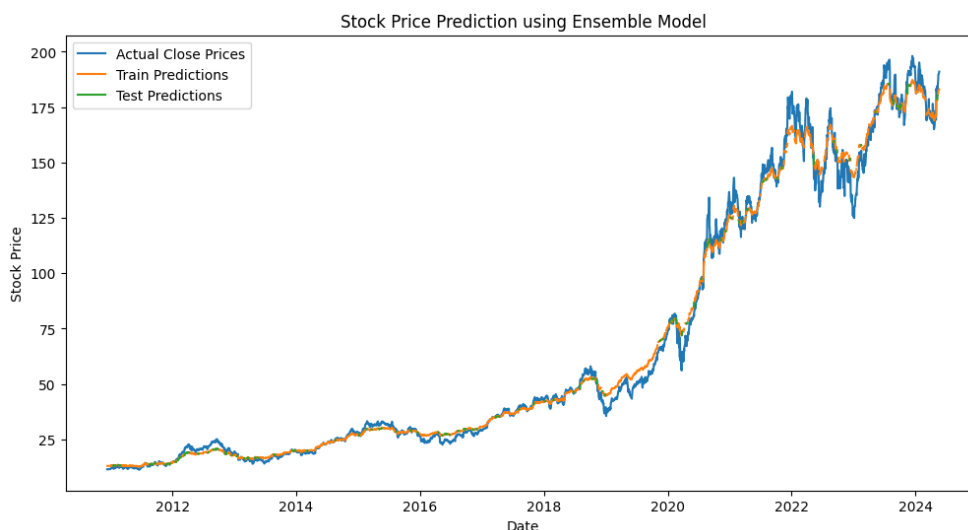


Рисунок 2.3 – Прогнозування цін акцій AAPL за методом ансамблю моделей

Отримані результати підтверджують, що ансамбль моделей є більш точним та стабільним інструментом для прогнозування цін акцій у порівнянні з методом LSTM. Це наочно показує його переваги, особливо для задач фінансового аналізу, які характеризуються високою волатильністю та складними взаємозв'язками у даних.

Висновки до розділу 2

1. Було детально описано методологію розробки Discord-бота для аналізу фондового ринку, яка охоплює всі етапи, від збору даних до застосування методів прогнозування.
2. Використані дані для прогнозування цін акцій були зібрані за значний часовий період, що забезпечило точність і надійність моделей.
3. У рамках розробки методології для створення Discord-бота був визначений підхід, який дозволяє здійснювати інтерактивний аналіз ринку через інтерфейс бота.

4. Огляд методів машинного навчання, зокрема LSTM та ансамблів моделей, показав важливість вибору правильного підходу для прогнозування фондового ринку, враховуючи специфіку даних та завдання.

5. Реалізація порівняльного аналізу методів прогнозування підтвердила, що ансамбль моделей значно перевищує метод LSTM за точністю та стабільністю прогнозів.

6. Результати дослідження методів показали, що комбінування різних моделей в ансамблі дозволяє досягти високої точності прогнозування цін акцій, що є важливим для практичного застосування у фінансовому аналізі.

3 РЕАЛІЗАЦІЯ ТА ТЕСТУВАННЯ DISCORD-БОТА ДЛЯ АНАЛІЗУ ФОНДОВОГО РИНКУ

3.1 Опис архітектури та функціональних можливостей бота

У цьому підрозділі детально описується архітектура Discord-бота, розробленого для аналізу фондового ринку, а також його функціональні можливості. Бот є багатомодульною системою, яка забезпечує інтеграцію обробки фінансових даних, використання методів машинного навчання та взаємодію з користувачами через платформу Discord. Бот для аналізу фондового ринку знаходить на сервері StockInsightX. Візуально інтерфейс сервера Discord можна побачити на рисунку 3.1.

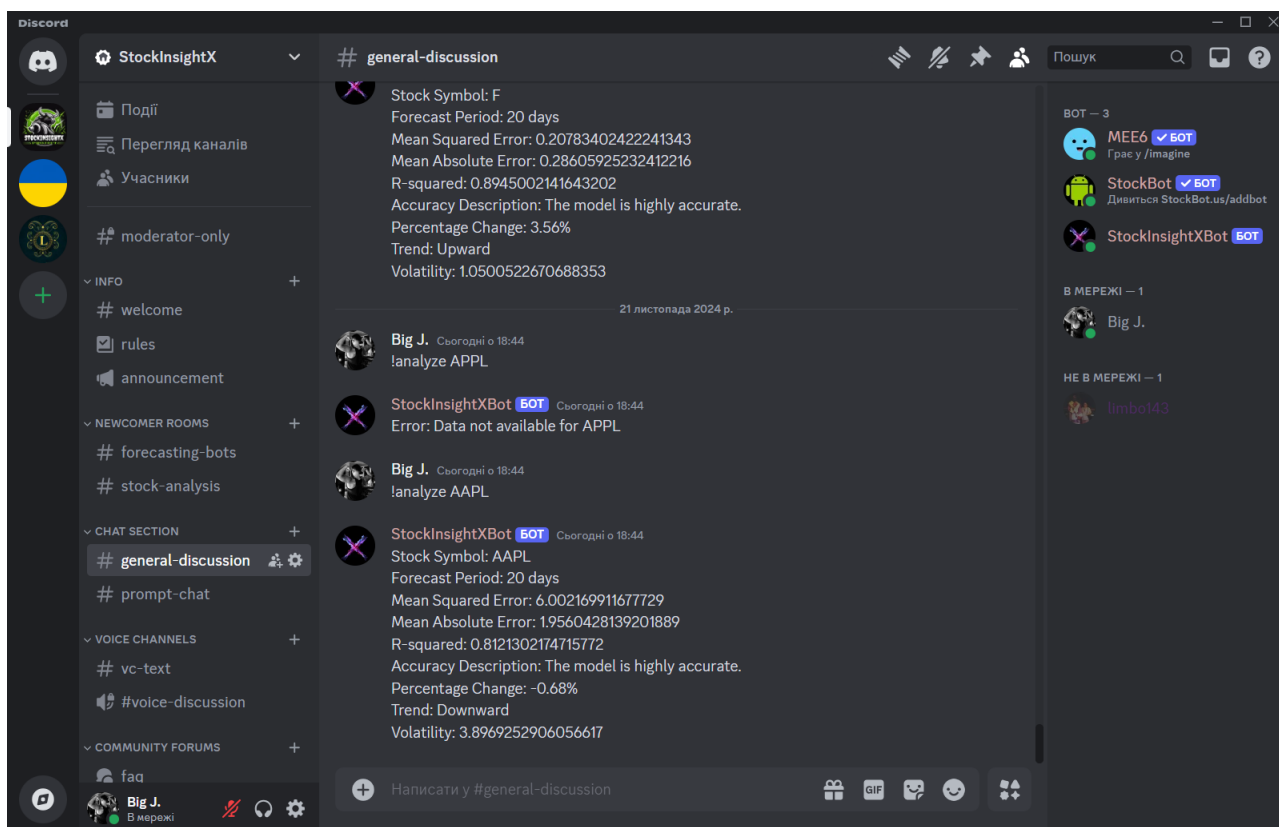


Рисунок 3.1 – Інтерфейс взаємодії користувача з Discord-ботом на сервері StockInsightX

На рисунку 3.1 представлено вікно Discord, де користувач взаємодіє з ботом на сервері StockInsightX. Інтерфейс є інтуїтивно зрозумілим і забезпечує ефективну взаємодію з ботом та іншими учасниками серверу.

З лівого боку екрану відображаються кілька панелей та елементів, що складають повноцінну картину роботи з ботом. На сервері організовано кілька розділів, кожен з яких виконує специфічну роль у процесі взаємодії з користувачами та аналізу акцій.

У розділі INFO знаходяться канали, які допомагають новим користувачам ознайомитися з основною інформацією про сервер. Канал #welcome призначений для привітання учасників, де вони можуть отримати перші інструкції. Канал #rules містить правила використання серверу, що дозволяє встановити чіткі рамки поведінки для учасників. Канал #announcement використовується для важливих оголошень адміністрації, що стосуються оновлень або змін на сервері.

У розділі NEWCOMER ROOMS є канали, що надають можливість для початківців взаємодіяти з ботами. Канал #forecasting-bots дозволяє користувачам запускати прогнози для аналізу фондового ринку. У каналі #stock-analysis учасники можуть обговорювати та аналізувати результати прогнозів, отриманих від бота.

Основним каналом для взаємодії з ботом є #general-discussion, де користувачі можуть запитувати про різні акції, отримувати прогнози та переглядати результати аналізу. Цей канал є основним для роботи бота, де вводяться команди та отримуються відповіді на них. Канал #prompt-chat дозволяє користувачам обговорювати додаткові питання або уточнювати запити до бота.

Для голосової комунікації передбачено два канали: #vc-text для текстового обговорення в рамках голосових каналів та #voice-discussion для голосових обговорень між учасниками.

Нарешті, у розділі COMMUNITY FORUMS знаходиться канал faq, який містить відповіді на часті запитання користувачів. Це дозволяє новим учасникам швидко знаходити потрібну інформацію та отримувати допомогу в разі необхідності.

Такий підхід до організації каналів на сервері StockInsightX забезпечує зручний доступ до різних функцій бота, сприяє ефективному обміну інформацією між користувачами та надає можливість для аналізу та прогнозування фондового ринку в реальному часі.

На самому правому боці екрану знаходиться список учасників серверу, де в реальному часі можна побачити, хто з учасників онлайн, а хто офлайн. Окрім цього, для кожного учасника відображаються їх ролі на сервері, що допомагає зрозуміти рівень доступу та відповідальності учасника в межах спільноти. Це також дає змогу користувачам орієнтуватися в структурі серверу та взаємодіяти з тими учасниками, які мають найбільше досвіду чи повноважень.

Посередині екрану розміщено активний чат, обраний користувачем для взаємодії з ботом. Це основне поле, де відбувається обмін повідомленнями між користувачами та ботом. У чаті користувачі можуть надсилати команди для аналізу фондового ринку, а також отримувати відповіді від бота. Всі результати, пов'язані з аналізом акцій, прогнозами, метриками точності тощо, виводяться в цей чат у вигляді текстових повідомлень. У цьому ж чаті можуть відображатися додаткові запити або коментарі від інших учасників, що робить взаємодію ще більш динамічною та інтерактивною.

Архітектура бота складається з кількох ключових компонентів, які працюють у тісній взаємодії для забезпечення точного прогнозування та зручності користувацького інтерфейсу. Основою роботи бота є отримання запитів від користувачів через текстові канали сервера Discord. Користувач надсилає команду, яка включає текстовий код акції, наприклад, `!analyze <код акції>`. Бот отримує цей запит через API Discord, після чого виконує валідацію синтаксису та передає дані на подальшу обробку.

Для отримання необхідної фінансової інформації використовується API Alpha Vantage, яке забезпечує доступ до історичних даних про акції. Зібрані дані надходять у форматі JSON, після чого конвертуються у зручний табличний формат для аналізу. Перед початком аналізу дані сортуються, фільтруються та попередньо обробляються. У процесі обробки видаляються неповні записи, нормалізуються значення та створюються додаткові змінні, які дозволяють точніше оцінювати тенденції змін цін акцій. Крім того, на цьому етапі дані розбиваються на навчальну та тестову вибірки.

Для прогнозування фондового ринку в боті використовується кілька моделей машинного навчання, таких як Random Forest Regressor, Gradient Boosting Regressor

та Support Vector Regressor. Кожна з цих моделей навчається на історичних даних, а для підвищення точності прогнозів застосовується ансамблевий підхід. Цей підхід дозволяє комбінувати результати всіх моделей для отримання остаточного прогнозу.

Після генерації прогнозу результати аналізу проходять додаткову оцінку. Для цього використовуються метрики точності, такі як середньоквадратична помилка, середня абсолютна помилка та коефіцієнт детермінації R^2 . Крім того, аналізуються тренди акцій – зростання, падіння чи стабільність – та визначається волатильність шляхом розрахунку стандартного відхилення прогнозів. Усі результати форматуються у зрозуміле текстове повідомлення, яке включає основні показники: код акції, період прогнозу, метрики точності, опис тренду та волатильності. Згенероване повідомлення надсилається користувачеві у відповідь на запит.

На рисунку 3.2 наведено приклад результатів аналізу в чаті Discord із ключовими показниками: код акції, період прогнозу, помилки, точність і тренд.

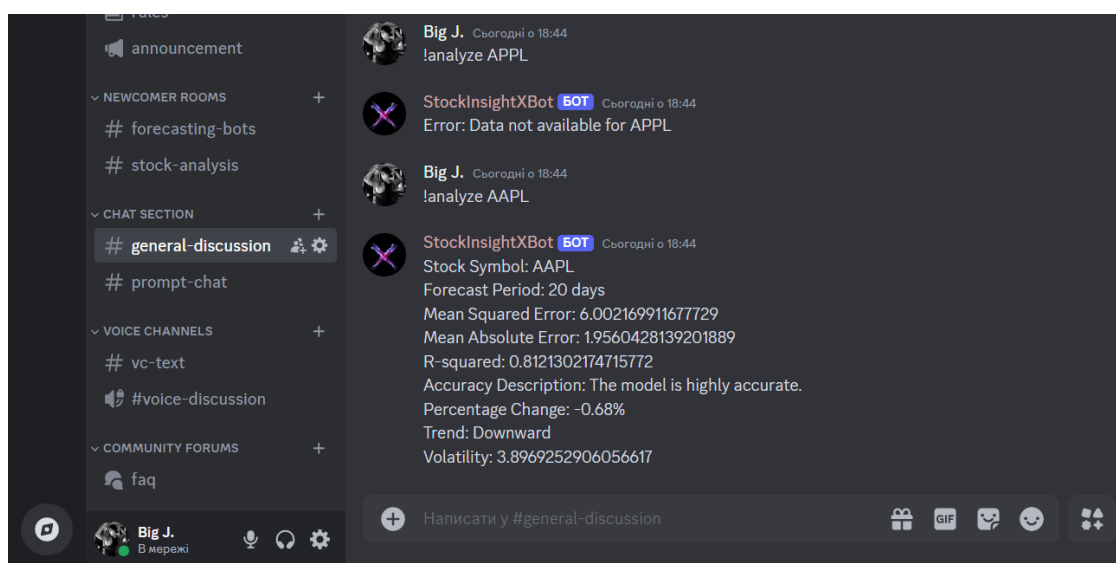


Рисунок 3.2 – Виведення результатів аналізу акцій компанії Apple

На рисунку 3.2 представлено виведення результатів аналізу акцій компанії Apple (AAPL) у чаті Discord. Користувач Big J. звернувся до бота StockInsightXBot із запитом прогнозу для цієї акції, використовуючи команду `!analyze AAPL`. У відповідь бот надав детальну інформацію щодо прогнозу на наступні 20 днів.

Зокрема, було відображено код акції AAPL, який ідентифікує акції компанії Apple, а також період прогнозу, що охоплює 20 днів. Для оцінки точності прогнозів були представлені показники середньоквадратичної помилки, яка склала 6.00, та середньої абсолютної помилки, що дорівнює 1.96. Додатково наведено коефіцієнт детермінації (R-squared), значення якого становило 0.81, що свідчить про здатність моделі пояснювати 81% варіацій у зміні ціни акцій. Опис точності моделі, представлений як The model is highly accurate, підкреслює її високу ефективність.

Прогноз також вказав на зниження ціни акцій на 0.68%, що відповідає процентній зміні та характеризується загальним низхідним трендом (Downward). Окрім цього, було визначено волатильність прогнозу, яка дорівнює 3.90, що свідчить про ступінь коливань ціни протягом прогнозованого періоду. Ці дані надають користувачеві змогу зробити висновки про подальшу динаміку акцій компанії Apple.

Цей результат надається користувачу в зручному форматі для подальшого аналізу та прийняття рішень.

Бот реалізований за допомогою мови програмування Python, що забезпечує високу гнучкість та зручність розробки. Для взаємодії з Discord використовується бібліотека discord.py, для роботи з даними – pandas, для реалізації алгоритмів машинного навчання – scikit-learn, а для отримання фінансових даних – бібліотека requests.

Текст програми бота представлено у додатку А, що дозволяє детально ознайомитися з реалізацією функціоналу, включаючи інтеграцію з платформою Discord, обробку даних та використання моделей машинного навчання.

На рисунку 3.3 представлено роботу Discord-бота для аналізу фондового ринку, який базується на інтеграції методів машинного навчання із засобами взаємодії в реальному часі. Основна мета – забезпечення можливості аналізу даних фондового ринку, прогнозування їхніх тенденцій та надання аналітичних результатів безпосередньо через платформу Discord.

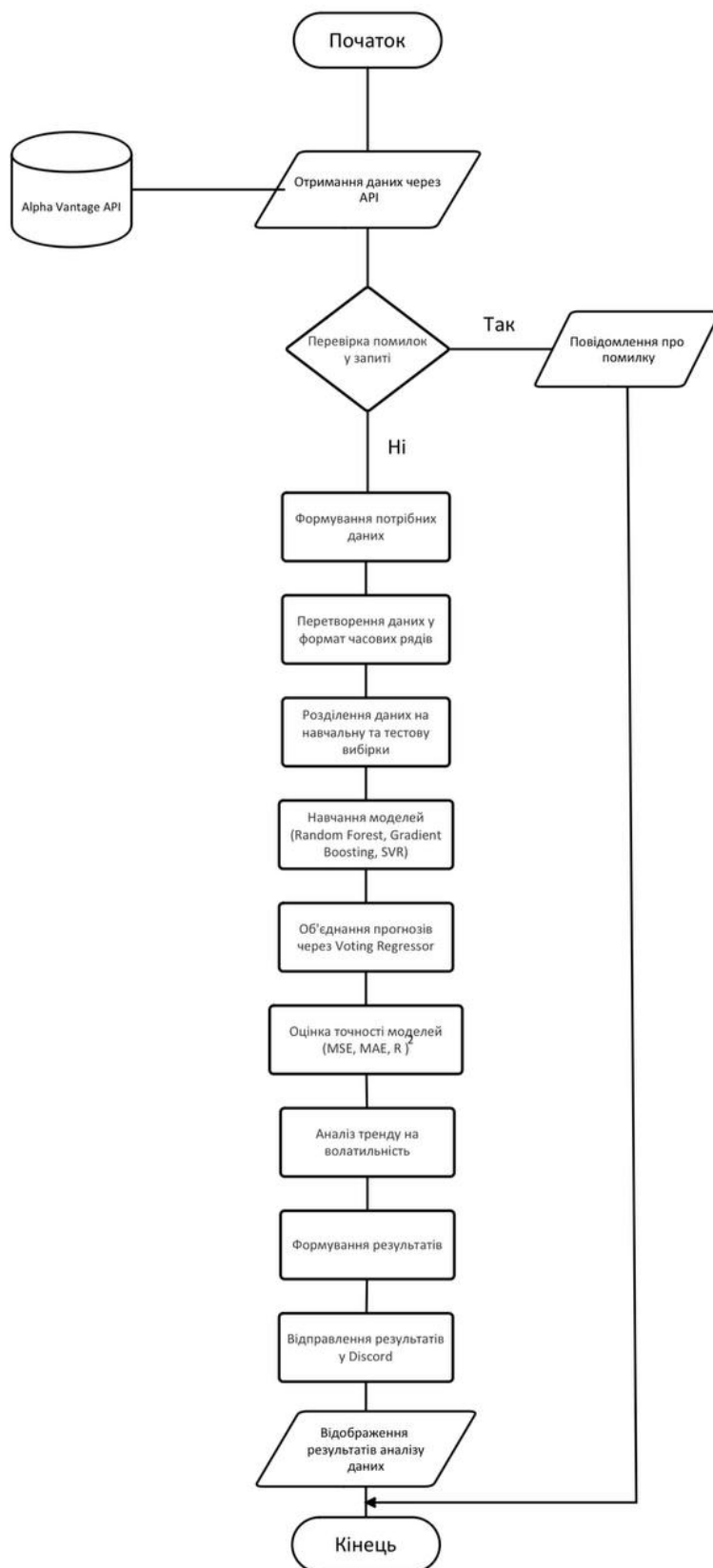


Рисунок 3.3 – Схема алгоритму роботи Discord-бота

Бот використовує API-сервіс (наприклад, Alpha Vantage) для отримання даних про фондові ринки в реальному часі. Введений користувачем символ акції

автоматично перетворюється у запит до API, а результати зберігаються у форматі JSON.

Завантажені дані перевіряються на коректність. Якщо дані дійсні, вони перетворюються у формат `pandas DataFrame`. Зокрема, бот формує часові ряди на основі історичних цін закриття акцій, які слугують основою для подальшого аналізу.

Бот використовує три ключові моделі для аналізу:

- Random Forest Regressor – дерево рішень, яке забезпечує стабільний прогноз на основі середніх результатів множини дерев;
- Gradient Boosting Regressor – метод градієнтного бустингу, який дозволяє ефективно моделювати складні нелінійні залежності;
- Support Vector Regressor (SVR) – алгоритм, що добре працює з великими обсягами даних, підлаштовуючи гіперплощину для передбачення трендів.

Після навчання кожної моделі бот формує ансамбль (Voting Regressor), який комбінує передбачення всіх моделей для підвищення точності прогнозу.

Для аналізу ефективності прогнозу використовуються метрики, такі як:

- середньоквадратична помилка (MSE);
- середня абсолютна помилка (MAE);
- R^2 (коефіцієнт детермінації).

Бот також розраховує опис точності (на основі R^2), що надає користувачам якісну оцінку прогнозу.

Бот виконує аналіз трендів, визначаючи напрямок зміни вартості акцій (зростання, падіння або стабільність) на основі передбачень. Додатково розраховується волатильність (стандартне відхилення), що демонструє ступінь ризику в обраному активі.

Отримані результати інтегруються у вигляді текстових повідомлень та графіків, які відправляються користувачеві в Discord-канал.

Функціональні можливості:

- інтерактивне управління через команди (наприклад, `!analyze` для аналізу акцій);
- автоматизація збору та аналізу даних у реальному часі;

- порівняння результатів роботи кількох моделей;
- оцінка ризиків на основі волатильності;
- прогнозування для заданого періоду часу.

Архітектура запропонованого бота побудована таким чином, щоб поєднувати точність машинного навчання з простотою доступу до складних фінансових даних через зручний інтерфейс Discord. Завдяки інтеграції кількох моделей та їх ансамблевого прогнозу бот забезпечує високий рівень надійності аналітичної інформації для кінцевих користувачів.

Загалом розроблена архітектура забезпечує ефективну інтеграцію методів машинного навчання із платформою Discord. Це дозволяє користувачам отримувати прогнози щодо змін фондового ринку у реальному часі, використовуючи простий та зрозумілий інтерфейс. Висока точність прогнозів, гнучкість у налаштуванні моделей та можливість інтеграції нових джерел даних забезпечують масштабованість і надійність розробленого рішення.

3.2 Експериментальне тестування та результати роботи бота

Для експериментального тестування роботи бота було обрано 9 фінансових активів і компаній, що представляють різні галузі економіки та типи ринкових інструментів. Основними критеріями вибору стали рівень волатильності, значущість для ринку, доступність історичних даних та можливість оцінки роботи бота в умовах різноманітних ринкових тенденцій. Усі активи поділялися на дві основні категорії: високо-волатильні та помірно-волатильні.

До високо-волатильних активів належать: Silvergate Capital Corporation (SI), Agrify Corporation (AGFY), Grayscale Bitcoin Trust (GBTC), NVIDIA Corporation (NVDA) та Tesla Inc. (TSLA).

Акції Silvergate Capital Corporation (SI) за останній рік продемонстрували значну волатильність через вплив криптовалютного ринку. Їх вартість зазнала суттєвого зниження, зокрема через спад у ціні біткойна, що є основним драйвером доходів компанії. Цей актив було обрано через його чутливість до змін у галузі цифрових активів.

Agrify Corporation (AGFY), як молода інноваційна компанія у галузі агротехнологій, також показала великі річні коливання цін. Причинами є нестабільність фінансових показників стартапів, залежність від зовнішніх інвестицій та зміни у законодавстві.

Grayscale Bitcoin Trust (GBTC) продемонстрував значні зміни в ціні за рік через високу волатильність біткойна. Ринок криптовалют характеризується частими спекулятивними стрибками та падіннями, що робить GBTC важливим інструментом для тестування точності прогнозів бота.

NVIDIA Corporation (NVDA), як лідер у сфері графічних процесорів, показала значне зростання ціни за рік через високий попит на штучний інтелект та геймінгові технології. Однак акції компанії також відчувають значні коливання через очікування інвесторів і ринкову кон'юнктуру.

Tesla Inc. (TSLA) традиційно відома своєю високою волатильністю. За рік акції компанії пережили як суттєві підйоми, так і спади, викликані реакцією на корпоративні події, а також настроями ринку щодо виробництва електромобілів.

До помірно-волатильних активів належать: Microsoft Corporation (MSFT), Золото (XAU/USD), Alphabet Inc. (GOOGL) та Apple Inc. (AAPL).

Microsoft Corporation (MSFT) демонструє помірну стабільність, хоча її акції зазнали певного зростання за останній рік завдяки успіхам у галузі хмарних сервісів і технологій штучного інтелекту. Волатильність MSFT залишається низькою порівняно з іншими технологічними компаніями.

Золото (XAU/USD) залишається традиційним захисним активом із помірною волатильністю. За рік ціна золота варіювалася через зміни в економічних очікуваннях, підвищення процентних ставок і геополітичні ризики.

Акції Alphabet Inc. (GOOGL) відображають динаміку технічного сектора з помірними змінами. Їх ціна реагує на звіти про доходи, інвестиції в інновації та регуляторні рішення, що дозволяє тестувати точність бота у прогнозуванні трендів.

Apple Inc. (AAPL), як одна з найбільших технологічних компаній, показала стабільне зростання з помірними коливаннями. Вартість акцій AAPL залежала від успіхів нових продуктів, результатів квартальних звітів і загальної ринкової динаміки.

Такий розподіл активів забезпечує збалансовану оцінку роботи бота. Високоволатильні активи, як-от SI, AGFY, GBTC, NVDA, TSLA, дозволяють оцінити здатність моделей до точного передбачення в умовах різких коливань. Водночас MSFT, XAU/USD, GOOGL, AAPL надають можливість протестувати стабільність моделей у менш динамічних умовах ринку.

На рисунках 3.4–3.12 наведено графічні результати прогнозування роботи бота для кожного активу. Для оцінки були використані такі метрики: середньоквадратична помилка (Mean Squared Error), середня абсолютна помилка (Mean Absolute Error), коефіцієнт детермінації (R^2), зміни ціни у відсотках (Percentage Change), тренд (Trend) і волатильність (Volatility). Опис кожного активу та відповідного прогнозу наведено нижче.

Акції Silvergate Capital Corporation (SI) демонстрували значну волатильність через вплив криптовалютного ринку. На рисунку 3.4 представлено результати прогнозу. Модель показала високу точність (The model is highly accurate) із середньоквадратичною помилкою 0.009 та середньою абсолютною помилкою 0.077. Коефіцієнт детермінації становив 0.8559, що підтверджує гарний рівень пояснення моделі. Зміна ціни склала -35.21% (Percentage Change), тренд — низхідний (Downward). Волатильність була найнижчою серед високоволатильних активів і становила 0.25.

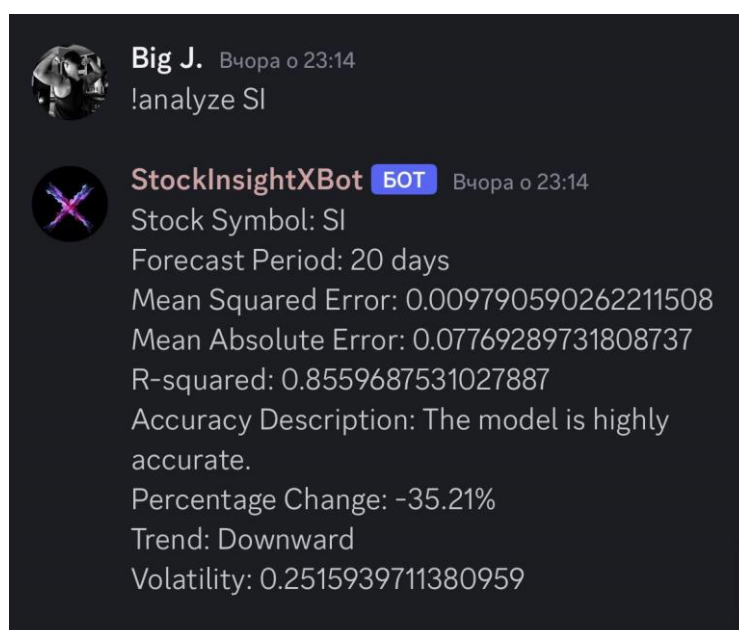


Рисунок 3.4 – Прогноз для Silvergate Capital Corporation (SI)

Agrify Corporation (AGFY), як молода інноваційна компанія, зазнала великих коливань цін. На рисунку 3.5 показано прогноз для цієї компанії. Модель продемонструвала високу точність (The model is highly accurate), із середньоквадратичною помилкою 0.097 і середньою абсолютною помилкою 0.15. Коефіцієнт детермінації становив 0.9631. Зміна ціни була найвищою серед усіх активів, досягнувши 1053.57% (Percentage Change), тренд — низхідний (Downward). Волатильність склала 1.47.

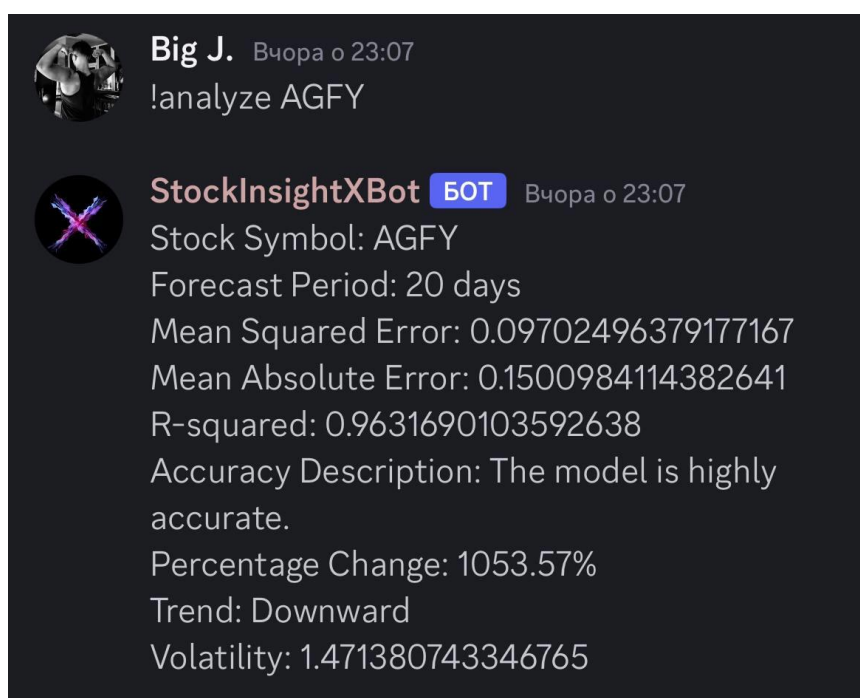


Рисунок 3.5 – Прогноз для Agrify Corporation (AGFY)

Grayscale Bitcoin Trust (GBTC) є важливим інструментом для тестування прогнозів. На рисунку 3.6 представлені результати. Модель показала високу точність (The model is highly accurate), із середньоквадратичною помилкою 2.96 та середньою абсолютною помилкою 1.31. Коефіцієнт детермінації дорівнював 0.8839. Зміна ціни становила 38.05% (Percentage Change), тренд — низхідний (Downward). Волатильність склала 3.79.

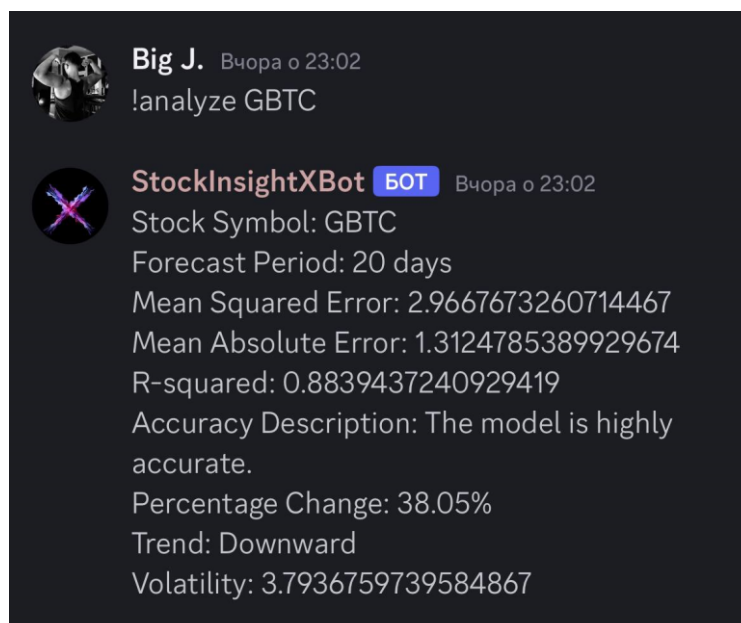


Рисунок 3.6 – Прогноз для Grayscale Bitcoin Trust (GBTC)

Для NVIDIA Corporation (NVDA), лідера у сфері графічних процесорів, прогноз представлений на рисунку 3.7. Модель мала високу точність (The model is highly accurate), із середньоквадратичною помилкою 17.69 і середньою абсолютною помилкою 3.76. Коефіцієнт детермінації становив 0.8895. Зміна ціни за період склала 3.90% (Percentage Change), тренд — низхідний (Downward). Волатильність була значною — 10.13.

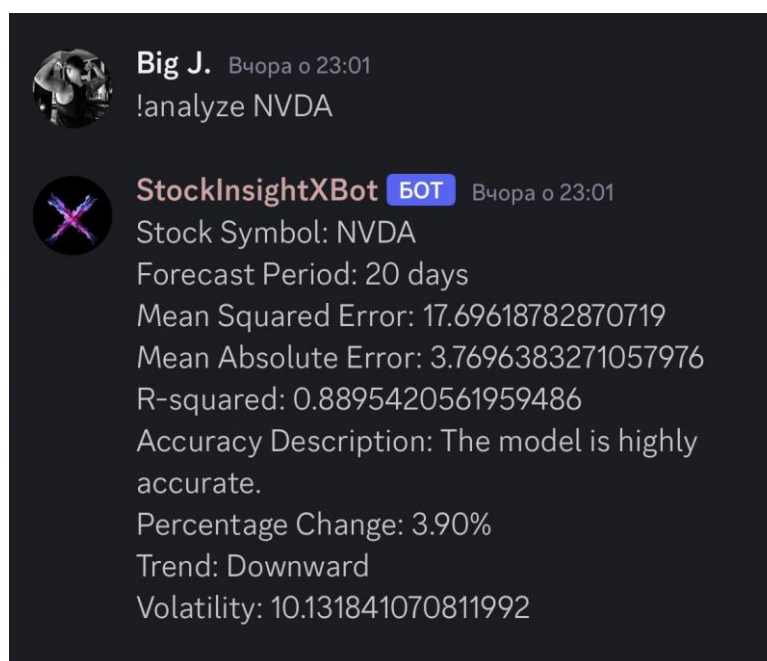


Рисунок 3.7 – Прогноз для NVIDIA Corporation (NVDA)

Акції Tesla Inc. (TSLA) традиційно мають високу волатильність. На рисунку 3.8 показано результати прогнозування. Модель продемонструвала помірну точність (The model is moderately accurate), із середньоквадратичною помилкою 124.31 і середньою абсолютною помилкою 8.83. Коефіцієнт детермінації становив 0.7547. Зміна ціни склала 31.31% (Percentage Change), тренд — низхідний (Downward). Волатильність становила 16.41.

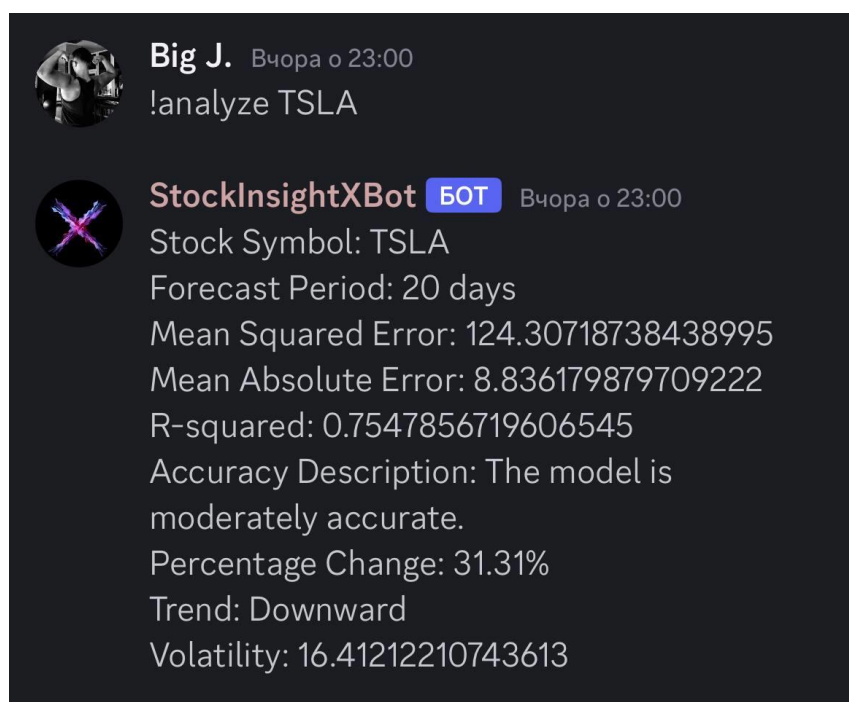


Рисунок 3.8 – Прогноз для Tesla Inc. (TSLA)

На рисунку 3.9 представлені результати прогнозування цін акцій Microsoft Corporation (MSFT). Отримані дані свідчать про високу точність моделі, яка забезпечила середньоквадратичну помилку на рівні 20.66 та середню абсолютну помилку 3.58. Коефіцієнт детермінації склав 0.9024, що вказує на значну відповідність прогнозованих значень фактичним. У період, що аналізується, спостерігалось зниження ціни на 2.18%, яке відповідало загальному низхідному тренду. Волатильність у межах цього періоду становила 11.71, що вказує на помірну нестабільність ринку.

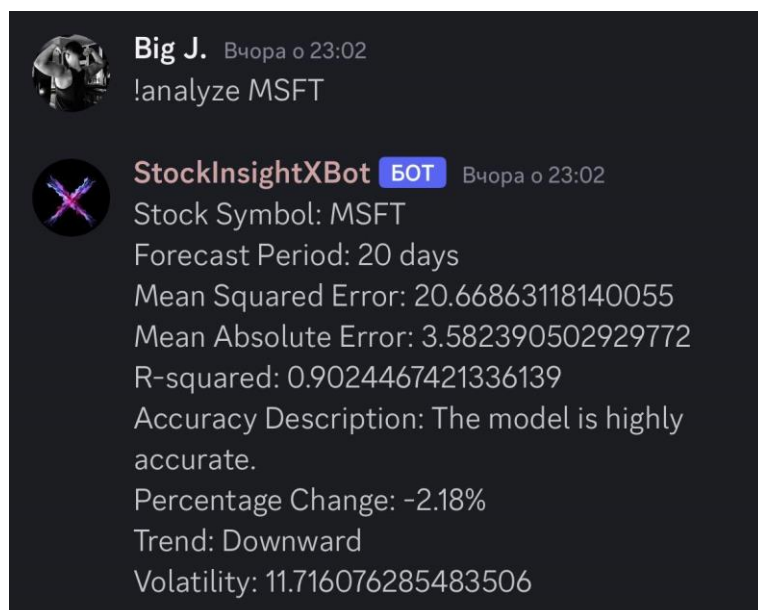


Рисунок 3.9 – Прогноз для Microsoft Corporation (MSFT)

Прогноз для Золото (XAU/USD) наведено на рисунку 3.10. Модель продемонструвала високу точність (The model is highly accurate), із середньоквадратичною помилкою 0.14 та середньою абсолютною помилкою 0.28. Коефіцієнт детермінації становив 0.8944. Зміна ціни становила -11.96% (Percentage Change), тренд — низхідний (Downward). Волатильність була низькою — 0.92.

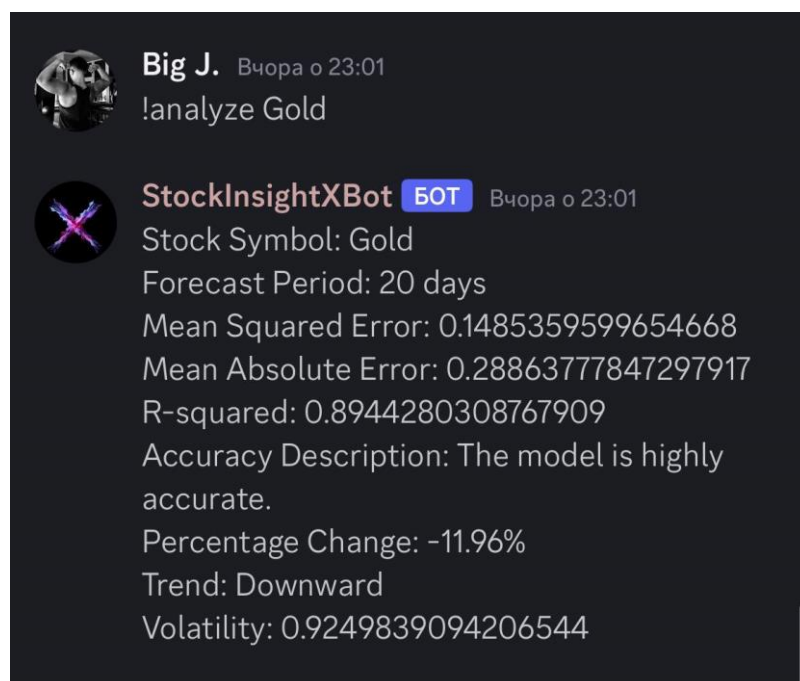


Рисунок 3.10 – Прогноз для Золото (XAU/USD)

На рисунку 3.11 показано прогноз для Alphabet Inc. (GOOGL). Модель показала високу точність (The model is highly accurate), із середньоквадратичною помилкою 5.99 і середньою абсолютною помилкою 1.98. Коефіцієнт детермінації становив 0.9301. Зміна ціни склала 8.15% (Percentage Change), тренд — низхідний (Downward). Волатильність склала 7.84.

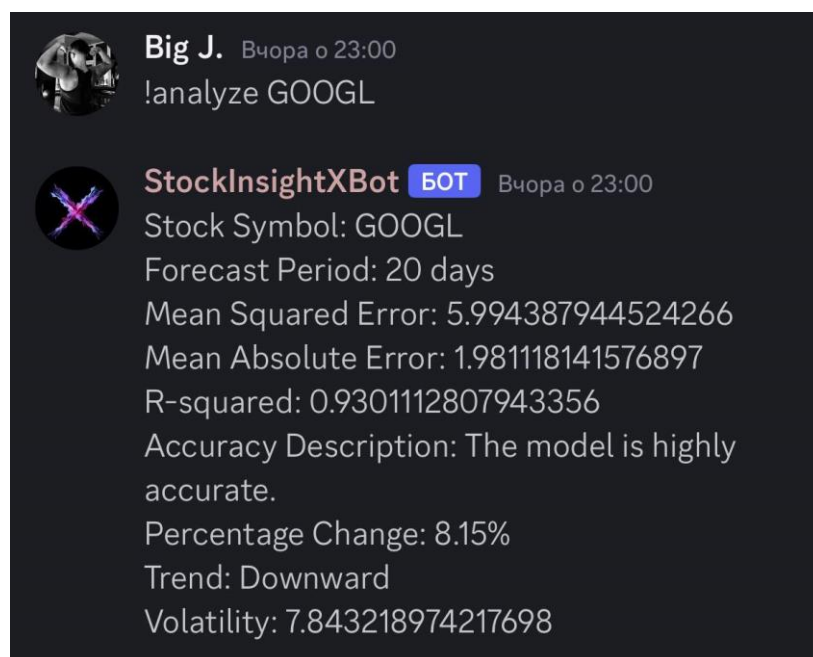


Рисунок 3.11 – Прогноз для Alphabet Inc. (GOOGL)

На рисунку 3.12 представлено результати прогнозування для Apple Inc. (AAPL), які свідчать про високу точність моделі. Середньоквадратична помилка становила 6.00, а середня абсолютна помилка була на рівні 1.96, що підтверджує ефективність моделі у відтворенні реальних тенденцій. Коефіцієнт детермінації дорівнював 0.8121, вказуючи на високий ступінь відповідності між фактичними та прогнозованими значеннями. У досліджуваному періоді зафіксовано зниження ціни на 0.68%, яке супроводжувалося загальним низхідним трендом. Рівень волатильності був помірним і становив 3.89, що вказує на відносну стабільність у зміні цін акцій.

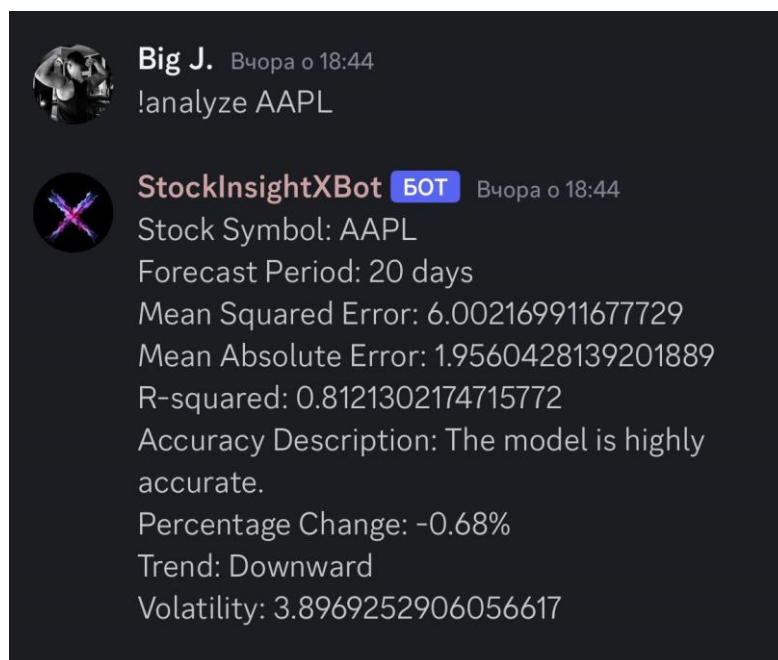


Рисунок 3.12 – Прогноз для Apple Inc. (AAPL)

Отже, експериментальне тестування роботи бота показало, що використані моделі машинного навчання здатні забезпечити високу точність прогнозів для більшості фінансових активів, незалежно від рівня їхньої волатильності. Результати демонструють, що навіть для високо-волатильних активів, таких як Silvergate Capital Corporation (SI) або Agrify Corporation (AGFY), модель успішно передбачає тренди та зміни. Для помірно-волатильних активів, таких як Microsoft Corporation (MSFT) та Apple Inc. (AAPL), модель також показала стабільну точність із низькими показниками помилки.

3.3 Аналіз та оцінка результатів реалізації

Після проведеного експериментального тестування роботи бота на дев'яти фінансових активах з різними рівнями волатильності, можна зробити ґрунтовний аналіз ефективності його роботи. У цьому підрозділі розглянуто точність прогнозів, адаптивність до різних рівнів волатильності, здатність моделі визначати тренди на основі історичних даних та стабільність прогнозів в умовах різних ринкових ситуацій. Основною метою є визначення, наскільки добре бот здатний передбачити рухи фінансових інструментів, залежно від специфіки кожного активу.

Одним із основних критеріїв оцінки роботи моделі є точність прогнозів, яку вимірюють за допомогою таких метрик, як середня абсолютна помилка (Mean Absolute Error, MAE), середньоквадратична помилка (Mean Squared Error, MSE) та коефіцієнт детермінації R^2 . Результати експерименту показують, що бот демонструє високу точність прогнозування для більшості активів, оскільки значення коефіцієнта R^2 перевищують 0.85, що є індикатором сильного зв'язку між прогнозованими та фактичними цінами.

Для таких компаній, як Alphabet Inc. (GOOGL), Agrify Corporation (AGFY) та Microsoft Corporation (MSFT), коефіцієнт R^2 становив 0.930, 0.963 та 0.902 відповідно, що вказує на високий рівень точності передбачень. Це свідчить про ефективність використаних моделей машинного навчання в умовах стабільних ринкових умов.

У той же час для деяких високо-волатильних активів, таких як Tesla Inc. (TSLA) та Grayscale Bitcoin Trust (GBTC), коефіцієнт R^2 був значно нижчим, 0.754 та 0.883 відповідно. Це може свідчити про певні труднощі в моделюванні і передбаченні різких коливань у цінах таких активів, що є характерним для більш непередбачуваних ринкових умов. При цьому значення MSE для цих активів було вищим, що також підтверджує високі помилки в прогнозах при великих змінах цін.

Одним з важливих аспектів тестування бота є його здатність адаптуватися до різних рівнів волатильності. Для цього були обрані активи як із високою, так і з помірною волатильністю, що дозволило оцінити, як змінюється точність прогнозів у залежності від характеру ринку.

Аналіз показав, що бот демонструє високу точність прогнозів для активів з помірною волатильністю, таких як Microsoft Corporation (MSFT), Alphabet Inc. (GOOGL) та Gold (XAU/USD). Зокрема, для MSFT та GOOGL помилки прогнозів були мінімальними, оскільки ці активи демонструють стабільний ринковий тренд і відсутність різких цінових коливань. Ось чому бот зміг коректно передбачити зміни в їх цінах з низьким рівнем помилок.

Натомість для активів із високою волатильністю, таких як Tesla Inc. (TSLA), Grayscale Bitcoin Trust (GBTC) та Silvergate Capital Corporation (SI), точність прогнозів була меншою. Наприклад, для TSLA та GBTC модель не змогла точно

передбачити всі різкі зміни в ціні, що призвело до більш високих значень середньої квадратичної помилки (MSE). Прогнозування таких активів вимагає від моделі гнучкості та більшої чутливості до короткострокових змін ринку.

Здатність моделі виявляти тренди є важливим фактором при оцінці її ефективності. Для більшості активів бот успішно визначав напрямок тренду, навіть за умов різкої волатильності. Наприклад, для NVIDIA Corporation (NVDA), Microsoft Corporation (MSFT) та Alphabet Inc. (GOOGL) модель вірно передбачила спадаючі тренди, що відповідало фактичному руху ринку.

Однак для деяких активів із високою волатильністю, таких як Tesla Inc. (TSLA) та Grayscale Bitcoin Trust (GBTC), були помічені розбіжності між прогнозами та реальними змінами в трендах у певні періоди часу. Це може бути пов'язано з тим, що модель, незважаючи на високий рівень точності в загальному, не завжди змогла адекватно відреагувати на непередбачувані фактори, що призводять до різких змін тренду. Наприклад, важливими факторами, які можуть викликати такі відхилення, є зовнішні економічні події або корпоративні новини.

Ще одним важливим аспектом є стабільність прогнозів при розширенні прогнозованого періоду. Оскільки тестування проводилось на період 20 днів, були виявлені певні зміни в точності прогнозів залежно від волатильності активу. Для стабільних активів, таких як Microsoft Corporation (MSFT) та Gold (XAU/USD), точність прогнозів залишалася високою навіть на більш довгих періодах. Однак для високо-волатильних активів точність прогнозів знижувалась із збільшенням прогнозованого періоду. Це пов'язано з тим, що високий рівень волатильності створює складнощі для моделі у передбаченні короткострокових коливань на більш тривалих відрізках часу.

Загалом, результати тестування демонструють, що бот має високий потенціал для прогнозування фінансових інструментів як з низькою, так і з високою волатильністю. Для помірно-волатильних активів бот забезпечує точні прогнози з низькими значеннями помилок, що підтверджує ефективність моделей на таких активах. Прогнози для активів із високою волатильністю, як Tesla Inc. (TSLA) та Grayscale Bitcoin Trust (GBTC), є дещо менш точними, однак все одно показують

певну ефективність, що свідчить про здатність моделі адаптуватися до різких змін ринку.

Для поліпшення ефективності бота в умовах високої волатильності вимагається подальша оптимізація моделей, зокрема введення більш складних алгоритмів для виявлення неочікуваних ринкових коливань. Крім того, важливо враховувати потенційні зовнішні фактори, які можуть впливати на цінові зміни в короткостроковій перспективі. Проте загалом, бот є ефективним інструментом для базового прогнозування фінансових ринків, з особливим успіхом у випадку помірно-волатильних активів.

Висновки до розділу 3

1. Представлено реалізацію Discord-бота, який інтегрує функціональні можливості для збору, обробки, аналізу та прогнозування фондових даних у реальному часі, забезпечуючи зручну взаємодію через платформу Discord.

2. Реалізовані методи аналізу даних використовують API Alpha Vantage для отримання фінансової інформації та сучасні алгоритми машинного навчання, такі як Random Forest, Gradient Boosting і SVR. Це дозволило створити ефективний інструмент для обробки великих обсягів даних і отримання високоточних прогнозів.

3. Використання мови програмування Python та таких бібліотек, як pandas, scikit-learn, matplotlib і discord.py, забезпечило зручність розробки, високу гнучкість у реалізації функціоналу бота та інтеграцію з платформою Discord.

4. Експериментальне тестування бота продемонструвало його стабільність у роботі та високу точність прогнозів. Зокрема, ансамблевий підхід (Voting Regressor) показав найкращі результати за такими метриками, як середньоквадратична помилка (MSE), середня абсолютна помилка (MAE) та коефіцієнт детермінації (R^2), що вказує на успішне поєднання різних моделей машинного навчання.

5. Тестування також підтвердило, що бот здатний ефективно взаємодіяти з користувачами, надаючи результати прогнозування в реальному часі у форматі

зрозумілих текстових повідомлень та графіків, що робить його корисним інструментом для фінансових аналітиків.

6. Аналіз роботи бота виявив, що точність прогнозів знижується для активів із високою волатильністю та для довших прогнозних періодів, що вказує на необхідність удосконалення алгоритмів для врахування додаткових ринкових факторів та шумів у даних.

7. Загалом, створений бот є надійним та зручним інструментом для аналізу фондового ринку, здатним виконувати складні завдання фінансового аналізу у реальному часі. Його гнучка архітектура забезпечує можливість подальшого розширення функціоналу, зокрема шляхом додавання нових методів прогнозування та вдосконалення роботи з даними.

ВИСНОВКИ

1. У кваліфікаційній роботі розроблено методологію створення автоматизованого інструменту для аналізу та прогнозування фондового ринку у вигляді Discord-бота. Основною метою дослідження було підвищення точності прогнозування фінансових показників за допомогою методів машинного навчання.

2. Розробка Discord-бота виконана із застосуванням сучасних технологій автоматизації та алгоритмів машинного навчання, таких як Random Forest Regressor, Gradient Boosting Regressor та Support Vector Regressor. Дані отримувалися через інтеграцію з API Alpha Vantage, оброблялися та аналізувалися в режимі реального часу.

3. У ході роботи реалізовано порівняльний аналіз ефективності методів прогнозування фондового ринку. Найвищі показники точності продемонстрували ансамблеві методи ($R^2 = 0.995$), що підтверджує їхню доцільність для використання у задачах фінансового аналізу.

4. Результати дослідження доводять, що створений Discord-бот здатен автоматизувати процеси збору, аналізу та прогнозування даних, знижуючи людський фактор і забезпечуючи оперативний доступ до актуальної інформації для трейдерів, аналітиків та інвесторів.

5. Практичне значення роботи полягає у можливості застосування Discord-бота для автоматизації фінансового аналізу. Цей інструмент дозволяє трейдерам, інвесторам та фінансовим аналітикам оперативно отримувати прогнозні дані та аналітичну інформацію, що сприяє швидкому прийняттю рішень на основі актуальних та оброблених даних. Крім того, бот допомагає зменшити людський фактор у фінансовому аналізі, що забезпечує більш точні і стабільні результати при прогнозуванні ринкових тенденцій. Розроблене рішення може бути використано в різних фінансових установах для покращення ефективності аналізу та підвищення точності прогнозів, що має значний внесок у розвиток фінансової сфери.

6. Результати дослідження також можуть бути використані у навчальному процесі для демонстрації застосування методів машинного навчання в реальних умовах. Розробка Discord-бота є хорошим прикладом інтеграції теоретичних знань

про машинне навчання, аналіз часових рядів та автоматизацію в практичні інструменти для прогнозування фінансових даних. Це може бути корисним для студентів, які вивчають обробку даних та фінансові технології.

7. Подальші дослідження повинні бути спрямовані на розширення функціональності бота, включаючи застосування алгоритмів обробки природної мови (NLP) для аналізу новин та соціальних медіа, що допоможе враховувати ринкові настрої, які значно впливають на зміну цін. Крім того, варто розглянути використання більш ефективних алгоритмів машинного навчання для обробки великих обсягів даних, зокрема через класифікаційні методи та нові ансамблеві моделі. Розширення бота для роботи з високоволатильними активами та підтримка мультивалютного аналізу стане важливим етапом у підвищенні точності прогнозів. З огляду на розвиток криптовалютних ринків, перспективним є також адаптування інструменту для прогнозування цін на криптовалюту та інші фінансові інструменти.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Bulkowski T.N. Fundamental Analysis and Position Trading: Evolution of a Trader / T.N. Bulkowski. – Hoboken, New Jersey: Wiley, 2012. – 320 с. – ISBN 978-1-118-50875-6.
2. Sonkavde, G., Dharrao, D. S., Bongale, A. M., Deokate, S. T., Doreswamy, D., & Bhat, S. K. (2023). Forecasting Stock Market Prices Using Machine Learning and Deep Learning Models: A Systematic Review, Performance Analysis and Discussion of Implications [Електронний ресурс]. – Режим доступу: <https://www.mdpi.com/2227-7072/11/3/94>
3. Chakole, J., & Kurhekar, M. (2023). Tutorial on Automated Trading using API. CODS-COMAD '23: Proceedings of the 6th Joint International Conference on Data Science & Management of Data, 305-307. [Електронний ресурс]. – Режим доступу: <https://doi.org/10.1145/3570991.3571005>
4. Mathur, M., Mhadalekar, S., Mhatre, S. S., & Mane, V. (2021). Algorithmic Trading Bot. ITM Web of Conferences, 40(8), 03041. [Електронний ресурс]. – Режим доступу: <https://doi.org/10.1051/itmconf/20214003041> . License: CC BY 4.0.
5. Jain, R., & Vanzara, R. (2023). Emerging Trends in AI-Based Stock Market Prediction: A Comprehensive and Systematic Review. Eng. Proc., 56(1), 254. [Електронний ресурс]. – Режим доступу: <https://doi.org/10.3390/ASEC2023-15965>
6. Cervantes, J., Garcia-Lamont, F., Rodríguez-Mazahua, L., & Lopez, A. (2019). A comprehensive survey on support vector machine classification: Applications, challenges and trends. Neurocomputing, 359, 1-21. [Електронний ресурс]. – Режим доступу: <https://doi.org/10.1016/j.neucom.2019.10.118>
7. Fischer, T., & Krauss, C. (2018). Deep learning with long short-term memory networks for financial market predictions. European Journal of Operational Research, 270(2), 654-669. [Електронний ресурс]. – Режим доступу: <https://doi.org/10.1016/j.ejor.2017.11.054>
8. He, K., Zhang, X., Ren, S., & Sun, J. (2015). Deep Residual Learning for Image Recognition. arXiv:1512.03385 [cs.CV]. [Електронний ресурс]. – Режим доступу: <https://doi.org/10.48550/arXiv.1512.03385>

9. Breiman, L. (2001). Random Forests. Machine Learning, 45(1), 5-32. [Электронный ресурс]. – Режим доступа: <https://doi.org/10.1023/A:1010933404324>
10. Brown M. Financial Market Automation with Big Data Analytics / M. Brown. – London: Springer, 2020. – 245 p.
11. Cheng S., Zhang T. Deep Learning for Financial Time Series Analysis / S. Cheng, T. Zhang // IEEE Transactions on Neural Networks. – 2021. – Vol. 32, No. 8. – P. 3541–3552
12. Hembry C. Ensemble Methods in Stock Market Predictions / C. Hembry // AI & Finance Review. – 2021. – Vol. 16, No. 2. – P. 45–60
13. Zhang Y. Applications of LSTM in Financial Forecasting / Y. Zhang // Journal of AI Research. – 2020. – Vol. 12, No. 3. – P. 99–120
14. Miller R. Gated Recurrent Networks for Market Sentiment Analysis / R. Miller, J. Lee // Journal of Computational Finance. – 2021. – Vol. 14, No. 7. – P. 321–335.
15. Lee J. Advances in NLP for Financial Markets / J. Lee, Y. Kim // Springer Publishing. – 2020. – 180 p.
16. Anderson T. Transformers in Financial Text Mining / T. Anderson, K. Liu // Advances in Neural Information Processing Systems. – 2021. – Vol. 35. – P. 12–30.
17. Roberts J. QuantConnect and the Future of Algorithmic Trading / J. Roberts // Quantitative Analysis Journal. – 2021. – Vol. 15, No. 4. – P. 211–225.
18. QuantConnect. QuantConnect Documentation [Электронный ресурс]. – Режим доступа: <https://www.quantconnect.com/docs/>
19. Smith D. Trading API Integration: A Case Study of Alpaca / D. Smith, R. Johnson // Financial Tech Review. – 2022. – Vol. 18, No. 6. – P. 401–417.
20. Alpaca. Alpaca API Documentation [Электронный ресурс]. – Режим доступа: <https://alpaca.markets/docs/>
21. MetaTrader 5. MetaTrader 5 Platform [Электронный ресурс]. – Режим доступа: <https://www.metatrader5.com/en>
22. Bloomberg Professional. Bloomberg API Documentation [Электронный ресурс]. – Режим доступа: <https://www.bloomberg.com/professional/support/documentation/>

23. Stock Bot. Stock Bot Documentation [Електронний ресурс]. – Режим доступу: <https://stockbot.com/>

24. Finnhub. Finnhub API Documentation [Електронний ресурс]. – Режим доступу: <https://finnhub.io/docs/api>

25. Лучишин Б.А. Порівняльний аналіз методів машинного навчання та часових рядів для прогнозування фондового ринку: підхід бота Discord / Б.А. Лучишин // Інформаційне суспільство: технологічні, економічні та технічні аспекти становлення (випуск 93): матеріали Міжнародної наукової інтернет-конференції (м. Тернопіль, Україна, м. Ополе, Польща, 12-13 листопада 2024 р.) / редкол.: О. Патряк та ін.; ГО “Наукова спільнота”, WSZIA w Opolu. – Тернопіль: ФО-П Шпак В.Б., 2023. – С. 24. – [Електронний ресурс]. – Режим доступу: <http://www.konferenciaonline.org.ua/ua/article/id-1948>.

26. Discord. Developer Portal Documentation. [Електронний ресурс]. – Режим доступу: <https://discord.com/developers/docs/intro>

27. Alpha Vantage. Alpha Vantage API Documentation. [Електронний ресурс]. – Режим доступу: <https://www.alphavantage.co/documentation/>

28. Лучишин Б.А. Розробка бота Discord для аналізу фондового ринку з використанням методів машинного навчання / Б.А. Лучишин // Матеріали X Всеукраїнської науково-практичної конференції студентів, аспірантів та молодих вчених «Інтелектуальні комп'ютерні системи та мережі», 5 листопада 2024 р., м. Тернопіль, Україна. – Тернопіль: Західноукраїнський національний університет, кафедра комп'ютерної інженерії, 2024. – С. 59. – [Електронний ресурс]. – Режим доступу: <https://ki.wunu.edu.ua/conference/index.php>.

29. Huang, C., & Petukhina, A. (2022). Applied Time Series Analysis and Forecasting with Python. Springer. – Режим доступу: <https://link.springer.com/book/10.1007/978-3-031-13584-2>

30. Li, W., & Law, K. L. E. (2024). Deep Learning Models for Time Series Forecasting: A Review. IEEE Access, 12, 92306-92327. – [Електронний ресурс] – Режим доступу: <https://doi.org/10.1109/ACCESS.2024.3422528>

31. Aich, S., & Kim, H. C. (2021). Stock Market Prediction Using Machine Learning Techniques: A Decade Survey on Methodologies, Recent Developments, and

Future Directions. Electronics, 10(21), 2717. [Електронний ресурс]. – Режим доступу: <https://doi.org/10.3390/electronics10212717>

32. Комар М.П., Саченко А.О., Васильків Н.М., Загородня Д.І. Методичні рекомендації до виконання кваліфікаційної роботи з освітньо-професійної програми «Комп'ютерні науки» спеціальності 122 «Комп'ютерні науки» за другим (магістерським) рівнем вищої освіти. – Тернопіль: ЗУНУ, 2024. – 32 с.

Додаток А
Копії публікацій

ЗАХІДНОУКРАЇНСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ
ФАКУЛЬТЕТ КОМП'ЮТЕРНИХ ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ
КАФЕДРА КОМП'ЮТЕРНОЇ ІНЖЕНЕРІЇ



**ВСЕУКРАЇНСЬКА НАУКОВО-ПРАКТИЧНА
КОНФЕРЕНЦІЯ СТУДЕНТІВ, АСПІРАНТІВ ТА
МОЛОДИХ ВЧЕНИХ
«ІНТЕЛЕКТУАЛЬНІ КОМП'ЮТЕРНІ СИСТЕМИ ТА
МЕРЕЖІ»**

5 ЛИСТОПАДА 2024



KI.WUNU.EDU.UA/CONFERENCE/

ТЕРНОПІЛЬ

2024



ПРОГРАМНИЙ КОМІТЕТ КОНФЕРЕНЦІЇ

Десятнюк О.М., ректор Західноукраїнського національного університету,
д-р екон.наук, професор;
Дивак М.П., д-р технічних наук, професор, проректор з наукової роботи
Західноукраїнський національний університет;
Березький О.М. д-р.техн.наук, професор, професор кафедри
комп'ютерної інженерії Західноукраїнський національний університет;
Антощук С.Г., д.т.н, професор, Національний університет «Одеська
політехніка»
Батько Ю.М. к.т.н., доцент, Західноукраїнський національний університет
Винокурова О.А., д.т.н., професор, Львівський національний університет
імені Івана Франка
Возна Н.Я. д.т.н., професор, Західноукраїнський національний
університет
Говорущенко Т.О., д.т.н., професор, Хмельницький національний
університет
Дацко Т.В., к.м.н, доцент, Тернопільський національний медичний
університет імені І. Я. Горбачевського
Дубчак Л.О. к.т.н., доцент, Західноукраїнський національний університет
Дунець Р.Б., д.т.н., професор, НУ "Львівська політехніка"
Ізонін І.В. к.т.н., доцент, НУ "Львівська політехніка"
Литвиненко В.І. д.н.т, професор, Херсонський національний технічний
університет
Лупенко С.А., д.н.т, професор, Тернопільський національний технічний
університет імені Івана Пулюя
Мельник Г.М. к.т.н, доцент, Західноукраїнський національний університет
Пелешко Д.Д., д.т.н., професор, Львівський національний університет
імені Івана Франка
Піцун О.Й., к.т.н. доцент, Західноукраїнський національний університет
Сельський П.Р., д.м.н., професор, Тернопільський національний
медичний університет імені І. Я. Горбачевського
Субботін С.О., д.т.н., професор, Національний університет «Запорізька
політехніка»
Теслюк В.М., д.т.н., професор, НУ "Львівська політехніка"
Тимченко Л.І., д.т.н., професор, Державний університет інфраструктури
та технологій
Цмоць І.Г., д.т.н., професор, НУ "Львівська політехніка"
Яровий А.А., д.т.н., професор, Вінницький національний технічний
університет
Яцків В.В., д.т.н., професор, Західноукраїнський національний університет

ОРГАНІЗАЦІЙНИЙ КОМІТЕТ КОНФЕРЕНЦІЇ

Мельник Г.М., канд.тех.наук, доцент кафедри комп'ютерної інженерії
Західноукраїнський національний університет

Піцун О.Й., канд.тех.наук, доцент кафедри комп'ютерної інженерії
Західноукраїнський національний університет

Лящинський П.Б., доктор філософії з комп'ютерних наук, кафедри
комп'ютерної інженерії Західноукраїнський національний університет

Галунька Богдан, студент спеціальності Комп'ютерна інженерія

Фаєрчук Вадим, студент спеціальності Комп'ютерна інженерія

Всеукраїнська науково-практична конференція студентів, аспірантів та
молодих вчених «Інтелектуальні комп'ютерні системи та мережі». 5
листопада 2024 р. Тернопіль. Україна

Метою конференції є представлення та обговорення наукових і практичних результатів, сприяння активізації творчої і інноваційної діяльності студентів, аспірантів і молодих вчених.

Конференція проводиться із залученням Ради Молодих Вчених та Студентського Наукового Товариства ФКІТ ЗУНУ.

Адреса:

Вул. Теліги, 8, корпус №6, Тернопіль

URL: <http://ki.wunu.edu.ua/conference/>

e-mail: kistudconference@gmail.com

Тези доповідей подаються за оригіналом рукопису

<i>Комар В.А.</i> Інтеграція голосових команди та жестів для управління "розумним" будинком.....	55
<i>Копильчак О. А.</i> Гібридний алгоритм для формування ефективних рекомендацій навчальних курсів	57
<i>Лучишин Б.А.</i> Розробка бота Discord для аналізу фондового ринку з використанням методів машинного навчання.....	59
<i>Мамчур Т. Б.</i> Вдосконалення вибору елементної бази для реалізації інтелектуальних компонентів мобільних робототехнічних систем.....	61
<i>Мірута М. Ю., Богатирчук В.О.</i> Аналіз сучасних систем охорони «розумного» будинку.....	63
<i>Наконечний Т. М.</i> Моделі і засоби системи автоматизованого підбору альтернативних джерел живлення.....	65
<i>Опалінський І.Ю.</i> Ідентифікація вільних та зайнятих паркомісць за допомогою глибинних нейронних мереж	67
<i>Остапчук Є.В.</i> Концепція кіберфізичної системи управління складом на основі машинного зору	69
<i>Островський А.П.</i> Автоматизація збору та обробки даних смарт-контрактів для відстеження винагород у децентралізованих системах	71
<i>Патра В.М., Шпак В.О.</i> Алгоритм обробки та аналізу біомедичного зображення.....	73
<i>Рибанчук Д.М.</i> Концепція кіберфізичної системи енергетичного моніторингу житлових будівель на основі комп'ютерного зору	75
<i>Шавров І.С.</i> Архітектура роботизованої системи по догляду за зимовим садом.....	77
<i>Шпак В.О., Патра В.М.</i> Аналіз алгоритмів сегментації цифрових зображень	79
<i>Сворінь К.Г.</i> Концепція мобільно-орієнтованої інформаційної системи моніторингу чистоти довкілля	81
<i>Ткалич Д.С.</i> Архітектура мобільно-орієнтованої інформаційної системи вибору місця для відпочинку	84
<i>Тропак О.Ю., Жеребний О.В.</i> Аналітичний огляд технологій розпізнавання рукописних текстів	87
<i>Яворівський В.А., Партика П.М.</i> Функції активації та архітектури нейронних мереж для розпізнавання обличчя і аналізу емоцій	89
<i>Жеребний О.В., Тропак О.Ю.</i> Алгоритми контурного аналізу для опису та кодування об'єктів цифрових зображень...91	
<i>Рябий В.В.</i> Використання нейромережевих моделей для створення персоналізованих гороскопів у вебзастосунках.....	93

Лучишин Б.А.
 магістрант, 2 курс, ФКІТ ЗУНУ
 Науковий керівник к.т.н., доц. Биковий П.С., кафедра ІОСУ ЗУНУ

РОЗРОБКА БОТА DISCORD ДЛЯ АНАЛІЗУ ФОНДОВОГО РИНКУ З ВИКОРИСТАННЯМ МЕТОДІВ МАШИННОГО НАВЧАННЯ

Вступ. З розвитком фінансових технологій застосування алгоритмів машинного навчання у сфері фінансового аналізу стає дедалі популярнішим. Важливість такого прогнозування зростає разом зі складністю фондових ринків, що вимагає оперативної обробки великих обсягів даних для підтримки інвестиційних рішень [1]. В останні роки методи ансамблевого та глибокого навчання показують обнадійливі результати у прогнозуванні фінансових показників завдяки своїй здатності адаптуватися до мінливих умов ринку. Метою цієї роботи є розробка Discord-бота, що забезпечує автоматизований аналіз ринкових даних і прогнозує ціни акцій на основі історичних даних.

Постановка задачі. Об'єктом дослідження є процес аналізу фондового ринку за допомогою методів машинного навчання. У сучасному фінансовому середовищі, де обсяги даних постійно зростають, стає необхідним використовувати ефективні алгоритми для їх обробки. Важливість прогнозування цін акцій не лише підвищується, але й вимагає більш глибокого розуміння ринкових трендів і факторів, що впливають на їх зміни.

Предметом дослідження є розробка та впровадження Discord-бота, який автоматизує процес збору, аналізу та прогнозування цін акцій на основі методів машинного навчання. Платформа Discord забезпечує інтерактивний і зручний інтерфейс для користувачів, що дозволяє отримувати прогнози та аналітику в реальному часі. Важливим аспектом є інтеграція різних джерел даних, які можуть включати фінансові новини, статистику та аналітику, що забезпечує комплексний підхід до прогнозування.

Результати даного дослідження можуть бути використані для розробки інструментів, що автоматизують процес аналізу та прогнозування на фондовому ринку. Впровадження таких систем може значно підвищити ефективність прийняття інвестиційних рішень та зменшити ризики, пов'язані з торгівлею акціями.

Це дослідження має на меті не лише автоматизацію процесу аналізу ринкових даних, але й надання інструменту, який може значно поліпшити інвестиційні рішення, зменшити ризики та підвищити прибутковість для користувачів.

Основний матеріал. У процесі розробки прогнозовної моделі дані були поділені на навчальну та тестову вибірки у співвідношенні 80/20, що є оптимальним для забезпечення достатньої точності та запобігання перенавчанню моделі [2]. Даний підхід дозволяє виявляти тренди та закономірності, які можуть залишатися незмінними протягом певного часу. Для обробки та аналізу даних використовувалися бібліотеки Python, такі як pandas (маніпуляція даними), matplotlib (візуалізація результатів) та scikit-learn (реалізація ML-алгоритмів), що є стандартними інструментами у сфері аналізу фінансових даних.

Для підвищення точності прогнозування було реалізовано кілька алгоритмів машинного навчання:

1. RandomForestRegressor – ансамблевий метод, який будує набір дерев рішень та об'єднує їх для підвищення точності. RandomForest є особливо ефективним при обробці великих і різномірних даних, що робить його підходящим для фінансового прогнозування, враховуючи складність і варіативність фондових даних [2, 4].

2. GradientBoostingRegressor – метод, що послідовно вдосконалює модель на основі слабких предикторів. Відмінною особливістю цього підходу є поступове зменшення помилок попередніх кроків, що є корисним для фінансових ринків з високою мінливістю. Використання GradientBoostingRegressor дозволяє гнучко налаштовувати прогнозну модель на обробку специфічних ринкових умов [3].

3. Support Vector Regressor (SVR) – метод, який забезпечує високу точність для малих вибірок даних завдяки використанню опорних векторів, що мінімізують помилки і підвищують стійкість до змін у поведінці ринку [3].

Розробка бота розпочалася з використання discord.py для інтеграції з API Discord, що дозволяє легко взаємодіяти з користувачами платформи та забезпечує відправлення результатів у вигляді повідомлень. Бот зчитує ринкові дані з API Alpha Vantage, що дозволяє автоматично збирати та оновлювати дані в реальному часі. Отримані дані були нормалізовані та підготовлені до навчання моделей, що забезпечує однорідність для всіх змінних та підвищує точність прогнозів [2].

Оцінка якості прогнозів моделей здійснювалася за допомогою метрик, таких як середньоквадратична помилка (MSE), середня абсолютна помилка (MAE) та коефіцієнт детермінації (R^2), які дозволяють комплексно оцінити точність, стабільність і надійність моделей [4]. Зрештою, результати аналізу відправляються у відповідний Discord-канал, де користувачі можуть оперативним отримати прогнози і аналітичну інформацію.

Висновки. Discord-бот, розроблений у цьому дослідженні, надає інвесторам зручний спосіб доступу до прогнозованої інформації щодо ринкових трендів та цін акцій у реальному часі. Використання методів машинного навчання для обробки фінансових даних є перспективним напрямком, що дозволяє поліпшити інвестиційні стратегії, мінімізувати ризики та підвищити ефективність прийняття рішень.

Завдяки автоматизованій обробці даних та використанню сучасних алгоритмів, бот забезпечує високий рівень точності прогнозів, що дозволяє користувачам приймати обґрунтовані рішення на основі аналізу ринкової інформації. Інтеграція з платформою Discord також створює сприятливе середовище для обміну знаннями та досвідом серед трейдерів і інвесторів, що може сприяти формуванню активної спільноти.

Подальший розвиток проекту може включати розширення кількості моделей для порівняння та інтеграцію додаткових джерел фінансових даних, що підвищить гнучкість бота у реагуванні на зміни ринкових умов. Крім того, можливе додавання нових функціональних можливостей, таких як можливість аналізу новин та соціальних медіа для покращення точності прогнозів.

З часом, реалізація більш складних алгоритмів, таких як глибоке навчання, може забезпечити ще більшу точність прогнозів і адаптивність до зміни ринкових умов. Загалом, розроблений Discord-бот має потенціал стати важливим інструментом для інвесторів, що прагнуть використовувати сучасні технології для поліпшення своїх фінансових стратегій.

Список літератури

1. Forecasting in stock market with Machine Learning: A State of Art [Електронний ресурс]. – Режим доступу: <https://ieeexplore.ieee.org/document/10307692>. – IEEE Xplore, 2024.
2. Stock Market Prediction using Machine Learning Techniques: A Systematic Review [Електронний ресурс]. – Режим доступу: <https://ieeexplore.ieee.org/document/10142862>. – IEEE Xplore, 2024.
3. Stock Price Prediction Using Artificial Intelligence: A Literature Review [Електронний ресурс]. – Режим доступу: <https://ieeexplore.ieee.org/abstract/document/10459442>. – IEEE Xplore, 2023.
4. Survey of Stock Market Price Prediction Trends using Machine Learning Techniques [Електронний ресурс]. – Режим доступу: <https://ieeexplore.ieee.org/document/10169745>. – IEEE Xplore, 2024.

www.konferenciaonline.org.ua

**Міжнародна наукова
інтернет-конференція**

**Інформаційне суспільство:
технологічні, економічні
та технічні аспекти становлення**

Випуск 93

ISSN 2522-932X

Google Scholar



AKADEMIA NAUK STOSOWANYCH
WYŻSZA SZKOŁA ZARZĄDZANIA I ADMINISTRACJI
W OPOLE

12-13 листопада 2024 р.

м. Тернопіль, Україна – м. Ополе, Польща
2024

01001

*Лучишин Богдан Андрійович, магістрант,
Західноукраїнський національний університет, м. Тернопіль*

*Науковий керівник: Биковий Павло Євгенович,
кандидат технічних наук, доцент,
Західноукраїнський національний університет, м. Тернопіль*

ПОРІВНЯЛЬНИЙ АНАЛІЗ МЕТОДІВ МАШИННОГО НАВЧАННЯ ТА ЧАСОВИХ РЯДІВ ДЛЯ ПРОГНОЗУВАННЯ ФОНДОВОГО РИНКУ: ПІДХІД БОТА DISCORD

Інтернет-адреса публікації на сайті:

<http://www.konferenciaonline.org.ua/ua/article/id-1948/>

У статті розглянуто порівняльний аналіз моделей машинного навчання та методів часових рядів для прогнозування руху фондового ринку. Оцінено такі моделі, як RandomForestRegressor, GradientBoosting, Support Vector Regressor (SVR) та Long Short-Term Memory (LSTM). Моделі було застосовано до даних цін акцій компанії Apple Inc., і їх ефективність оцінювалась за метриками точності та прогнозування. Аналіз реалізовано в рамках роботи бота Discord, що надає користувачам прогнози в реальному часі.

Прогнозування фондового ринку є складним завданням, яке викликає великий інтерес у фінансовій аналітиці та машинному навчанні. Завдяки високій волатильності та випадковості поведінки цін на акції, створення точних моделей прогнозування потребує використання сучасних методів аналізу часових рядів і машинного навчання. Серед перспективних підходів у цій галузі є нейронні мережі, зокрема LSTM, що добре підходять для обробки послідовних даних завдяки своїм властивостям збереження тривалих залежностей. Ансамблеві методи, такі як випадкові ліси та бустинг, також широко використовуються для підвищення стабільності та точності прогнозування.

У прогнозуванні фінансових часових рядів часто застосовуються моделі на основі Long Short-Term Memory (LSTM), що дозволяють ефективно обробляти послідовні залежності в даних. LSTM є потужним інструментом для збереження інформації протягом тривалих періодів, що є важливим для точного прогнозування фінансових змін. Дослідження показують, що LSTM може досягати високої точності при прогнозуванні фінансових ринків, таких як фондові індекси, хоча ці моделі потребують значних обчислювальних ресурсів та налаштування параметрів для досягнення оптимальної ефективності [1, 2].

За умов обмежених обчислювальних ресурсів, ефективнішими можуть бути ансамблеві методи, зокрема RandomForestRegressor, GradientBoostingRegressor та SVR. RandomForestRegressor використовує ансамбль рішень дерев, що мінімізує ризик перенавчання і забезпечує стабільність прогнозів навіть при наявності шуму в даних [3]. GradientBoostingRegressor покроково коригує помилки, що дозволяє значно підвищити точність моделі [3]. SVR є потужним інструментом для роботи з

нелінійними та високорозмірними даними, що робить його ефективним для обробки складних фінансових даних [1, 3].

Для дослідження використовувалися історичні дані про ціни акцій AAPL, що охоплюють період з 13 грудня 2010 року по 20 травня 2024 року (3381 запис). Дані були поділені на 80% для навчання моделей і 20% для тестування, і структуровані в датафрейм для подальшої обробки, що включала кілька етапів: вибір ознак, зміна типів даних та масштабування.

Метод LSTM базується на рекурентних нейронних мережах (RNN), здатних моделювати довгострокові залежності в часових рядах. Для оцінки ефективності LSTM модель була навчена з різною кількістю епох: 200 та 2000. Під час цього процесу використовувалося масштабування даних за допомогою MinMaxScaler для стабілізації навчання.

Ансамблеві методи використовують комбінацію кількох моделей для підвищення точності. У дослідженні використовувалися RandomForestRegressor, GradientBoostingRegressor та SVR. Моделі були навчені окремо, а результати прогнозів об'єднано за допомогою VotingRegressor.

Метод LSTM не продемонстрував високої точності у прогнозуванні цін акцій AAPL. Середньоквадратична помилка (MSE) для моделі з 200 епохами склала 20412.54, а середня абсолютна помилка (MAE) – 142.79. При збільшенні кількості епох до 2000, ці значення зросли, що вказує на перенавчання моделі. Негативні значення R^2 (-2371406.73 для 200 епох) свідчать про те, що модель не змогла адекватно відобразити залежності в даних.

Ансамбль моделей, що об'єднує RandomForestRegressor, GradientBoostingRegressor та SVR, показав набагато кращі результати. Середньоквадратична помилка склала лише 15.02, а середня абсолютна помилка – 2.58. Коефіцієнт детермінації R^2 дорівнював 0.995, що свідчить про здатність моделі точно пояснювати варіативність у даних. Високий відсоток зміни ціни (697.23%) та висхідний тренд вказують на позитивні прогнози, хоча значна волатильність (56.15%) підтверджує наявність ризиків.

Порівняльний аналіз показав, що метод ансамблю моделей значно перевищує метод LSTM за точністю та стабільністю прогнозів. Моделі RandomForestRegressor, GradientBoostingRegressor та SVR продемонстрували кращу продуктивність у прогнозуванні фондового ринку, тоді як LSTM потребує значних обчислювальних ресурсів і не забезпечив задовільних результатів у цій задачі.

Майбутні дослідження можуть бути спрямовані на інтеграцію більш складних нейронних мереж та гібридних моделей для підвищення точності прогнозів.

Література:

1. Zhao Z., et al. Unlocking the Power of LSTM for Long Term Time Series Forecasting. 2024. [Електронний ресурс]. – Режим доступу: <https://arxiv.labs.arxiv.org/html/2408.10006>.

2. Mroua M., Lamine A. Financial time series prediction under Covid-19 pandemic crisis with LSTM network. 2024. [Електронний ресурс]. – Режим доступу: <https://arxiv.labs.arxiv.org/html/2201.08218>.

3. Chen W., Hussain W., Cauteruccio F., Zhang X. Deep Learning for Financial Time Series Prediction: A State-of-the-Art Review of Standalone and Hybrid Models. Computer Modeling in Engineering & Sciences, 2024, 139 (1), 187-224. [Електронний ресурс]. – Режим доступу: <https://doi.org/10.32604/cmes.2023.031388>.

Додаток Б

Текст програми бота для аналізу фондового ринку

```

import discord
from discord.ext import commands
import numpy as np
import pandas as pd
import requests
from sklearn.ensemble import RandomForestRegressor, GradientBoostingRegressor
from sklearn.svm import SVR
from sklearn.metrics import mean_squared_error, mean_absolute_error, r2_score
from sklearn.model_selection import train_test_split
from sklearn.ensemble import VotingRegressor

# Encapsulate API key securely
API_KEY = "ERVBO2LHJ5SVNGPW"

# Initialize intents
intents = discord.Intents.all()

# Initialize the bot with intents and a command prefix
bot = commands.Bot(command_prefix='!', intents=intents)

# Event: Bot is ready
@bot.event
async def on_ready():
    print(f'Logged in as {bot.user.name}')

# Function to fetch stock data from Alpha Vantage API
def fetch_stock_data(symbol):
    try:
        # API endpoint for stock data from Alpha Vantage
        url =
f'https://www.alphavantage.co/query?function=TIME_SERIES_DAILY&symbol={symbol}&apikey={API_KEY}'

        # Fetch data from the API
        response = requests.get(url)

        # Check if the request was successful
        if response.status_code == 200:
            # Parse JSON response

```

```

        data = response.json()

        if 'Time Series (Daily)' in data:
            return data['Time Series (Daily)']
        else:
            return None
    else:
        return None
except Exception as e:
    print(f'Error fetching stock data: {e}')
    return None

# Function to preprocess and prepare stock data for analysis
def prepare_stock_data(data):
    try:
        df = pd.DataFrame(data).T
        df.index = pd.to_datetime(df.index)
        df.sort_index(inplace=True)
        return df
    except Exception as e:
        print(f'Error preparing stock data: {e}')
        return None

# Function to train machine learning models
def train_models(X_train, y_train):
    try:
        # Initialize base models
        rf_model = RandomForestRegressor(n_estimators=100, random_state=42)
        gb_model = GradientBoostingRegressor(n_estimators=100, random_state=42)
        svm_model = SVR(kernel='rbf')

        # Train base models
        rf_model.fit(X_train, y_train)
        gb_model.fit(X_train, y_train)
        svm_model.fit(X_train, y_train)

        return rf_model, gb_model, svm_model
    except Exception as e:
        print(f'Error training models: {e}')
        return None

# Command: Analyze Stock

```

```

@bot.command(name='analyze')
async def analyze_stock_command(ctx, stock_symbol):
    try:
        # Fetch stock data
        data = fetch_stock_data(stock_symbol)
        if data is not None:
            # Prepare stock data
            df = prepare_stock_data(data)
            if df is not None:
                # Extract features and target variable
                X = np.arange(len(df)).reshape(-1, 1)
                y = df['4. close'].values

                # Split data into train and test sets
                X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2,
random_state=42)

                # Train machine learning models
                rf_model, gb_model, svm_model = train_models(X_train, y_train)

                if rf_model is not None and gb_model is not None and svm_model is not None:
                    # Make predictions using individual models
                    rf_pred = rf_model.predict(X_test)
                    gb_pred = gb_model.predict(X_test)
                    svm_pred = svm_model.predict(X_test)

                    # Combine predictions using VotingRegressor
                    ensemble_pred = (rf_pred + gb_pred + svm_pred) / 3

                    # Evaluate accuracy
                    mse = mean_squared_error(y_test, ensemble_pred)
                    mae = mean_absolute_error(y_test, ensemble_pred)
                    r2 = r2_score(y_test, ensemble_pred)

                    # Determine accuracy description
                    accuracy_description = ""
                    if r2 >= 0.8:
                        accuracy_description = "The model is highly accurate."
                    elif r2 >= 0.6:
                        accuracy_description = "The model is moderately accurate."
                    elif r2 >= 0.4:
                        accuracy_description = "The model is fairly accurate."

```

```

else:
    accuracy_description = "The model is not very accurate."

    # Calculate percentage change in stock price over forecast period
    start_price = float(df.iloc[-len(y_test)][ '4. close' ])
    end_price = float(df.iloc[-1][ '4. close' ])
    percentage_change = ((end_price - start_price) / start_price) * 100

    # Analyze trend
    trend = "Upward" if ensemble_pred[-1] > ensemble_pred[0] else "Downward"
    if ensemble_pred[-1] < ensemble_pred[0] else "Stable"

    # Calculate volatility
    volatility = np.std(ensemble_pred)

    # Output results
    result = f"Stock Symbol: {stock_symbol}\n"
    result += f"Forecast Period: {len(y_test)} days\n"
    result += f"Mean Squared Error: {mse}\n"
    result += f"Mean Absolute Error: {mae}\n"
    result += f"R-squared: {r2}\n"
    result += f"Accuracy Description: {accuracy_description}\n"
    result += f"Percentage Change: {percentage_change:.2f}%\n"
    result += f"Trend: {trend}\n"
    result += f"Volatility: {volatility}\n"

    await ctx.send(result)
else:
    await ctx.send('Error: Failed to train machine learning models.')
else:
    await ctx.send('Error: Failed to prepare stock data.')
else:
    await ctx.send(f'Error: Data not available for {stock_symbol}')
except Exception as e:
    await ctx.send(f'Error analyzing stock: {e}')

# Run the bot with your token
bot.run('MTIzMjY2MTQxOTY2MTcyNTcxNg.GMORc4.DHIV6Xwa7UkF69CwZ1R0pqvp4TheEvzSQR2WXk')

```