

**Київський національний університет імені Тараса Шевченка**

**Економічний факультет**

**Кафедра економічної кібернетики**

**КВАЛІФІКАЦІЙНА РОБОТА БАКАЛАВРА**

**«Моделювання інвестиційного ризику компаній AI-сегменту»**

студента 4 курсу  
спеціальності 051 «Економіка»  
ОПП «Економічна кібернетика»  
денної форми навчання  
Легкоступова Данііла Олеговича

**Науковий керівник:**

доктор економічних наук, професор  
Камінський Андрій Борисович

Засвідчую, що в цій роботі немає  
запозичень із праць інших авторів без  
відповідних посилань

Студент \_\_\_\_\_  
(підпис)

Роботу допущено до захисту перед ЕК рішенням

кафедри економічної кібернетики від 12 червня

2025 р., протокол № 15

Завідувач кафедри: доктор економічних наук,  
професор Ляшенко Олена Ігорівна \_\_\_\_\_

(підпис)

## Зміст

<b>ВСТУП</b> .....	<b>3</b>
<b>1 Світовий інвестиційний ринок: Структура та ризики</b> .....	<b>5</b>
<b>1.1 Фондові ринки</b> .....	<b>5</b>
<b>1.2 Ринки облігацій з фіксованим купоном</b> .....	<b>6</b>
<b>1.3 Ринки деривативів</b> .....	<b>8</b>
<b>1.4 Товарні ринки</b> .....	<b>9</b>
<b>1.5 Валютні ринки</b> .....	<b>10</b>
<b>1.6 Нерухомість та альтернативні інвестиції</b> .....	<b>10</b>
<b>1.7 Ризики на світовому інвестиційному ринку</b> .....	<b>12</b>
<b>1.8 Ризики, специфічні для ринку</b> .....	<b>14</b>
<b>2 Торгівля на фондовому ринку за допомогою штучного інтелекту: Учасники, технології та тренди</b> .....	<b>16</b>
<b>2.1 Впровадження торгівлі штучним інтелектом на інституційному     рівні</b> .....	<b>17</b>
<b>2.2 Моделі та методи штучного інтелекту, що використовуються в     трейдингу</b> .....	<b>23</b>
<b>2.3 Тенденції, показники та ефективність торгівлі за допомогою     штучного інтелекту</b> .....	<b>28</b>
<b>3 Моделювання інвестиційного ризику компаній AI-сегменту</b> .....	<b>35</b>
<b>3.1 Набір даних та етапи підготовки для побудови моделі</b> .....	<b>36</b>
<b>3.2 Дослідження важливості ознак та створення цільових ознак для     регресії та класифікації</b> .....	<b>44</b>
<b>3.3 Random Forest</b> .....	<b>50</b>
<b>3.3 Gradient Boosting</b> .....	<b>54</b>
<b>3.4 Logistic regression</b> .....	<b>57</b>
<b>3.5 Порівняння результатів класифікації</b> .....	<b>59</b>
<b>Висновок</b> .....	<b>64</b>
<b>Джерела</b> .....	<b>66</b>
<b>Додаток</b> .....	<b>69</b>

## ВСТУП

Штучний інтелект (ШІ) - це галузь комп'ютерних наук, яка займається розробкою комп'ютерних систем з ціллю, щоб виконувати задачі, які вимагають людського інтелекту заради їх вирішення. Водночас значна частина завдань штучного інтелекту виконується завдяки машинному навчанню.

Машинне навчання (МШ) – це різновид штучного інтелекту, який використовує результат навчання на наборі даних для створення моделі, здатної виконувати складні задачі. [33]

Можливість штучного інтелекту відкриває дуже великий спектр потенційних сфер застосування машинного навчання, які можна уявити практично у всіх сферах життя. Все більше компаній визнають цей аспект, в результаті чого використання штучного інтелекту динамічно зростало протягом останніх кількох років і, згідно з прогнозами, буде продовжувати зростати. Розрізняють два типи компаній:

Компанії, які інтегрують штучний інтелект у свій бізнес-процес для підвищення ефективності, продуктивності та якості обслуговування клієнтів, використовують ШІ в якості інструмента для створення доданої вартості та конкурентної переваги. Поряд з ними існують компанії, для яких розробка продуктів або технологій на основі штучного інтелекту є основою їхньої бізнес-моделі. Саме ці компанії формують ядро AI-сегменту та вважаються «прихованими чемпіонами» галузі. Компанії обох типів, що активно працюють з технологіями ШІ, вже сьогодні отримують вищу оцінку на фондовому ринку та демонструють випереджаюче зростання ринкової капіталізації порівняно з традиційними підприємствами. Штучний інтелект стрімко набуває популярності, стаючи інструментом для оптимізації та автоматизації робочих процесів у великій кількості галузей по всьому світу. Зростаючий попит на ШІ-рішення стимулює компанії впроваджувати інноваційні розробки, що веде до значного розширення можливостей цієї технології.

На сьогоднішній день штучний інтелект використовується для виконання різноманітних завдань, від збору та аналізу даних до розпізнавання мови, оптимізації ланцюгів постачання та візуалізації в медичній сфері. Цей спектр можливостей робить штучний інтелект невід'ємним елементом розвитку сучасних технологій, стимулюючи інвестиції та дослідження в цій галузі.

Метою даної дипломної роботи є дослідження ринку штучного інтелекту, його загальний аналіз, а саме де він використовується, які існують ризики, а також моделювання інвестиційного ризику шляхом машинного навчання. Ключовим аспектом даної роботи є аналіз тенденцій розвитку акцій компаній за допомогою машинного навчання, розбивка цих акцій у різні кластери, що допоможе для різних користувачів, як нових у цій сфері, так і більш досвідчених, краще зорієнтуватись, куди можна інвестувати свої кошти.

Об'єкт дослідження: система управління ризиками інвестуванні в компанії, які займаються штучним інтелектом.

Предмет дослідження: розробка та застосування методів штучного інтелекту для побудови системи управління інвестиційними ризиками.

Методи дослідження: машинне навчання, а саме кластеризаційні моделі.

Наукова та/або практична новизна роботи: наукова новизна даної роботи полягає в застосуванні методів машинного навчання саме у такій сфері. Дослідження буде спрямоване на створення оптимального розподілу капіталу для своїх користувачів.

Структура роботи: дана робота складається з вступу, трьох основних розділів, висновку, списку використаних джерел та додатку з посиланням на Google Colab, в якому була виконена побудова моделі.

## **1 Світовий інвестиційний ринок: Структура та ризики**

Глобальні фінансові ринки становлять складну й динамічну систему, що забезпечує перерозподіл капіталу в масштабах усього світу. Вони виконують ключову функцію у сучасній економіці, оскільки з'єднують інвесторів і позичальників із різних країн, сприяючи мобільності фінансових ресурсів та стимулюючи розвиток інвестиційної діяльності. Через фінансові ринки відбувається торгівля різноманітними інструментами, що, у свою чергу, забезпечує необхідну ліквідність, формування ринкових цін і створює передумови для сталого економічного зростання. Зазвичай світовий інвестиційний простір поділяють на кілька основних сегментів: ринок акцій, боргових інструментів (з фіксованим доходом), деривативів, валют та альтернативних активів. Кожен із цих ринків має власну структуру, специфіку інструментів і набір учасників, проте всі вони взаємопов'язані, формуючи єдину глобальну фінансову систему. У межах цієї системи капітал прямує туди, де він може бути використаний найефективніше, а ризики — розподіляються між учасниками або передаються далі через відповідні механізми. У подальшій частині роботи буде розглянуто, як саме функціонують ключові сектори інвестиційного ринку, які інструменти та учасники є для них характерними, а також яку роль вони відіграють у світовій економіці. [2]

### **1.1 Фондові ринки**

Фондові ринки посідають одне з центральних місць у структурі глобальних фінансових ринків і є, ймовірно, найбільш упізнаваними серед широкого загалу. Саме тут відбувається розміщення акцій компаній, які уособлюють частку їх власності, а також здійснюється їх подальший обіг серед інвесторів. Первинне публічне розміщення акцій (ІРО) дозволяє компаніям залучати акціонерний капітал, що надалі може бути спрямований на розвиток, модернізацію виробництва, інноваційні проєкти тощо. У такий спосіб

фінансові ресурси приватних інвесторів трансформуються у джерело зростання бізнесу.

Для інвесторів участь у фондовому ринку - це, перш за все, можливість отримання прибутку. Це може бути як приріст капіталу (зростання вартості акцій), так і регулярні виплати у формі дивідендів. Торгівля акціями здійснюється як на традиційних фондових біржах (зокрема, Нью-Йоркська, Лондонська, Токійська), так і через електронні торговельні системи. Характерною рисою цих платформ є прозорий процес формування цін, що сприяє ефективному розподілу ресурсів і підвищенню довіри до ринку.

Учасниками фондових ринків є як фізичні особи (роздрібні інвестори), так і великі організації — інвестиційні фонди, пенсійні фонди, страхові компанії, які займаються інвестуванням зазвичай не своїх коштів, а являються посередниками між «ринком» і приватними інвесторами. Крім них, важливу роль відіграють професійні трейдери, брокери та маркет-мейкери. Останні забезпечують безперервне котирування цін купівлі й продажу, підтримуючи тим самим ліквідність ринку. Брокери, у свою чергу, виконують посередницьку функцію, з'єднуючи продавців і покупців.

Загалом, фондовий ринок має ключове значення для економіки: він забезпечує бізнес фінансовими ресурсами, сприяє підвищенню прозорості корпоративного управління та виступає індикатором загального економічного стану. Компанії, акції яких котируються на біржі, зобов'язані регулярно публікувати фінансову звітність, що забезпечує певний рівень контролю з боку інвесторів та регуляторів. [2] [5]

## **1.2 Ринки облігацій з фіксованим купоном.**

Ринки з фіксованим купоном, зосереджені на боргових цінних паперах, насамперед облігаціях, виконують важливу роль у фінансовій системі. У цьому сегменті уряди, муніципалітети та корпорації залучають кошти шляхом

випуску облігацій: емітент отримує авансовий капітал, зобов'язуючись виплачувати інвесторам періодичні купони й повернути номінал у кінці терміну.

Світовий ринок таких інструментів включає державні (наприклад, казначейські облігації США або суверенні позики інших країн), корпоративні, муніципальні облігації. Уряди використовують їх для покриття бюджетного дефіциту або фінансування проєктів інфраструктури, а корпорації — для підсилення операційної діяльності чи масштабування бізнесу. Інвестори ж націлені на отримання стабільного доходу, особливо в умовах фіксованих процентів.

Головними учасниками ринку зазвичай є інституційні інвестори, до яких входять - пенсійні фонди, страхові компанії, банки, а також взаємні і суверенні фонди добробуту. Також до цієї групи долучаються індивідуальні інвестори, які шукають надійні джерела доходу. На первинному ринку ключову роль відіграють дилери й андеррайтери, а на вторинному — брокери й електронні торгові платформи.

Облігації визнані менш волатильними, ніж акції, тому державні папери стабільних країн вважаються низькоризиковими активами. Це робить їх привабливими для консервативних портфелів. Вони також є важливим ажіотом у контрольній політиці: центральні банки активно взаємодіють із ринком облігацій у рамках монетарної стратегії. Якщо казати про Україну, то за даними Міністерства фінансів України, станом на 1 січня 2025, обсяг ОВДП в обігу становив понад 1,86 трлн грн, а інвестиції фізичних осіб зросли майже на 50 % протягом 2024. Це свідчить про високий попит на боргові інструменти всередині країни. [3]

### 1.3 Ринки деривативів

Ринок деривативів - це сфера, де торгують спеціальними фінансовими контрактами, вартість яких залежить від інших активів, наприклад, акцій, облігацій, товарів, валют, відсоткових ставок або біржових індексів. Ці контракти можуть або надавати право, або зобов'язувати сторони купити чи продати такий актив у майбутньому на визначених умовах. До найпоширеніших видів деривативів належать ф'ючерси та опціони, які торгуються на біржах, а також свопи й форварди, що зазвичай укладаються напряму між сторонами без участі біржі. Учасники використовують похідні інструменти з кількома важливими цілями:

1. Хеджування ризиків – стратегія захисту інвестора або компанії від потенційних фінансових втрат, пов'язаних з змінами цін, курсом валют чи відсотковою ставкою.

2. Спекуляція - це коли трейдери виконують торгові операції з впевненістю, що в найближчому майбутньому ціна на даний актив зміниться.

Масштаби ринку деривативів вражаючі — його умовна вартість у багато разів перевищує сукупний обсяг світових ринків акцій і облігацій. Ключові учасники: комерційні й інвестиційні банки (позабіржові дилери), хедж-фонди, власні торговельні фірми, менеджери активів, корпоративні казначейства, а також біржі й клірингові центри, які відповідають за розрахунки й зменшення контрагентського ризику.

Попри переваги, деривативи можуть концентрувати ризики — прикладом цього стала роль похідних інструментів у кризі 2008 року, коли недостатнє управління призвело до системних збитків. [6]

## 1.4 Товарні ринки

Товарні ринки - це ринки, на яких торгують необробленими фізичними товарами, такими як метали, енергоресурси та сільськогосподарська продукція. На цих ринках може здійснюватися спотова торгівля (негайна поставка фізичного товару) або торгівля деривативами (контракти на майбутню поставку або обмін цінами). На спотових товарних ринках виробники (наприклад, фермери, гірничодобувні компанії або нафтовики) продають фізичні товари, а кінцеві споживачі або виробники купують їх; ціни тут визначаються попитом і пропозицією на фізичні товари в режимі реального часу. Однак значна частка світової торгівлі товарами відбувається через похідні контракти (ф'ючерси, форварди та опціони), які використовують фізичні товари як базові активи. Найбільші біржі, такі як Чиказька товарна біржа (CME), Міжконтинентальна біржа (ICE) та Лондонська біржа металів, котирують стандартизовані товарні ф'ючерси та опціони, що дозволяє учасникам ринку легко торгувати з прозорим ціноутворенням.

Основними учасниками товарних ринків є виробники та споживачі товарів, комерційні трейдери та компанії, що займаються торгівлею товарами, фінансові інвестори, такі як сировинні фонди або хедж-фонди, а також спекулянти, які прагнуть отримати прибуток від коливання цін. Ці ринки виконують життєво важливу економічну роль, встановлюючи ціни на основні товари та ресурси в усьому світі, а також забезпечують механізм страхування цінових ризиків. Наприклад, фіксуючи ціну за допомогою ф'ючерсів, фермер може застрахуватися від ризику падіння цін на зернові під час збору врожаю, тоді як харчова компанія може захиститися від різкого зростання витрат на сировину. Ціни на сировинні товари часто чутливі до глобальних економічних тенденцій та геополітичних подій (таких як війни чи торговельна політика) і мають тенденцію до циклічності. Сировинні товари також розглядаються як альтернативна інвестиція, яка може бути захищена від інфляції або девальвації валюти - коли загальні ціни зростають або валюти слабшають, матеріальні

активи, такі як золото або нафта, часто зростають у ціні. Сировинні товари часто стають притулком для інвесторів у періоди інфляції та економічної невизначеності. Таким чином, товарні ринки є ключовими для глобальної торгівлі та фінансової стабільності, гарантуючи, що виробники можуть знайти покупців, а споживачі - забезпечити постачання за прозорими цінами. [9]

### **1.5 Валютні ринки**

Валютний ринок (Forex) — це глобальна мережа, де торгуються національні валюти одна проти одної. Це найоб'ємніший і найліквідніший фінансовий ринок: щоденні обсяги торгів перевищують 7,5 трильйонів доларів. Тут немає єдиної біржі — всі операції виконуються через банки, брокерів і трейдерів по всьому світу цілодобово.

Курси валют формуються шляхом постійного співвідношення попиту й пропозиції між парами валют. Метою цього ринку є конвертація валют, тобто компанії обмінюють іноземну виручку на національну валюту, щоб покривати операційні витрати. Головними учасниками, які задіяні в валютних операціях являються комерційні та інвестиційні банки, вони є основними постачальниками ліквідності на ринку. Також до великих учасників відносяться центральні банки, транснаціональні корпорації, хедж-фонди та компанії, які управляють активами

Висока ліквідність та малий спред роблять ринок привабливим. Але курси валюти сильно реагують на макроекономічні фактори (процентні ставки, інфляцію, геополітику), тому Forex слугує індикатором стану світової економіки. [4]

### **1.6 Нерухомість та альтернативні інвестиції**

Нерухомість посідає особливе місце серед інвестиційних активів завдяки своїй подвійній природі — вона водночас є матеріальним (реальним) активом і складовою частиною глобального фінансового ринку. Інвестування

в нерухомість передбачає придбання, утримання, управління, а також продаж або здачу в оренду об'єктів житлового й комерційного призначення, земельних ділянок з метою отримання доходу. На відміну від фінансових інструментів, таких як акції чи облигації, нерухомість є фізичним активом, однак вона відіграє стратегічну роль у формуванні довгострокових інвестиційних портфелів.

За деякими оцінками, загальна ринкова вартість світової нерухомості перевищує капіталізацію ринків акцій і боргових інструментів, що свідчить про її значення як глобального класу активів. Прибутковість для інвесторів формується як за рахунок орендної плати, так і завдяки зростанню ринкової вартості майна з часом. Попри локальність ринків нерухомості, їх сукупна вага в структурі глобальних інвестицій є істотною. Місцева специфіка — попит, пропозиція, економічні умови в конкретному регіоні — суттєво впливає на дохідність об'єктів.

Серед основних учасників ринку — приватні інвестори, девелопери, пенсійні фонди, страхові компанії та інвестиційні фонди нерухомості (REIT). Останні дають можливість вкладати кошти в нерухомість опосередковано через ринок акцій, що підвищує доступність цього інструменту для широкого кола інвесторів.

Нерухомість відносять до альтернативних інвестицій, адже вона відрізняється високими витратами на входження, низькою ліквідністю та залежністю від зовнішнього фінансування (іпотека). Водночас вона приваблива через здатність забезпечувати диверсифікацію ризиків і частково хеджувати інфляцію — орендна плата й вартість активів зазвичай зростають разом із загальним рівнем цін.

Крім нерухомості, до альтернативних інвестицій належать приватні інвестиції в компанії (у тому числі венчурні), інфраструктурні активи, хедж-

фонди, предмети мистецтва та інші нетрадиційні активи. Такі інструменти менш регульовані та прозорі, часто вимагають тривалого інвестиційного горизонту і доступні лише для кваліфікованих інвесторів. У результаті вони використовуються великими гравцями з відповідним аналітичним і управлінським ресурсом.

Також сюди відносяться ETF – біржові інвестиційні фонди, які об'єднують кошти багатьох інвесторів для купівлі набору активів (акцій чи товарів) і торгується на біржі, як звичайна акція. ETF працює по принципу, що інвестори купують частку фонду, фонд повторює структуру певного індексу або активу, а отже ціна ETF змінюється протягом короткого проміжку часу і її можна придбати чи продати на біржі в будь-який момент. [34]

Таким чином, нерухомість і альтернативні активи не тільки доповнюють традиційні ринки акцій і облігацій, але й дозволяють інвесторам впроваджувати гнучкіші стратегії розподілу активів, знижуючи залежність від волатильності публічних ринків. [7]

### **1.7 Ризики на світовому інвестиційному ринку**

Загалом ризики поділяють на систематичні та несистематичні. До систематичних ризиків відносяться ризики, які впливають на весь ринок або більшість активів одночасно. Таких ризиків неможливо уникнути навіть шляхом диверсифікації. Несистематичні ризики пов'язані з окремою компанією, сектором або галуззю. Такі ризики можна знизити або усунути шляхом диверсифікації.

Світовий інвестиційний ринок постійно перебуває під впливом різноманітних ризиків, які можуть дестабілізувати його роботу й негативно позначитися на прибутковості для інвесторів. Ці ризики можна поділити на дві великі групи: макроекономічні - тобто ті, що пов'язані з загальним станом економіки та політики, та ринкові, які виникають усередині самої фінансової

системи або зумовлені динамікою ринку. Водночас, ці фактори часто переплітаються: наприклад, економічний шок може спричинити фінансову нестабільність, а кризи на ринках можуть посилити макроекономічні проблеми. Для інвесторів, аналітиків і державних органів важливо враховувати ці ризики при формуванні стратегій захисту фінансової стабільності.

Одним із найважливіших макроекономічних викликів для ринків є інфляція. При її зростанні знижується реальна купівельна спроможність грошей, тобто доходи інвесторів фактично знецінюються. Особливо це впливає на ринки облігацій: якщо папери мають фіксований дохід, то інфляція «з'їдає» їхню вартість. У таких умовах центральні банки часто реагують підвищенням ключових відсоткових ставок, що саме по собі створює нову хвилю ризиків: дорожчають кредити, інвестори переоцінюють цінність активів, а вартість раніше випущених облігацій знижується, бо нові пропонують вищу дохідність. Крім того, ріст ставок може негативно вплинути на фондовий ринок: збільшується ставка дисконту, що знижує вартість очікуваних прибутків компаній, а отже - і ціни акцій.

Ще одним важливим макроекономічним фактором є економічні цикли, зокрема рецесії. Коли темпи економічного зростання сповільнюються або настає спад, компанії скорочують витрати, зменшуються прибутки, зростає кількість боргових дефолтів. У такі періоди інвестори уникають ризикових активів, що спричиняє падіння цін на акції, зростання дохідності по боргових інструментах і загальну нестабільність на ринках. На прикладі 2023 року аналітики вже фіксували, що основними загрозами для фінансових перспектив залишаються уповільнення глобального ВВП і стійка інфляція.

Варто згадати і про геополітичні ризики, які останніми роками стали особливо актуальними. Збройні конфлікти, політичні кризи, дипломатичні протистояння можуть миттєво вплинути на ринки: порушуються торговельні

ланцюги, зростають ціни на сировину, падає довіра до компаній і держав. Міжнародний валютний фонд у своїх звітах неодноразово підкреслював, що серйозні геополітичні потрясіння (наприклад, війни чи терористичні акти) мають довгостроковий негативний вплив на фінансові ринки: волатильність зростає, капітал швидко виводиться з ринків, які вважаються вразливими. Більш того, через фінансову взаємозалежність країн такі події можуть поширюватися: криза в одній країні здатна викликати ефект доміно для інших.

Не менш важливо враховувати політичні й регуляторні ризики. Раптові зміни у податковій політиці, введення торгових обмежень або зміна урядових економічних курсів можуть створити значну невизначеність. Наприклад, перехід до протекціоністської моделі торгівлі або посилення контролю над фінансовими потоками часто сприймаються інвесторами як негативні сигнали. Те ж саме стосується і невдалих дій центральних банків - втрата довіри до них може призвести до дестабілізації інфляційних очікувань і процентних ставок, що вдарить по всіх класах активів. [8] [14]

### **1.8 Ризики, специфічні для ринку.**

Фінансові ринки - це не лише інструмент для ефективного розподілу капіталу, а й складна система, яка має свої внутрішні загрози. Однією з основних проблем, що виникає зсередини самої ринкової структури, є волатильність. Якщо коротко, це коливання цін. Невеликі коливання - це нормально і навіть корисно для ринку. Проте коли ціни змінюються занадто різко й непередбачувано, це може створити паніку. Інвестори починають масово продавати активи, і це спричиняє ще більші цінові обвали. Особливо вразливими до цього є ті, хто працює з кредитним плечем - вони змушені швидко позбуватись активів, щоб покрити втрати, і запускають «ланцюгову реакцію» нестабільності. Такі ситуації іноді навіть спричиняють технічні збої в торгівлі.

Ще одне явище, яке загрожує ринку - це цінові бульбашки. Вони виникають тоді, коли ціни на активи починають зростати швидше, ніж дозволяють реальні економічні показники. Це часто пов'язано з надмірним оптимізмом інвесторів, спекуляціями або доступними кредитами. Спочатку здається, що це чудова можливість заробити, але коли бульбашка «лопається», ринок стрімко падає, і багато учасників зазнають суттєвих збитків. Яскравий приклад - іпотечна криза у США в 2007 році, яка стала початком глобальної фінансової кризи. Ще один приклад - крах так званих «доткомів» у 2000 році.

Зі спекуляціями й бульбашками тісно пов'язаний і системний ризик. Це вже загроза не окремим компаніям, а всій фінансовій системі. Суть у тому, що проблеми однієї установи (наприклад, великого банку) можуть перекинутись на інші - через борги, фінансові зв'язки чи втрату довіри. Якщо збанкрутує системно важливий учасник ринку, це може викликати лавину негативних наслідків. Такий сценарій ми вже бачили під час банкрутства Lehman Brothers у 2008 році, коли криза в одній установі швидко поширилась на весь світовий фінансовий сектор.

Ще одним з критичних аспектів являється ризик ліквідності. Ризик виникає коли виникають проблеми з швидким продажем активів і без великих цінових втрат, а під час самої кризи, коли всі починають на ринку продавати одночасно, а немає такої кількості охочих придбати товар – ринок обвалюється. Прикладом для цього може слугувати березень 2020 року, коли глобальна падемія почала масово розповсюджуватись у світі, навіть державні облігації США втрачали свою ліквідність, пока не втрутився центральний банк. Тому можна сказати, що світовий інвестиційний банк – чутлива і взаємопов'язана система. До неї входять валюти, акції, облігації, нерухомість деривативи та інші інструменти, які залежать один від одного. Зовнішні шоки (наприклад, війна, інфляція, рецесія) та внутрішні ринкові фактори (спекуляції, відсутність ліквідності чи системні збої) можуть одночасно впливати на всю систему.

Щоб уникати таких ситуацій, потрібне постійне управління ризиками, ефективне регулювання та баланс між підтримкою ринкових механізмів і запобіганням надмірним коливанням. Без цього навіть найсильніші ринки можуть опинитися на межі кризи. [10] [11] [12]

## **2 Торгівля на фондовому ринку за допомогою штучного інтелекту: Учасники, технології та тренди.**

На сьогоднішній день штучний інтелект суттєво змінює підходи до торгівлі на фондовому ринку як на рівні великих фінансових установ, так і серед роздрібних інвесторів. Хедж-фонди, інвестиційні банки та платформи для трейдингу все активніше впроваджують ШІ у свої стратегії. Наприклад, багато фондів використовують моделі машинного навчання для передбачення руху ринку, тоді як мобільні додатки для трейдингу інтегрують аналітику на основі ШІ для підтримки індивідуальних користувачів.

Компанії застосовують широкий спектр технологій - від класичних алгоритмів машинного навчання до складних нейронних мереж і навіть навчання моделей з підкріпленням, тобто навчання шляхом повторення дій. Головна перевага таких систем полягає у здатності швидко аналізувати великі обсяги даних, знаходити закономірності, непомітні для людини, і приймати торгові рішення майже миттєво.

У межах цього дослідження я розглянув, як провідні світові компанії використовують ШІ у трейдингу, які саме моделі вони застосовують, і в чому полягає ефективність таких систем. Окрему увагу приділено аналізу статистичних даних щодо масштабів використання ШІ у торгівлі, обсягів активів під управлінням (AUM), прибутковості ШІ-систем у порівнянні з традиційними підходами, а також прогнозам щодо подальшого зростання цього напрямку.

У роботі також розглянуто регіональні відмінності у впровадженні ШІ, темпи зростання ринку в різних країнах, особливості регулювання і потенційні ризики. В цілому, використання штучного інтелекту у фінансовій галузі формує нові стандарти у сферах, а не тільки банально демонструє високі показники ефективності. [13]

## **2.1 Впровадження торгівлі штучним інтелектом на інституційному рівні.**

Провідні банки, хедж-фонди та компанії з управління активами по всьому світу інтегрували штучний інтелект у свої торгові операції. Провідні інституційні гравці, їхні регіони, технології штучного інтелекту, які вони використовують, та основні переваги їхніх систем:

**J.P. Morgan Chase** (США) створили програму, яка виконує торгові операції на основі ШІ під назвою LOXM. Вона використовує глибоке навчання з підкріпленням і предиктивну аналітику. Тобто алгоритм на основі минулих і теперешніх операцій навчається, щоб в майбутньому приймати кращі рішення. Навчена на мільярдах минулих транзакцій, LOXM може виконувати великі замовлення на акції в режимі реального часу за оптимальними цінами, мінімізуючи вплив на ринок. Під час випробувань ця система ШІ перевершила як ручні, так і існуючі автоматизовані методи, забезпечивши значну економію коштів і на ~15% підвищивши ефективність виконання порівняно з традиційними алгоритмами. J.P. Morgan почав розгортати LOXM в Європі в 2017 році, а пізніше в Азії та США після перших успіхів в оптимізації швидкості/ціни для великих обсягів торгів. [15]

**Royal Bank of Canada - RBC** (Канада) – розробив власну торгову платформу під назвою Aiden у партнерстві зі своєю лабораторією Borealis AI, використовуючи глибоке навчання з підкріпленням для здійснення торгівлі акціями. Алгоритми штучного інтелекту Aiden підлаштовуються під ринкові

умови в режимі реального часу, навчаючись на кожній поточній угоді, щоб зменшити помилки і покращити кращі рішення в майбутньому. Система була вперше запущена в 2020 році і довела свою здатність орієнтуватися на волатильних ринках, не вимагаючи частого перепрограмування людиною. Наприклад, Aiden VWAP (Volume-Weighted Average Price) використовує глибоку нейронну мережу, що враховує понад 200 ринкових вхідних даних, для динамічного виконання ордерів з мінімальним відхиленням від еталонного значення. Автономно досліджуючи торгові стратегії в безпечних межах, Ейден постійно знаходить кращі моделі виконання, допомагаючи клієнтам досягати цін, ближчих до еталонних, і знижувати торгові витрати. Адаптивність ШІ дозволяє трейдерам зосередитися на більш важливих завданнях, в той час як рутинне виконання оптимізується. [16]

**Two Sigma** (США) є однією з найвідоміших компаній, що активно використовує штучний інтелект у торгівлі – це нью-йоркський хедж-фонд, який спеціалізується на кількісних стратегіях та алгоритмах на базі ШІ. Компанія активно впроваджує машинне навчання для пошуку ринкових закономірностей, зокрема застосовує як градієнтно-підсилені дерева, так і глибокі нейронні мережі.

Two Sigma працює з великими обсягами даних, аналізуючи цінові ряди, фінансові новини та текстову інформацію. Завдяки цьому компанія успішно реалізує стратегії статистичного арбітражу - тобто виявляє тимчасові відхилення у цінах активів і заробляє на їх вирівнюванні.

Окрема увага приділяється обробці природної мови (NLP): за допомогою цієї технології Two Sigma видобуває корисні сигнали з текстів - новин, аналітики, повідомлень у ЗМІ, що дозволяє швидко реагувати на зміни ринку. Компанія торгує різними активами - акціями, ф'ючерсами тощо - і робить це в автоматичному режимі з високою частотою.

III-моделі в Two Sigma не статичні - вони адаптуються до змін ринкових умов, що забезпечує компанії високу результативність у порівнянні з класичними фондами. Їхній підхід базується на постійному вдосконаленні моделей, масштабованій інфраструктурі та поєднанні наукового аналізу з фінансовою практикою. [18]

**Bridgewater Associates** (США) - найбільший у світі хедж-фонд (базується в штаті Коннектикут), який управляє близько 171 мільярда доларів під управлінням. Bridgewater вже давно є піонером у систематичній торгівлі і за останнє десятиліття створив Лабораторію штучного інтелекту (AI Lab), в якій працюють десятки інженерів та дослідників машинного навчання. Лабораторія III досліджує застосування машинного навчання (включаючи глибоке навчання) до макро- та чисто альфа-стратегій Bridgewater. Системи штучного інтелекту Bridgewater поглинають величезні масиви економічних і ринкових даних (ціни, економічні показники, новини тощо) і вчаться прогнозувати рух активів або оптимізувати портфелі відповідно до інвестиційних принципів фірми. Одним з напрямків є використання навчання з підкріпленням для здійснення торгів та оптимізації маршрутів, подібно до інших банків. Хоча Bridgewater тримає деталі в таємниці, відомо, що вона використовує III для покращення управління ризиками та постійного . Мета полягає в тому, щоб систематично фіксувати неефективність на глобальних ринках, поєднуючи людське розуміння з пошуком закономірностей за допомогою штучного інтелекту. [17]

**Renaissance Technologies** (США) - великий кількісний хедж-фонд, який базується в Нью-Йорку, що управляє понад 130 мільярдами доларів (включаючи його знаменитий фонд Medallion). RenTech був попередником у застосуванні математики, статистики та машинного навчання в торгівлі.

Моделі RenTech, були розроблені командами докторів наук, які використовують алгоритми машинного навчання на петабайтах ринкових і

альтернативних даних для виявлення прогностичних сигналів. Вони були першими, хто застосував методи, подібні до сучасного штучного інтелекту: нелінійні моделі, дерева рішень і нейронні мережі, навчені на історичних прикладах. Як повідомляється, Renaissance використовує ансамблеві моделі і постійно перенавчає їх на нових даних, щоб адаптуватися до ринкових змін, що дає їй перевагу в короткостроковій торгівлі. Результатом стала надзвичайна продуктивність - її фонд Medallion, як відомо, згенерував понад 100 мільярдів доларів прибутку з 1988 по 2020 рік.

Станом на 2024 рік RenTech залишається лідером у торгівлі на основі штучного інтелекту: згідно з даними регуляторних органів, компанія управляє цінними паперами на суму близько 69 мільярдів доларів США та загальним активом під управлінням близько 106 мільярдів доларів США. Фірма приписує значну частину свого успіху автоматизованим системам навчання, які розвивають стратегії швидше, ніж будь-який підхід, заснований на людському факторі. [19]

**Citadel LLC** (США) - хедж-фонд зі штаб-квартирою в Чикаго. На 2024 рік компанія керувала активами розміром більше 60 мільярдів доларів США і застосовувала штучний інтелект і машинне навчання у щоденній роботі.

Citadel працює як мультистратегічний фонд і маркет-мейкер, тобто компанія не лише управляє інвестиціями, але й забезпечує ліквідність на ринку. Торгові системи активно використовують моделі глибокого машинного навчання для обробки та виявлення закономірностей у даних, особливо увага приділяється ринку акцій. Окрім цього, навчання з підкріпленням застосовується для адаптації параметрів торгів у режимі реального часу. Це дозволяє компанії максимально ефективно виконувати угоди, особливо у високочастотній торгівлі (HFT).

Хоча остаточне рішення про угоди приймається портфельними менеджерами, але моделі штучного інтелекту відіграють ключову роль у зборі, аналізі та поданні інформації. Вони обробляють великі обсяги даних: новини, котирування, зміну глибини ринку та внутрішню динаміку портфелів. Завдяки цьому Citadel швидко реагує на зміни, що й робить її однією з найефективніших інвестиційних компаній сучасності. [20]

**Man Group** (Великобританія) - одна з найбільших європейських хедж-фірм (базується в Лондоні) з активами під управлінням 144 мільярди доларів США, чий підрозділ Man AHL систематично впроваджує штучний інтелект. Man AHL почав торгувати моделями на основі машинного навчання у своїх клієнтських портфелях ще у 2014 році. Сьогодні вона використовує такі методи, як глибоке навчання для розпізнавання образів і обробка природної мови для текстових даних (наприклад, аналіз новинних настроїв). Man також досліджував навчання з підкріпленням для здійснення торгівлі та оптимальної маршрутизації ордерів.

У 2023 році Man Group запустила нову програму «Alpha AI», щоб координувати свої зусилля в області ШІ і ML між торговими командами. Фірма повідомляє, що багато її стратегій (керовані ф'ючерси, довгі/короткі акції) тепер включають компонент ШІ для покращення вилучення сигналів або управління ризиками. Крім того, системи спостереження зі штучним інтелектом у Man відстежують ринки на предмет аномалій або можливого шахрайства (інсайдерська торгівля). Використання штучного інтелекту в Man Group допомогло підвищити продуктивність і тримати квантові стратегії на передньому краї індустрії.

**XTX Markets** (Велика Британія) - це лондонська фінансова компанія, яка з 2015 року активно працює в сфері алгоритмічної торгівлі, особливу увагу приділяє валютному сектору, а також на ринках акцій та облігацій. Незважаючи на відносно недавню появу на ринку, XTX швидко увійшла до

списку провідних електронних маркет-мейкерів світу, що значною мірою пов'язано з глибокою інтеграцією штучного інтелекту у всі аспекти своєї діяльності.

Компанія розробила високоточні моделі глибокого навчання, які тренуються на величезних масивах історичних тикових (міліметрових) ринкових даних. Ці моделі дають змогу в режимі реального часу формувати ціни, управляти ризиками та ліквідністю, аналізуючи поведінку ринку, макроекономічні сигнали та динаміку заявок. Усе це дозволяє утримувати вузькі спреди між купівлею та продажем, зменшувати витрати і забезпечувати прибуткову торгівлю навіть за високої волатильності.

На відміну від традиційних торгових підходів, алгоритми ХТХ працюють майже автономно — з мінімальним людським втручанням. Вони постійно самооновлюються та адаптуються до ринкових змін, що дає компанії суттєву перевагу над конкурентами, особливо у високочастотній торгівлі. ХТХ має глобальну присутність — її офіси відкриті в Європі, Азії та Північній Америці, що ще раз підкреслює ефективність та масштабованість моделі бізнесу, побудованої на основі AI. [21]

**Nomura Securities** (Японія). Nomura Securities, один з провідних азійських банків, активно впроваджує штучний інтелект у свої торгові операції. Банк розглядає це як стратегічну інвестицію у своє майбутнє, а не звичайний теперешній тренд на ШІ. Банк запустив тестову платформу для консультацій по інвестуванню на базі ШІ, використовуючи моделі машинного навчання для прогнозування ринкових тенденцій та управління значними клієнтськими замовленнями. Торгові алгоритми Nomura тепер можуть динамічно коригувати графіки виконання ордерів, базуючись на ШІ-прогнозах короткострокового впливу на ціни. Це дозволяє їм бути на крок попереду. Також Nomura активно розвиває цифрові ініціативи, зокрема, "Цифровий офіс Nomura" у Токіо, що слугує інкубатором для розробки нових ШІ-додатків.

Хоча, можливо, ми бачимо менше деталей про їхні досягнення порівняно з американськими банками, Nomura та інші японські фінансові гіганти, такі як Mitsubishi UFJ та SBI Holdings, все частіше звертаються до ШІ для вибору акцій, виконання угод та контролю ризиків. Це чітко показує їхнє прагнення модернізувати свої торгові майданчики. Усе це відображає ширшу тенденцію: ШІ дедалі активніше проникає у фінансові центри Азії, а саме Токіо, Гонконг, Сінгапур. [22]

## **2.2 Моделі та методи штучного інтелекту, що використовуються в трейдингу**

Торгові системи за допомогою штучного інтелекту використовують різні типи моделей, кожен з яких підходить для різних завдань:

**Машинне Навчання (Machine Learning)** – один найпоширеніший підходів AI в торгівлі. До них відносяться такі моделі, як алгоритми регресії, дерева рішень, випадкові ліси, машини з градієнтом та подібні, які навчаються на історичних даних для прогнозування результатів (наприклад, наступного дня повернення, напрямку цін, волатильність). Багато хедж-фонди почали з цих «традиційних» моделей ML. Наприклад, фонд може використовувати модель підвищення градієнта для прогнозування прибутковості акцій на основі сотень вхідних ознак (показників оцінки, технічних показників, показників настроїв тощо). Приблизно половина або більше квантів торгових фірм включили такі алгоритми ML в свій інвестиційний процес в останні роки. Привабливість полягає в тому, що ML може знайти нелінійні відносини та взаємодії, які прості лінійні моделі або люди можуть пропустити. Тим не менш, ці моделі, як правило, все ще вимагають людського керівництва для проектування функцій і тільки так само добре, як надані дані. [23]

**Глибоке навчання (Deep Learning)** - це підхід до штучного інтелекту, заснований на використанні багатоварових нейронних мереж, які за

принципом своєї роботи нагадують людський мозок. Такі нейронні мережі здатні самостійно знаходити складні зв'язки та закономірності в даних, без того, щоб людина вручну прописувала, на які саме ознаки потрібно звертати увагу. Активне використання глибокого навчання в сфері трейдингу почалося наприкінці 2010-х років, коли стало більше доступних даних та зросла потужність комп'ютерної техніки. Це дозволило застосовувати ці моделі для вирішення практичних задач на фінансових ринках. Наприклад, сьогодні активно використовують згорткові нейронні мережі (CNN), щоб аналізувати різноманітні зображення, такі як графіки цін чи навіть супутникові фото. Рекурентні нейронні мережі (LSTM чи RNN), зі свого боку, дуже ефективні для прогнозування часових рядів - наприклад, коли потрібно передбачити майбутні рухи цін на акції чи валюту.

Загалом глибоке навчання добре себе проявляє там, де потрібно обробляти великі обсяги неструктурованих даних — текстів, новин, аудіо, зображень, - і швидко перетворювати їх у конкретні торгові сигнали. Опитування, проведені у 2025 році, показують, що більшість квантових фондів уже або використовують глибоке навчання у своїй діяльності, або активно його тестують.

Основна перевага цих підходів - можливість враховувати дуже складні залежності у великій кількості даних. Проте є і свої складності: глибокі нейронні мережі часто називають «чорними ящиками», оскільки не завжди зрозуміло, яким чином вони доходять до певних висновків. До того ж, якщо не налаштувати такі моделі правильно, вони можуть перенавчитись — тобто занадто добре адаптуватись саме під навчальні дані, погано реагуючи на нову інформацію. [24]

**Обробка природної мови (NLP)** - підгалузь ШІ, що фокусується на розумінні людської мови, NLP став життєво важливим для торгових стратегій, які аналізують текстові дані. Аналіз настроїв новин, соціальних мереж

(наприклад, твітів) і звітності компаній зараз є поширеним. Моделі ШІ класифікують заголовки новин як позитивні, негативні або нейтральні для акцій, дозволяючи трейдерам миттєво реагувати на новини. Наприклад, ШІ EquBot (використовується в AIEQ ETF) використовує НЛП IBM Watson для читання мільйонів новин і фінансових звітів, імітуючи «команду з 1000 аналітиків, які працюють 24/7», скануючи в пошуках інсайтів. Аналогічно, BloombergGPT (LLM) та інші мовні моделі з фінансовою освітою зараз використовуються для генерування торгових сигналів або узагальнення складних звітів. Моделі NLP (від старих методів, таких як мішок слів і TF-IDF, до нових моделей на основі трансформації і LLM в стилі GPT) допомагають трейдерам не тільки оцінювати настрої, а й виявляти посилання на проблеми з ланцюгами поставок, зміни в керівництві та інші текстові підказки, які корелюють з рухом акцій. Приблизно 30% хедж-фондів у 2024 році повідомили, що використовують NLP та альтернативні текстові дані як частину свого інвестиційного процесу, і очікується, що цей показник зростатиме. Основна перевага полягає в систематичному перетворенні якісної текстової інформації на кількісні сигнали. [25]

**Навчання з підкріпленням (RL)** – це метод ШІ, який працює по принципу навчання на основі минулих та поточних даних, щоб навчитися краще працювати алгоритм, задля максимізації винагороди. У торгівлі «агентом» може бути алгоритм, який вирішує, як виконати ордер або як розподілити портфель, а винагородою може бути більший прибуток або менші торгові витрати. RL набув популярності для виконання торгових операцій та оптимізації стратегій. Такі банки, як JPMorgan і РБК, продемонстрували, що глибоке RL може досягти успіху у виконанні великих ордерів з мінімальним впливом на ринок (як показано на прикладі LOXM і Aiden). Алгоритм RL спостерігає за станом ринку і виконує дії (розміщення частин ордерів), отримуючи зворотній зв'язок у вигляді вартості виконання або прослизання; з часом він вивчає політику виконання, яка перевершує статичні алгоритми.

Деякі хедж-фонди також експериментували з RL для управління динамічними портфелями або хеджуванням опціонів, де модель навчається, коли потрібно перебалансувати або скоригувати позиції за невизначених умов. Хоча, за оцінками, лише 10-20% торгових фірм впровадили навчання з підкріпленням у виробництво (станом на середину 2020-х років), його використання зростає в спеціалізованих додатках, таких як маркет-мейкінг (агент RL навчається оптимально котирувати спреди між попитом і пропозицією) і високочастотна торгівля (де мікрорішення можуть бути сформульовані як проблема RL). Перевагою RL є його здатність обробляти послідовні рішення і адаптуватися до мінливого середовища; недоліком є те, що він часто вимагає великих навчальних даних і ретельної розробки винагороди, щоб уникнути непередбачуваних стратегій. [26]

**Генеративний Штучний Інтелект (Generative AI)** - це новий підхід у сфері AI, здатний не лише аналізувати дані, а й створювати їх. Його використовують у фінансах для симуляції ринкових сценаріїв і стрес-тестування стратегій. Наприклад, генеративні моделі можуть «програти» ринок у рідкісні кризи, яких немає в історичних даних - це допомагає трейдерам передбачити й підготуватися до екстремальних ситуацій. Деякі компанії застосовують GenAI для створення штучних фінансових даних, коли реальних недостатньо. Це особливо корисно, якщо потрібно підготувати моделі до незвичної ринкової поведінки чи специфічних активів. Великі мовні моделі, які ми знаємо з ChatGPT, теж використовуються в трейдингу. Наприклад, можна запропонувати моделі написати код торгової стратегії - і вона розробить його - або поставити їй питання про попит на нафту на базі актуальних звітів, а модель дасте розгорнуту аналітичну відповідь, тому це вже не простий генератор текстів, а помічник в ухваленні рішень.

За даними IOSCO (2024), багато учасників ринку вважають, що GenAI майже обов'язково поліпшить алгоритмічну торгівлю. Зокрема, він допоможе обробляти мультимодальні дані (тексти й зображення разом), швидко

оцінювати настрої учасників ринку і приймати рішення на базі ширшої інформації.

Хоча безпосереднє практичне застосування генеративного ШІ в торгівлі все ще на ранній стадії, вже видно перспективи розвитку. GenAI допомагає узагальнювати дані - наприклад, створювати «швидкі резюме» по ринковій ситуації - і генерувати нові стратегічні ідеї на основі аналізу великого масиву даних. У найближчі роки очікується, що інтерес до таких інструментів у фінансовому середовищі зростатиме, хоча і з обережністю - через ризики недостовірності або «галюцинації» моделей. [27]

В алгоритмічній торгівлі застосовують також **Еволюційні алгоритми та федеративне навчання** - хоча вони й залишаються в більш специфічних нішах.

Еволюційні алгоритми імітують природний відбір: систему починають з великої кількості «кандидатів» на стратегії, потім проходять сортування за якістю (прибутковістю чи стабільністю), а найуспішніші стикаються між собою, отримуючи нові варіації. Так працює, наприклад, платформа Aidyia: вона поєднує глибоке навчання з еволюційними алгоритмами, створюючи й постійно розвиваючи автономних «торгових агентів». Це дозволяє автоматично знаходити ефективні параметри стратегій.

Натомість федеративне навчання - це спосіб навчання моделей без централізації даних. Наприклад, кілька банків можуть одночасно тренувати алгоритми для виявлення шахрайства, обмінюючись лише оновленнями моделей, а не самими даними. Це особливо актуально для фінансових інституцій, які не можуть відкривати свої дані через регуляторні обмеження або внутрішню політику безпеки. Федеративне навчання дозволяє їм співпрацювати для покращення моделі, не порушуючи конфіденційність клієнтської інформації.

Зараз такі технології все ще не є загальноприйнятими, але їх потенціал великий - особливо для об'єднаних проєктів між інституціями, біржами та брокерами, наприклад, для колективного виявлення аномальної поведінки на ринку. [28]

### Тенденції, показники та ефективність торгівлі за допомогою штучного інтелекту.

Стрімке зростання ролі штучного інтелекту в трейдингу відображено в різних показниках і дослідженнях.

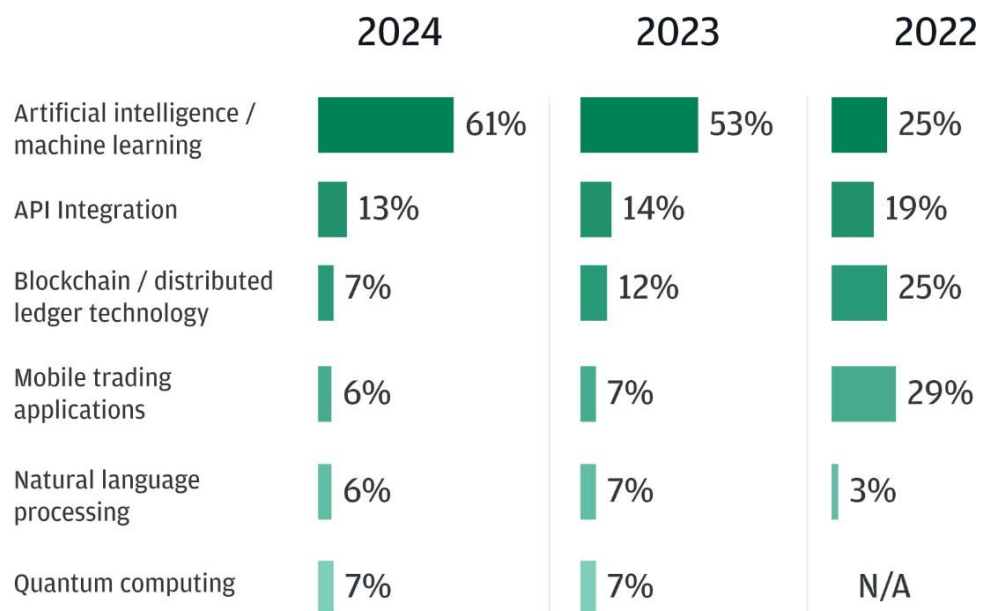


Рис. 1. Технології, що формують майбутнє торгівлі.

За останні кілька років торгівля на основі штучного інтелекту перетворилася з нішевої в майже мейнстрім. Станом на 2022 рік приблизно 25% усіх хедж-фондів використовували ШІ/машинне навчання для генерування інвестиційних ідей або управління портфелями. Фактично, галузеві опитування показали, що до початку 2024 року ті, хто не використовує ШІ, будуть у меншості. Це кардинальна зміна в порівнянні з тим, що було всього кілька років тому. У щорічному глобальному опитуванні J.P. Morgan, присвяченому торгівлі, 61% інституційних трейдерів у 2024 році назвали

AI/ML найвпливовішою технологією на наступні три роки, порівняно з 53% у 2023 році і лише 25% у 2022 році. Це свідчить як про поточне проникнення, так і про зростаючий ентузіазм щодо ШІ в торгівлі. Крім того, опитування 50 найбільших хедж-фондів 2023 року показало, що 9 з 10 менеджерів фондів планують використовувати ШІ для досягнення прибутковості - по суті, це майже повсюдне визнання важливості ШІ. [30]

Таке широке впровадження на елітному рівні просочується вниз: навіть менші фонди і власні торгові магазини все частіше розгортають бібліотеки ШІ і наймають фахівців з аналізу даних. Що стосується роздрібною торгівлі, то кількість торгових додатків та інструментів на основі штучного інтелекту стрімко зростає, а рівень їхнього використання зростає. Брокери повідомляють, що зростає частка роздрібних замовлень, які генеруються або маршрутизуються алгоритмічними системами (часто на основі штучного інтелекту), особливо з появою простих у використанні роботів-консультантів і бот-платформ. Коротше кажучи, штучний інтелект у торгівлі досяг переломного моменту, коли більшість учасників ринку (за обсягом торгів) використовують штучний інтелект у своїх процесах. Обсяг торгів та вплив на AUM: Значна частина глобальної торговельної діяльності зараз виконується або керується алгоритмами штучного інтелекту. За оцінками, понад 60% обсягу торгів у різних класах активів здійснюється не людьми-трейдерами, а автоматизованою торгівлею та алгоритмами машинного навчання. Сюди входить високочастотна торгівля, яка сама по собі може становити половину обсягу фондового ринку в США і значну частину в Європі/Азії - ці стратегії HFT майже повністю автоматизовані і все частіше використовують ШІ для розпізнавання шаблонів.

У хедж-фондах Bloomberg Intelligence повідомила, що квантові стратегії, керовані ШІ, становили понад 40% обсягів торгів у хедж-фондах у 2024 році, що свідчить про те, наскільки сильно торгові потоки залежать від рішень, прийнятих за допомогою ШІ. З точки зору активів, хедж-фонди, що

використовують ШІ, управляють значними і постійно зростаючими активами - американські хедж-фонди сукупно мають понад 5 трильйонів доларів США, і значна частина з них припадає на фонди з процесами, вдосконаленими за допомогою ШІ (наприклад, Renaissance, Bridgewater, Citadel та ін.). Деякі з перших фондів зі штучним інтелектом, запущених у середині 2010-х років, збільшили свої активи, і нові фонди, орієнтовані на штучний інтелект, продовжують з'являтися. Для роздрібних інвесторів ETF на основі AI, такі як AI Powered Equity ETF (AIEQ), принесли вибір акцій на основі AI на публічні ринки. Стратегія AIEQ (керована ШІ на базі IBM Watson від EquBot) щодня аналізує мільйони точок даних про 6000+ американських акцій, імітуючи армію аналітиків. За перший рік свого існування компанія залучила понад 100 мільйонів доларів, продемонструвавши апетит інвесторів до стратегій штучного інтелекту. У міру того, як ШІ стає стандартом, ми бачимо, як традиційні активи принаймні частково управляються за допомогою ШІ (наприклад, платформа Aladdin від BlackRock використовує ШІ для управління ризиками трильйонних активів). Суть в тому, що штучний інтелект - це не просто кричущий додаток, він впливає на значну частку капіталу на ринках сьогодні. Ефективність проти традиційних підходів: Одним із рушіїв впровадження ШІ є обіцянка (і докази) покращення продуктивності. Дослідження показують, що фонди, які використовують штучний інтелект, можуть досягти кращих показників прибутковості та ризику, ніж аналоги. Наприклад, дослідження Cerulli Associates показало, що хедж-фонди, які використовують ШІ, за три роки (2017-2020) сукупно принесли ~34% доходу, що майже втричі перевищує ~12% доходу світової індустрії хедж-фондів за цей період. Це випередження проілюстровано на графіку нижче, де порівнюються 3-річні кумулятивні прибутки хедж-фондів, керованих ШІ, та середньостатистичного хедж-фонду:

## Cumulative 3-Year Returns (through May 2020)

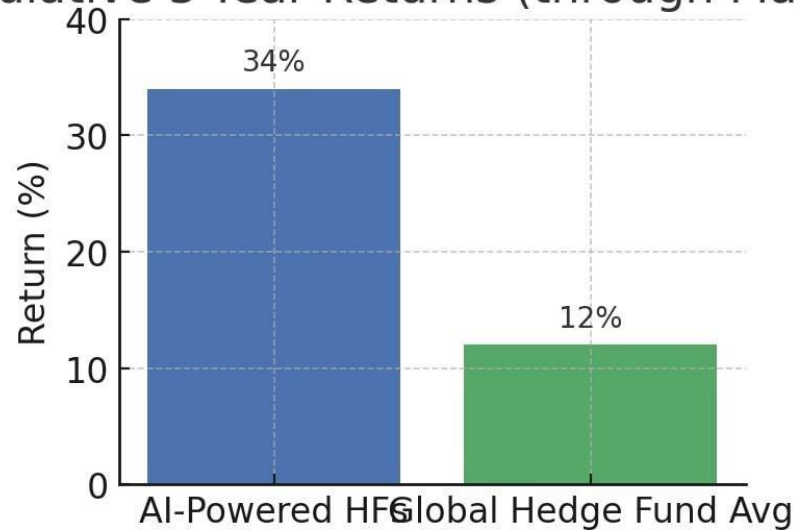


Рис. 2. Порівняння кумулятивної трирічної дохідності (2017-2020) хедж-фондів зі штучним інтелектом із середнім показником світових хедж-фондів. (Фонди на чолі зі штучним інтелектом отримали ~34% проти ~12% для галузі, що відображає потенціал штучного інтелекту для підвищення продуктивності.)

Ця перевага зумовлена багатьма факторами: Фонди зі штучним інтелектом можуть швидше обробляти інформацію та використовувати короточасні недоліки, а також часто мають кращий рівень управління ризиками (наприклад, закривають позиції на ранніх стадіях виникнення проблем). Згідно зі звітом Deloitte, у 2024 році хедж-фонди, які інтегрували ШІ для управління ризиками, мали на 15% менші просадки під час стресових ситуацій на ринку. Здатність ШІ швидко аналізувати історичні патерни та проводити стрес-тестування портфелів у режимі реального часу допомогла менеджерам уникнути глибших втрат. Існують також специфічні ніші, в яких ШІ досягає успіху - наприклад, згідно зі звітом FT, фонди, керовані ШІ на ринках, що розвиваються, у 2024 році досягли на ~25% вищої прибутковості, ніж аналоги, завдяки швидкій адаптації до змін режиму. Навіть на криптовалютних ринках, які, як відомо, відрізняються волатильністю, криптохедж-фонди, що використовують алгоритми ШІ, забезпечили середню

рентабельність інвестицій на рівні 18% у 2024 році, перевершивши багато дискреційних криптовалютних фондів.

Інший аспект - продуктивність і ефективність. ШІ може впоратися з багатьма завданнями, які раніше вимагали великих команд. В одному з опитувань було відзначено, що трейдери, які використовують алгоритмічну (з допомогою ШІ) торгівлю, були приблизно на 10% продуктивнішими - вони могли управляти більшою кількістю угод або активів на людину - ніж ті, хто використовував ручні методи. Це підвищення продуктивності відбувається завдяки тому, що ШІ автоматизує дослідження (сканування новин, запуск зворотних тестів) і навіть виконання, звільняючи трейдерів-людей, щоб вони могли зосередитися на стратегії. ШІ також має тенденцію до зменшення людських помилок (відсутність втоми і емоцій), що може означати менше дорогих помилок. Тим не менш, не всі фонди зі штучним інтелектом постійно показують кращі результати - ринкові умови відіграють певну роль, і деякі ранні фонди зі штучним інтелектом мали посередні результати. Але в цілому тенденція досить обнадійлива, і капітал вливається в AI-стратегії.

Індекс хедж-фондів зі штучним інтелектом EurekaHedge (відстежує фонди зі штучним інтелектом) за останні роки випередив багато традиційних індексів хедж-фондів.

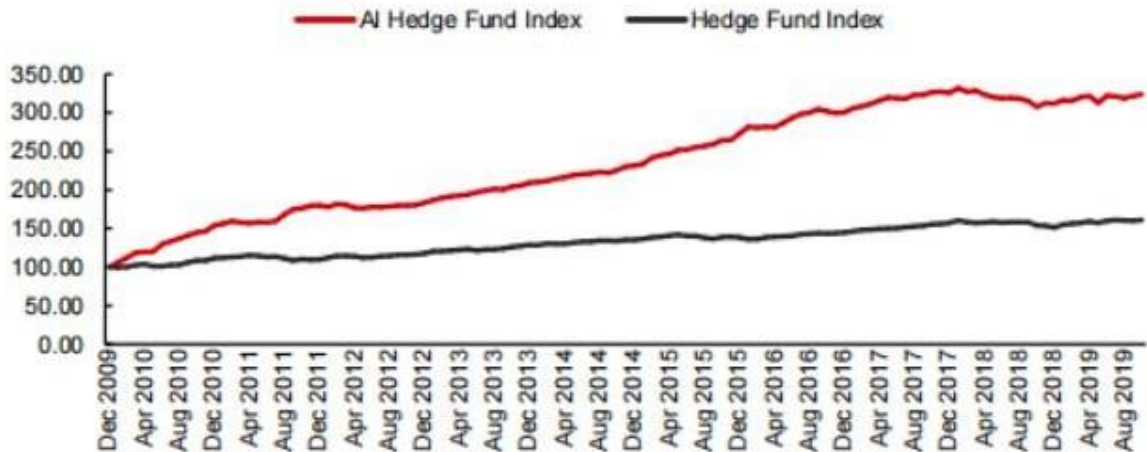


Рис. 3. Індекс ефективності хедж-фондів штучного інтелекту EurekaHedge. Взятий з публікації «Financial Relations between Hong Kong and the Mainland»

Великі компанії з управління активами також все частіше вкладають кошти в стратегії на основі ШІ, частково через побоювання втратити нове джерело альфа-інвестицій. Темпи зростання ринку: Ринок ШІ у фінансовій торгівлі зростає двозначними темпами і, за прогнозами, прискориться. Згідно з маркетинговими дослідженнями, світовий ринок торгових платформ зі штучним інтелектом (включаючи програмне забезпечення та послуги) оцінювався приблизно в \$11-12 млрд у 2024 році і, як очікується, досягне \$13,45 млрд у 2025 році. До 2030 року, за прогнозами, він становитиме близько \$33,45 млрд, що означає стабільні 20% середньорічні темпи зростання (CAGR) з 2025 по 2030 рік. Це зростання зумовлене все більшим впровадженням як в установах, так і в роздрібній торгівлі, постійним вдосконаленням технологій штучного інтелекту, а також зростанням складності ринків, що фактично вимагає допомоги штучного інтелекту. На наступному графіку показана прогнозована траєкторія зростання:

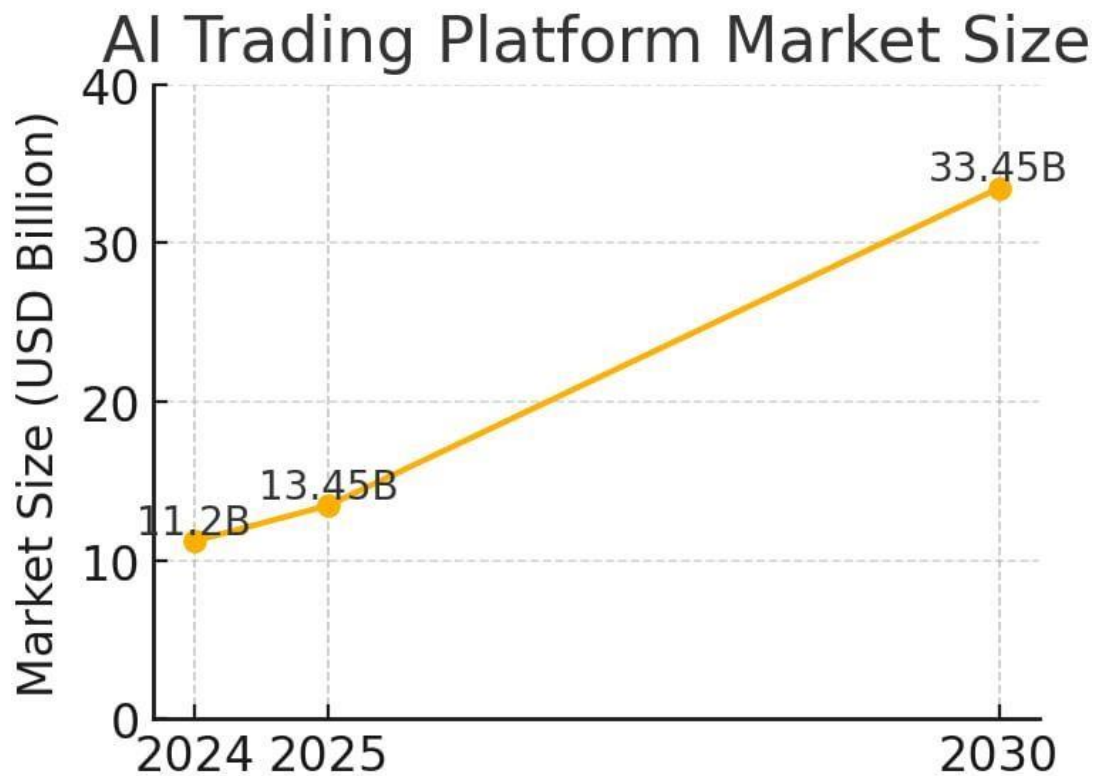


Рис. 4. Розмір і зростання світового ринку торгових платформ ШІ. Прогнозується, що ринок зросте з ~\$13,5 млрд у 2025 році до ~\$33,5 млрд до 2030 року ( $\approx 20\%$  CAGR), що відображає зростаючий попит на торгові рішення, керовані ШІ.

Якщо розбити цей показник за регіонами, то Північна Америка наразі домінує в інвестиціях у торгівлю штучним інтелектом, на неї припадає близько 37% доходу ринку у 2024 році.

Враховуючи концентрацію провідних фінансових установ і технологічних компаній у США та Канаді, а також розвинене регуляторне середовище, що заохочує інновації- Північна Америка домінує. Європа також відіграє важливу роль - великі фінансові центри, такі як Лондон, Франкфурт, Цюрих, активно впроваджують ШІ (часто в торгівлі, а також у сфері комплаєнсу). Але найшвидше зростання очікується в Азіатсько-Тихоокеанському регіоні, де такі країни, як Китай, Індія та Сінгапур, інвестують значні кошти у фінтех і ШІ для своїх ринків капіталу.

Прогнозується, що в Азіатсько-Тихоокеанському регіоні сектор торгівлі на основі ШІ зростатиме найвищими темпами середньорічного приросту серед регіонів. Наприклад, великі брокерські компанії та фонди Китаю швидко наймають експертів зі штучного інтелекту, а регулятори загалом підтримують розвиток фінтех-технологій. В Японії та Гонконгу традиційні фірми також переорієнтовуються на ШІ, щоб залишатися конкурентоспроможними. На ринках, що розвиваються, ШІ може випередити деякі застарілі системи, тому його впровадження може швидко прискоритися. [29] [31] [32]

### 3 Моделювання інвестиційного ризику компаній AI-сегменту

Метою технічного моделювання є оцінка інвестиційного ризику компаній, що працюють у сфері штучного інтелекту, з використанням кількісних та якісних показників, а також побудова моделі, яка дозволяє прогнозувати рівень ризику залежно від ринкових, фінансових та технологічних факторів. Зробити класифікацію компаній за їх рівнем ризику (низький / середній / високий), а також спробувати спрогнозувати ймовірність втрати прибутковості за регресійною ознакою.

Обрахунки виконуються на мові програмування **Python** в хмарній платформі **Google Colab**.

Для базової роботи з даними були використані бібліотеки **Pandas** та **NumPy**.

Для візуалізацій були добавлені бібліотеки **matplotlib.pyplot** для створення графіків та **seaborn** в якості надбудови над **matplotlib** заради приємнішого стіля теплової карти.

В якості моделей машинного навчання були використані **RandomForestClassifier** – модель для класифікації на основі ансамблю дерев рішень, яка працює за принципом голосування. **GradientBoostingClassifier** –

модель класифікації, яка навчається поетапно на помилках попередніх дерев.

**LogisticRegression** – лінійна модель для класифікації.

Для переробки даних були використані:

1. **StandardScaler** – задля стандартизації ознак.
2. **SimpleImputer** – для заповнення пропущених значень.

Для розбиття на вибірки був використаний **train\_test\_split** – він розбиває дані на тренувальну та тестову вибірки.

В якості підбору гіперпараметрів та валідації даних був використан **GridSearchCV** – автоматичний підбір кращих гіперпараметрів шляхом перебору всіх комбанацій з валідацією.

Для балансування класів був використаний **RandomOverSampler** – метод, який додає приклади з менш представлених класів, щоб збалансувати навчання. **Counter** – дозволяє швидко підраховувати кількість елементів у списку, був використаний для перевірки балансу класів.

В якості оцінки якості класифікації **classification\_report** сформував звіт з метриками Precision, Recall, F1-score, Support для кожного класу.

### 3.1 Набір даних та етапи підготовки для побудови моделі

Спершу був обраний набір даних. [1]

Табл. 1. Параметри таблиці з даними

Назва колонки	Опис
Company	Назва компанії
Major index membership	Належність до основного індексу
Market capitalization	Ринкова капіталізація

Income (ttm)	Чистий дохід за останні 12 місяців
Revenue (ttm)	Виручка за останні 12 місяців
Book value per share (mrq)	Балансова вартість на акцію за останній звітний квартал
Cash per share (mrq)	Грошові кошти на акцію за останній звітний квартал
Dividend (annual)	Дивіденди за рік
Dividend yield (annual)	Дивідендна дохідність за рік
Full time employees	Кількість штатних співробітників
Stock has options trading on a market exchange	Чи торгуються опціони на акції на біржі
Stock available to sell short	Можливість короткого продажу акцій
Analysts' mean recommendation (1=Buy 5=Sell)	Середня рекомендація аналітиків (1 – купувати, 5 – продавати)
Price-to-Earnings (ttm)	Коефіцієнт «ціна/прибуток» за останні 12 місяців
Forward Price-to-Earnings (next fiscal year)	Прогнозований коефіцієнт «ціна/прибуток» на наступний фінансовий рік
Price-to-Earnings-to-Growth	Коефіцієнт «ціна/прибуток/зростання» (PEG)
Price-to-Sales (ttm)	Коефіцієнт «ціна/виручка» за останні 12 місяців
Price-to-Book (mrq)	Коефіцієнт «ціна/балансова вартість» за останній звітний квартал
Price to cash per share (mrq)	Коефіцієнт «ціна/грошові кошти на акцію» за останній звітний квартал

Price to Free Cash Flow (ttm)	Коефіцієнт «ціна/вільний грошовий потік» за останні 12 місяців
Quick Ratio (mrq)	Швидкий коефіцієнт ліквідності за останній звітний квартал
Current Ratio (mrq)	Поточний коефіцієнт ліквідності за останній звітний квартал
Total Debt to Equity (mrq)	Співвідношення загального боргу до власного капіталу за останній звітний квартал
Long Term Debt to Equity (mrq)	Співвідношення довгострокового боргу до власного капіталу за останній звітний квартал
Distance from 20-Day Simple Moving Average	Відхилення поточної ціни від 20-денної простої ковзної середньої
Diluted EPS (ttm)	Розведений прибуток на акцію за останні 12 місяців
EPS estimate for next year	Прогноз прибутку на акцію на наступний рік
EPS estimate for next quarter	Прогноз прибутку на акцію на наступний квартал
EPS growth this year	Зростання прибутку на акцію за поточний рік
EPS growth next year	Зростання прибутку на акцію за наступний рік
Long term annual growth estimate (5 years)	Оцінка довгострокового щорічного зростання (5 років)
Annual EPS growth past 5 years	Середнє щорічне зростання прибутку на акцію за останні 5 років

Annual sales growth past 5 years	Середнє щорічне зростання виручки за останні 5 років
Quarterly revenue growth (YoY)	Зростання виручки за квартал у порівнянні з аналогічним періодом попереднього року
Quarterly earnings growth (YoY)	Зростання прибутку за квартал у порівнянні з аналогічним періодом попереднього року
Earnings date BMO = Before Market Open AMC = After Market Close	Дата оголошення результатів (BMO – до відкриття ринку, AMC – після закриття ринку)
Distance from 50-Day Simple Moving Average	Відхилення поточної ціни від 50-денної простої ковзної середньої
Insider ownership	Частка власності інсайдерів
Insider transactions (6-Month change in Insider Ownership)	Зміна частки власності інсайдерів за останні 6 місяців
Institutional ownership	Частка власності інституційних інвесторів
Institutional transactions (3-Month change in Institutional Ownership)	Зміна частки власності інституційних інвесторів за останні 3 місяці
Return on Assets (ttm)	Рентабельність активів за останні 12 місяців
Return on Equity (ttm)	Рентабельність власного капіталу за останні 12 місяців
Return on Investment (ttm)	Рентабельність інвестицій за останні 12 місяців
Gross Margin (ttm)	Валова маржа за останні 12 місяців

Operating Margin (ttm)	Операційна маржа за останні 12 місяців
Net Profit Margin (ttm)	Чиста маржа прибутку за останні 12 місяців
Dividend Payout Ratio (ttm)	Співвідношення виплат дивідендів до прибутку за останні 12 місяців
Distance from 200-Day Simple Moving Average	Відхилення поточної ціни від 200-денної простої ковзної середньої
Shares outstanding	Загальна кількість випущених акцій
Shares float	Кількість вільно обігаючих акцій
Short interest share / ratio	Частка коротких позицій / коефіцієнт коротких позицій
Short interest	Обсяг коротких позицій
Analysts' mean target price	Середня цільова ціна аналітиків
52-Week trading range	Діапазон торгів за останні 52 тижні
Distance from 52-Week High	Відхилення поточної ціни від 52-тижневого максимуму
Distance from 52-Week Low	Відхилення поточної ціни від 52-тижневого мінімуму
Relative Strength Index	Індекс відносної сили (RSI)
Relative volume	Відносний обсяг торгів
Average volume (3 month)	Середній обсяг торгів за 3 місяці
Volume	Поточний обсяг торгів
Performance (Week)	Динаміка ціни за останній тиждень
Performance (Month)	Динаміка ціни за останній місяць
Performance (Quarter)	Динаміка ціни за останній квартал
Performance (Half Year)	Динаміка ціни за останні півроку
Performance (Year)	Динаміка ціни за останній рік

Performance (Year To Date)	Динаміка ціни з початку поточного року
Beta	Бета-коефіцієнт волатильності
Average True Range (14)	Середній справжній діапазон за 14 днів
Volatility (Week, Month)	Волатильність за тиждень та місяць
Previous close	Ціна закриття попереднього торгового дня

Цей набір даних містить описові та фінансові характеристики 496 компаній із S&P 500, включно з:

1. Фінансові метрики (Ринкова капіталізація, Прибуток (ttm), Виручка, P/E, P/S, Ціна до балансу, EPS, Бета, Співвідношення боргу до власного капіталу та ін.)
2. Оцінка аналітиків: (Середня рекомендація аналітиків, Цільова ціна, Оцінка прибутку на акцію)
3. Динаміка та ризик: (Волатильність, Ефективність за різними періодами, Бета, Індекс відносної сили, Короткі відсотки)
4. Оцінки ризику / ліквідності: (Коефіцієнт швидкої ліквідності, коефіцієнт поточної ліквідності, вільний грошовий потік, ціна до грошових коштів, борг до власного капіталу)

Аналіз набору даних

```
df = pd.read_csv('content/sp500_companies_description.csv')
```

Company	Major Index Membership	Market Capitalization	Income (ttm)	Revenue (ttm)	Book value per share (bvps)	Cash per share (cps)	Dividend (annual)	Dividend yield (annual)	Full time employees	Performance (Quarter)	Performance (Half Year)	Performance (Year)	Performance (Year To Date)	Beta	Average True Range (14)	Volatility (Week, Month)	Previous close stock price	Current stock price	Performance (today)		
0	A	S&P 500	41.05B	1.32B	6.93B	18.95	0.90	0.64%	18300	-6.59%	7.56%	0.32%	-5.53%	1.04	3.27	2.37%	2.09%	140.90	141.37	0.33%	
1	AAL	S&P 500	10.65B	127.00M	48.97B	-8.91	13.97	-	129700	14.14%	17.76%	21.07%	28.59%	1.56	0.48	2.65%	2.73%	16.59	16.36	-1.41%	
2	AAP	S&P 500	7.60B	501.90M	11.15B	45.06	4.52	6.00	40000	9.20%	-25.89%	-34.12%	-10.13%	1.15	3.70	2.00%	2.39%	132.21	132.14	-0.05%	
3	AAPL	DMA, NDJ, S&P 500	2336.71B	95.17B	387.54B	3.57	3.38	0.92	0.60%	164000	7.70%	-1.31%	-2.24%	18.46%	1.29	3.30	1.67%	1.96%	152.87	153.91	0.68%
4	ABBV	S&P 500	261.85B	11.78B	58.95B	9.74	5.27	5.92	3.96%	50000	-8.84%	7.69%	1.43%	-7.65%	0.55	2.80	1.91%	1.90%	149.60	149.25	-0.23%
491	YUM	S&P 500	35.24B	1.32B	6.84B	-31.36	1.32	2.42	1.91%	23000	-2.25%	10.72%	10.79%	-0.31%	1.00	1.93	1.60%	1.55%	126.78	127.68	0.71%
492	ZBH	S&P 500	24.95B	290.20M	6.94B	57.27	1.04	0.96	0.79%	10000	1.74%	13.74%	6.70%	-3.14%	1.00	2.23	1.90%	1.68%	121.91	123.58	1.30%
493	ZBRA	S&P 500	15.61B	403.00M	5.78B	52.00	2.07	-	10500	10.54%	2.80%	-20.05%	18.51%	1.07	0.10	2.27%	2.91%	307.07	303.87	-1.04%	
494	ZION	S&P 500	6.81B	889.00M	2.71B	29.94	41.26	1.64	3.51%	9989	-9.10%	-24.01%	-31.43%	-12.90%	1.17	1.57	3.22%	2.32%	46.68	42.82	-8.27%
495	ZTS	S&P 500	77.24B	2.11B	8.08B	9.46	7.81	1.50	0.89%	13000	11.04%	5.95%	-6.31%	15.96%	0.77	3.04	1.76%	2.20%	168.64	169.94	0.77%

496 rows • 73 columns

Рис. 5. Вивантажений набір даних в Google Colab.

Логіка виключення непотрібних колонок: Унікальні для кожної компанії:

- Company — назва компанії → майже завжди унікальна
- Earnings date — дата звітності, важко узагальнити без складного препроцесингу
- Технічні, нестабільні або короткострокові змінні:
- Current stock price, Previous close, Performance (today)
- Distance from X-Day Moving Average
- Volatility (Week, Month) — можуть вводити шум без глибокого обліку періоду
- Дублікати за змістом або високочорельовані:
- Shares float, Shares outstanding — часто дублюють ринкову капіталізацію
- Average volume, Volume — для довгострокових оцінок не критично
- Price to cash per share, Cash per share, Book value — часто дублюють ROI або капіталізацію
- Складні для моделі без попередньої обробки:
- 52-Week trading range, Performance (Week), Analysts' target price — вимагають розбору в цифри/дельти
- Stock has options, Stock available to sell short — майже усі мають "Yes"

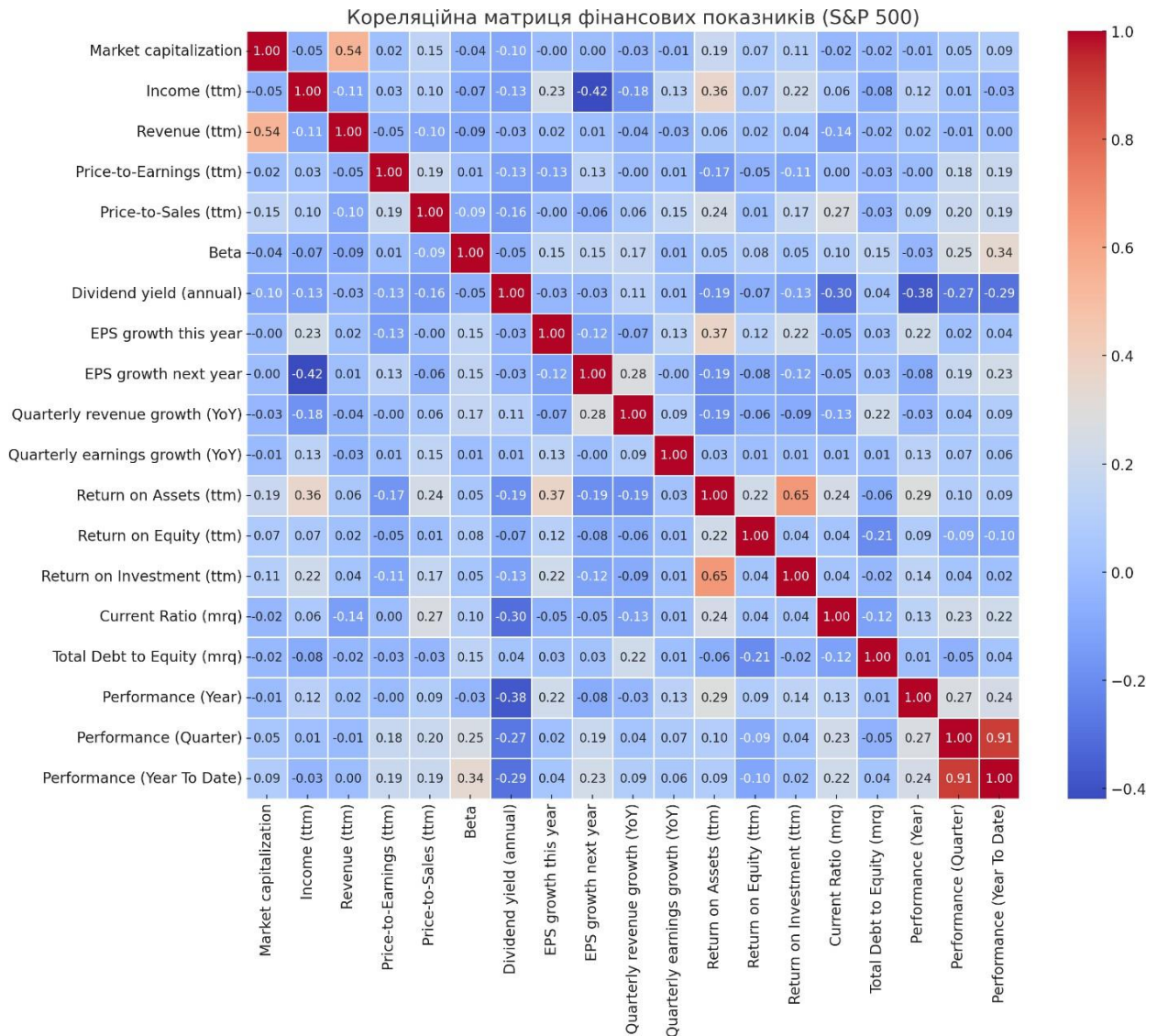


Рис. 6. Теплова карта, яка вказує кореляцію даних між собою. Графік з Google Colab.

Теплова карта показує, які змінні мають сильний зв'язок одна з одною:

Червоні клітинки — позитивна кореляція (чим ближче до 1, тим сильніше).

Сині клітинки — негативна кореляція (чим ближче до -1, тим сильніше обернена залежність).

Значення близько до 0 — відсутність суттєвого зв'язку.

## Дослідження важливості ознак та створення цільових ознак для регресії та класифікації

Спершу виконується важливість ознак та розподілення компаній на 3 категорії в залежності їх ризикованості. Методом ручового розбиття дані столбця Performance (Year) очищаються від зайвих символів «%» та перетворюються у числовий тип. Після цього завдяки `classige_risk()` розбиваю дані:

1. Якщо дохідність менше ніж -10% - класифікую як «Високий ризик»
2. Якщо дохідність на проміжку від -10% до 10% - надається «Середній ризик»
3. Якщо дохідність більш ніж 10% - «Низький ризик»

Отримані результати:

1. Середній ризик – 199 компаній.
2. Високий ризик – 166 компаній.
3. Низький ризик – 131 компанія.

Через велику кількість ознак в датасеті, було прийнято рішення провести попередній аналіз інформативності ознак, щоб зрозуміти, які саме характеристики компаній мають найбільший вплив на інвестиційний ризик та річну дохідність. Для цього я використав алгоритм Random Forest для класифікації, а також для регресії.

RandomForestClassifier для передбачення класу інвестиційного ризику (`risk_class`).

RandomForestRegressor для прогнозування безперервної метрики Performance (Year).

Моделі спершу були використані без попередньої оптимізації гіперпараметрів, а саме з  $n\_estimator=100$ ,  $random\_state=42$ ). Також дані були масштабовані за допомогою StandartScaler, щоб їх значення набували від 0 до 1, щоб уникнути домінування ознак з великим діапазоном значень.

Після тренування моделей була використана вбудована властивість `.feature_importances_`, яка показує наскільки кожна ознака важлива на прийняття рішення у дереві, і також дозволяє кількісно ранжувати ознаки за важливістю.

Для обох моделей була побудована гістограма важливості ознак, що дозволяє зрозуміти, які з параметрів найбільш інформативні, а які малозначущі.

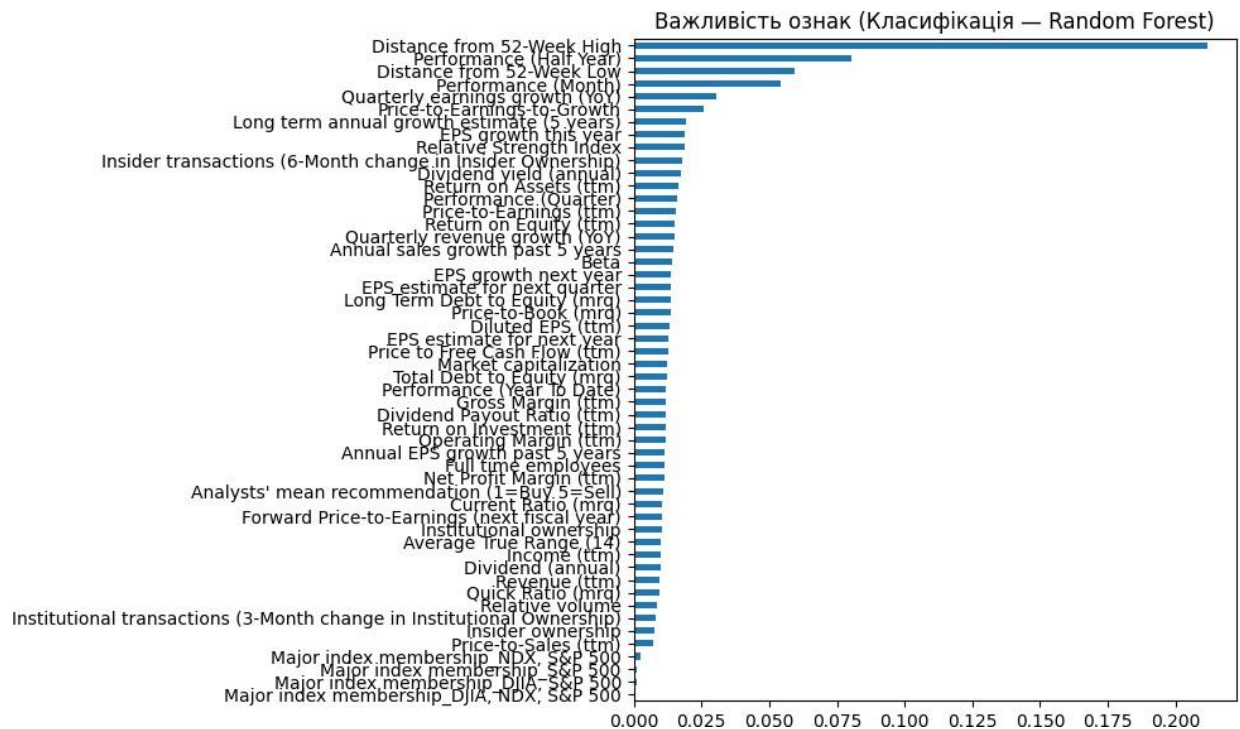


Рис. 7. Гістограма важливості ознак класифікації методом Random Forest.

З даної гістограми видно, що «Distance from 52-week High» являється найвпливовішим фактором і він вказує на те, наскільки ціна акції відвійшла від свого річного максимуму. «Distance from 52-week Low» - це близькість до мінімуму. «Performance (Year)» - фактична річна дохідність, яка прямо пов'язана з ризиком і є основою для формування цільової змінної. Ознаки які пов'язані з належністю до індексів «S&P 500», «DJIA», «Nasdaq» показали мініміальну важливість у моделі.

Отже, завдяки цій перевірці я відібрав найбільш інформативні фічі та зменшив розмір датасету, для коректної роботи.

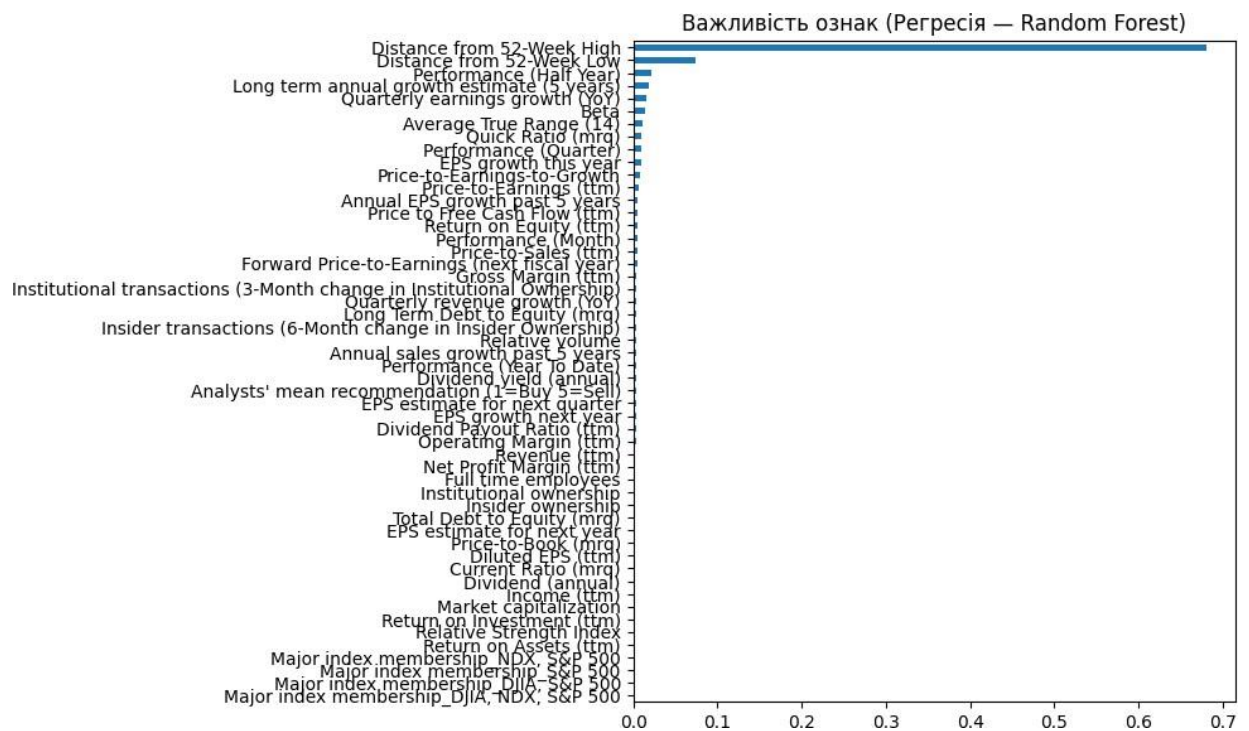


Рис. 8. Гістограма важливості ознак регресії.

У ході побудови регресійної моделі Random Forest з метою прогнозування річної дохідності було проведено аналіз важливості ознак. Було виявлено, що ознака «Distance from 52-week High» мала надмірно вискоюй вплив і значно перевищувала всі інші характеристики. Результат який вийшов призводить до зниження загальної стійкості і узагальнення моделі, також робить її дуже вразливою до шумів у конкретному показнику, і ускладнює

інтерпритацію та оцінку ризику. Через це було прийнято рішення не використовувати регресійний підхід у подальшому дослідженні.

Табл. 2. Результати тренувальної моделі

	Precision	Recall	F1-score	Support
Високий	1	1	1	134
Низький	1	1	1	97
Середній	1	1	1	165
Точність			1	396
Macro avg	1	1	1	396
Weighted avg	1	1	1	396

Результати даної таблиці показують на явне перенавчання тренувальної моделі – усі значення дорівнюють 1. Точність моделі також показала 1, що майже ніколи не буває у реальних задачах.

Метрика Precision відображає частку об'єктів, яку відносить до певного класу. В даному випадку 134/134, якщо говорити про «Високий».

Метрика Recall відображає яку частку об'єктів правильно знайшло.

F1-score – це середнє між Precision та Recall. Баланс між точністю і повнотою.

Support – кількість об'єктів у кожному класі.

Табл. 3. Результати тестової моделі

	Precision	Recall	F1-score	Support
Високий	0,85	0,91	0,88	32
Низький	0,85	0,68	0,75	34
Середній	0,64	0,74	0,68	34

Точність			0,77	100
Macro avg	0,78	0,77	0,77	100
Weighted avg	0,78	0,77	0,77	100

Спробуємо урівняти кількість прикладів для кожного з класів для тестової вибірки. Але спочатку додамо ще розділення і на валідаційну вибірку для майбутнього підбору гіперпараметрів моделей

Також є пробелма з наявністю пропусків у деяких записах, але через обмеженість вибірки вирішено заповнити ці значення середнім `SimpleImputer(strategy='mean')`.

Після цього роблю розбивку по типу тренувальна і тестова вибірки.

```
X_train_full, X_test, y_train_full, y_test = train_test_split(
    X_class_scaled, y_class, test_size=0.2, random_state=42,
    stratify=y_class
)
```

`X_class_scaled` – це матриця ознак, яка була попередньо масштабована методом «`StandartScaler`»

`Y_class` – цільова змінна, яка вказує на класи ризику.

`test_size=0.2` – показує, що 20% даних буде використано як тестову вибірку.

`random_state=42` – фіксує випадковість відтворення результатів.

`stratify=y_class` – відповідає, щоб розподіл класів у тренувальній і тестовій виборках буде пропорційно однаковим.

Розділення даних на тренувальну і валідаційну вибірку:

Вхідні дані `X_class_scaled` та цільова змінна `y_class` були розділені на 80% для навчання (`train`) і 20% для валідації (`val`).

Було використано параметр `stratify=y_class`, що гарантує однаковий розподіл класів у тренувальній і валідаційній вибірках (стратифікація).

Подальший поділ тренувальної вибірки:

Отримана тренувальна вибірка `X_train_full` додатково розділена на 80% — для остаточного навчання (`X_train`) і 20% — для валідації (`X_val`).

Знову ж таки, використовується стратифікація для збереження балансу класів у `y_train` і `y_val`.

Розміри отриманих підмножин:

Тренувальні: 316 зразків

Валідовані: 80 зразків

Тестові: 100 зразків

Кількість ознак (фіч) у кожному наборі — 39

Під час попереднього аналізу та обробки даних були виявлено дві ключові проблеми - наявність пропущених значень, а також незбалансованість класів.

У деяких ознаках були присутні пропуски і тому, щоб не було проблем з навчанням і модель не згенерувало з помилкою був використаний `OverSampling` – метод, який синтетично розмножує записи меншого класу, щоб їх кількість дорівнювала найбільшому.

Спочатку відбувається перевірка початкового розподілу класів у `y_train` і бачимо, що класи представлені нерівномірно. «Середній» 127 прикладів, «Високий» 105 прикладів, «Низький» має 84 приклади.

Застосовуючи `RandomOverSampler` (`oversampling`) робимо так, щоб усі класи мали однакову кількість зразків. `Oversampling` застосовується лише до тренувального набору, щоб уникнути витоку інформації валідації чи тесту.

Після цього бачимо, що три класи мають по 127 прикладів — тобто тренувальні дані збалансовані.

Під час підбору гіперпараметрів трьох моделей класифікації (`Random Forest`, `Gradient Boosting`, `Logistic Regression`) за допомогою `GridSearchCV` було використано `F1_max` як цільову для оцінювання.

Функція була обрана за цільову через те, що у задачі класифікації ми маємо три класи, причому їх розподіл нерівномірний (клас «низький» має значно менше прикладів, ніж «середній»).

Метрика `F1-score` є середнім гармонічним між `precision` (точністю) та `recall` (повнотою), що дозволяє знайти баланс між хибно позитивними та хибно негативними відповідями. `Max` варіант цієї метрики не враховує дисбаланс класів (усереднює `F1-score` для кожного класу без зважування), що дозволяє оцінити ефективність моделі в кожному класі окремо, навіть якщо деякі класи зустрічаються рідше.

Тому саме вона є найбільш адекватною для багатокласової класифікації (у нас їх 3), коли всі класи однаково важливі і при наявності дисбалансу в цільовій змінній.

### 3.2 Random Forest

**Random Forest** (випадковий ліс) - це ансамблевий алгоритм машинного навчання, який базується на великій кількості дерев рішень, що

працюють паралельно, а не послідовно, як у бустингу. Random Forest працює по принципу:

- Бутстрепінг (випадкове підвибіркування):
  - Для кожного дерева вибирається випадкова вибірка з тренувальних даних (із поверненням — тобто один запис може потрапити кілька разів).
- Побудова дерев:
  - Кожне дерево будується незалежно від інших.
  - На кожному розгалуженні дерева випадково обирається підмножина ознак (наприклад,  $\sqrt{\text{загальної кількості}}$ ), щоб зробити розділення.
  - Це створює різноманітність між деревами.
- Голосування:
  - Після того як усі дерева зробили свої передбачення:
    - Для класифікації — використовується голосування більшості.
    - Для регресії — береться середнє значення прогнозів дерев

### **Серед загальних переваг Random Forest можна назвати:**

- Високу точність навіть без складної настройки параметрів.
- Добре працює з великими наборами даних і багатьма змінними.
- Стійкий до перенавчання завдяки усередненню багатьох дерев.
- Вбудована оцінка важливості ознак.

### **Порівняння з градієнтним бустингом:**

Табл. 4 Порівняння Random Forest, Gradient Boosting та Logistic Regression

Характеристика	Random Forest	Gradient Boosting	Logistic Regression
Побудова моделей	Паралельна	Послідовна	Одна модель (лінійна комбінація ознак)
Підхід до помилок	Голосування багатьох дерев	Навчання на помилках попередніх	Мінімізація логістичної функції втрат
Схильність до overfitting	Низька	Може бути високою	Відносно низька, особливо з регуляризацією
Швидкість навчання	Швидше	Повільніше	Дуже швидке навчання

Отже Random Forest — це надійний, стійкий до перенавчання алгоритм, який підходить для широкого кола задач, особливо коли важлива стабільність і швидкий результат без складної настройки.

Табл. 5. Результати тренувальної вибірки методом Random Forest

	precision	recall	f1-score	support
високий	1.00	1.00	1.00	127
низький	1.00	1.00	1.00	127
середній	1.00	1.00	1.00	127
accuracy			1.00	381

macro avg	1.00	1.00	1.00	381
weighted avg	1.00	1.00	1.00	381

Табл. 6. Результати валідаційної вибірки методом Random Forest

	precision	recall	f1-score	support
високий	0,87	0,96	0,91	27
низький	0,72	0,62	0,67	21
середній	0,75	0,75	0,75	32
accuracy			0,79	80
macro avg	0,78	0,78	0,78	80
weighted avg	0,78	0,79	0,78	80

Табл. 7. Результати тестової вибірки методом Random Forest

	precision	recall	f1-score	support
високий	0,79	0,91	0,85	34
низький	0,82	0,69	0,75	26
середній	0,74	0,72	0,73	40
accuracy			0,78	100
macro avg	0,79	0,78	0,78	100
weighted avg	0,78	0,78	0,78	100

В таблицях подано результати моделі **Random Forest**, яка була навчена на збалансованій вибірці після oversampling, з використанням оптимальних

гіперпараметрів: 'max\_depth': None, 'max\_features': 'sqrt', 'min\_samples\_leaf': 1, 'min\_samples\_split': 2, 'n\_estimators': 100.

На тренувальній вибірці модель демонструє ідеальну продуктивність: усі три класи — "високий", "низький" і "середній" — мають precision, recall і F1-метрику рівні 1.00, що означає повне запам'ятовування навчальних даних. Це свідчить про явне перенавчання (overfitting), тобто модель ідеально працює на вже бачених прикладах, але, ймовірно, втратить точність на нових даних.

На валідаційній вибірці спостерігається помітне зниження продуктивності. Найвищий результат досягається для класу "високий" (F1 = 0.91), однак класи "низький" і "середній" мають нижчі F1-метрики — 0.67 і 0.75 відповідно. Загальна точність моделі становить 0.78, що є середнім показником, і вказує на втрату узагальнюючої здатності, особливо щодо менш виражених класів.

На тестовій вибірці, яка є незалежною, Random Forest показує аналогічну ситуацію: загальна точність — 0.78, F1-метрики варіюються від 0.73 до 0.85. Клас "високий" розпізнається найкраще (F1 = 0.85), у той час як клас "середній" має найнижчий показник (F1 = 0.73). Це ще раз підкреслює, що модель схильна більше «любити» один домінуючий клас, а інші — плутає частіше.

Загалом, модель Random Forest показала відмінне навчання на тренувальних даних, але її здатність до узагальнення на нові дані помітно гірша, особливо для менш представлених або складніших для розпізнавання класів. Це типовий приклад моделі, яка була надто гнучкою на етапі навчання, і хоч результати на тесті прийнятні, вони поступаються простішим або краще збалансованим підходам у класифікації.

## **Gradient Boosting**

**Градiєнтний бустинг** — це потужний ансамблевий метод машинного навчання, який поєднує багато слабких моделей (зазвичай рiшень дерев) у одну сильну. Працює по принципу:

1. Навчається перше дерево, яке робить початкові передбачення.
2. Обчислюється помилка (residuals) — рiзниця між передбаченням і фактичним значенням.
3. Наступне дерево вчиться на цих помилках, щоб їх виправити.
4. Моделі додаються поступово, кожна трохи покращує загальний результат.
5. Використовується learning rate, який контролює, наскільки сильно кожна нова модель впливає на фiнальний прогноз.

Табл.8. Результати тренувальної вибірки методом Gradient Boosting

	precision	recall	f1-score	support
високий	1.00	1.00	1.00	127
низький	0,97	0,98	0,97	127
середній	0,98	0,97	0,97	127
accuracy			0,98	381
macro avg	0,98	0,98	0,98	381
weighted avg	0,98	0,98	0,98	381

Табл.9. Результати валідаційної вибірки методом Gradient Boosting

	precision	recall	f1-score	support
високий	0,87	0,96	0,91	27
низький	0,70	0,76	0,73	21

середній	0,81	0,69	0,75	32
accuracy			0,80	80
macro avg	0,79	0,80	0,80	80
weighted avg	0,80	0,80	0,80	80

Табл.10. Результати тестової вибірки методом Gradient Boosting

	precision	recall	f1-score	support
високий	0,81	0,88	0,85	34
низький	0,79	0,73	0,76	26
середній	0,74	0,72	0,73	40
accuracy			0,78	100
macro avg	0,78	0,78	0,78	100
weighted avg	0,78	0,78	0,78	100

В таблицях представлено результати роботи моделі **Gradient Boosting** після навчання, валідації та тестування. Для навчання було використано збалансовані (через oversampling) дані, а оптимальні гіперпараметри підібрано наступним чином: `learning_rate = 0.01`, `max_depth = 3`, `n_estimators = 300`, `subsample = 0.6` та інші базові параметри регуляризації.

На тренувальній вибірці модель демонструє майже ідеальні результати: всі три класи мають F1-метрику не нижче 0.97, а загальна точність становить 0.98. Це свідчить про високу здатність моделі навчатись, хоча й з потенційною ознакою часткового перенавчання, враховуючи досконалі метрики.

На валідаційній вибірці результати є збалансованими: загальна точність становить 0.80, а F1-метрики коливаються від 0.73 до 0.91. Клас

"високий" розпізнається найкраще ( $F1 = 0.91$ ), клас "низький" — помірно ( $F1 = 0.73$ ), а "середній" — дещо гірше ( $F1 = 0.75$ ). Такий розподіл вказує на те, що модель має хорошу здатність до узагальнення, але складнощі можуть виникати при класифікації менш яскраво виражених або схожих класів.

На тестовій вибірці, яка повністю незалежна, модель показала стабільність: усі ключові метрики для трьох класів знаходяться на рівні 0.73–0.85, а загальна точність — 0.78. Це підтверджує, що модель добре переносить навички, отримані в навчанні, на нові дані. Зокрема, клас “середній” має  $F1 = 0.73$ , що є стабільним і прийнятним результатом.

Загалом, модель **Gradient Boosting** продемонструвала сильні результати на тренувальних даних та стійке узагальнення на валідації і тесті. Вона впевнено класифікує основний клас (“високий”), але потребує уваги до оптимізації на менш виражених класах (“низький” і “середній”). Це робить її надзвичайно придатною до використання в задачах, де важлива точність і контроль над перенавчанням.

### Logistic regression

**Логістична регресія** – це статистичний метод, який використовується для прогнозування ймовірності настання події. Научастіше використовується у задачах класифікації, де результатом являється один із двох можливих варіантів.

Табл. 11. Результати тренувальної вибірки методом Logistic regression

	precision	recall	f1-score	support
високий	0,95	0,97	0,96	127
низький	0,86	0,91	0,88	127
середній	0,89	0,83	0,86	127
accuracy			0,90	381

macro avg	0,90	0,90	0,90	381
weighted avg	0,90	0,90	0,90	381

Табл.12. Результати валідаційної вибірки методом Logistic regression

	precision	recall	f1-score	support
високий	0,81	0,93	0,86	27
низький	0,75	0,86	0,80	21
середній	0,84	0,66	0,74	32
accuracy			0,80	80
macro avg	0,80	0,81	0,80	80
weighted avg	0,81	0,80	0,80	80

Табл.13. Результати тестової вибірки методом Logistic regression

	precision	recall	f1-score	support
високий	0,88	0,85	0,87	34
низький	0,75	0,81	0,78	26
середній	0,77	0,75	0,76	40
accuracy			0,80	381
macro avg	0,80	0,80	0,80	381
weighted avg	0,80	0,80	0,80	381

В таблицях надані результати класифікаційної моделі логістичної регресії, навченої на збалансованих (oversampled) даних для багатокласової задачі. Було підібрано оптимальні гіперпараметри: 'C': 1, 'max\_iter': 1000, 'penalty': 'l1', 'solver': 'saga', що дозволяє моделі використовувати L1-регуляризацію (спарсність ознак) при досить великій кількості ітерацій.

На тренувальній вибірці модель показала високу якість класифікації — точність, повнота та F1-метрика для кожного класу коливаються в межах 0.86–0.96, а загальна асигасу становить 0.90. Це свідчить про успішне навчання на збалансованих даних без явного перенавчання.

На валідаційній вибірці результати дещо гірші: асигасу — 0.80, що свідчить про нормальну здатність моделі до узагальнення. Клас “високий” розпізнається добре (F1 = 0.86), тоді як для “низького” і “середнього” класів точність і recall нижчі (F1 = 0.76 та 0.74 відповідно), що вказує на меншу здатність моделі чітко відокремлювати ці класи.

На тестовій вибірці (яка є остаточним незалежним набором) логістична регресія демонструє аналогічну до валідації поведінку: загальна асигасу — 0.80, причому найкраще розпізнається клас “високий” (F1 = 0.87), а гірше — “середній” (F1 = 0.76). Це вказує на стабільність моделі, однак помітна схильність до кращого розпізнавання одного класу на шкоду іншим.

Загалом, логістична регресія показує прийнятну якість, стабільну поведінку та хорошу узагальненість, проте поступається більш складним моделям (як Gradient Boosting), особливо в класифікації “середнього” та “низького” класів. Це типовий результат для лінійної моделі, яка працює краще при чітко розділених класах і менше схильна до перенавчання, але може мати обмеження в складних структурах даних.

### Порівняння результатів класифікації

Табл.14. Порівняння результатів

Модель	Тренувальна	Валідаційна	Тестова	Надмірне навчання
Random Forest	1,00	0,78	0,78	Так

Gradient Boosting	0,98	0,80	0,78	Так
Logistic regression	0,90	0,80	0,80	Менше

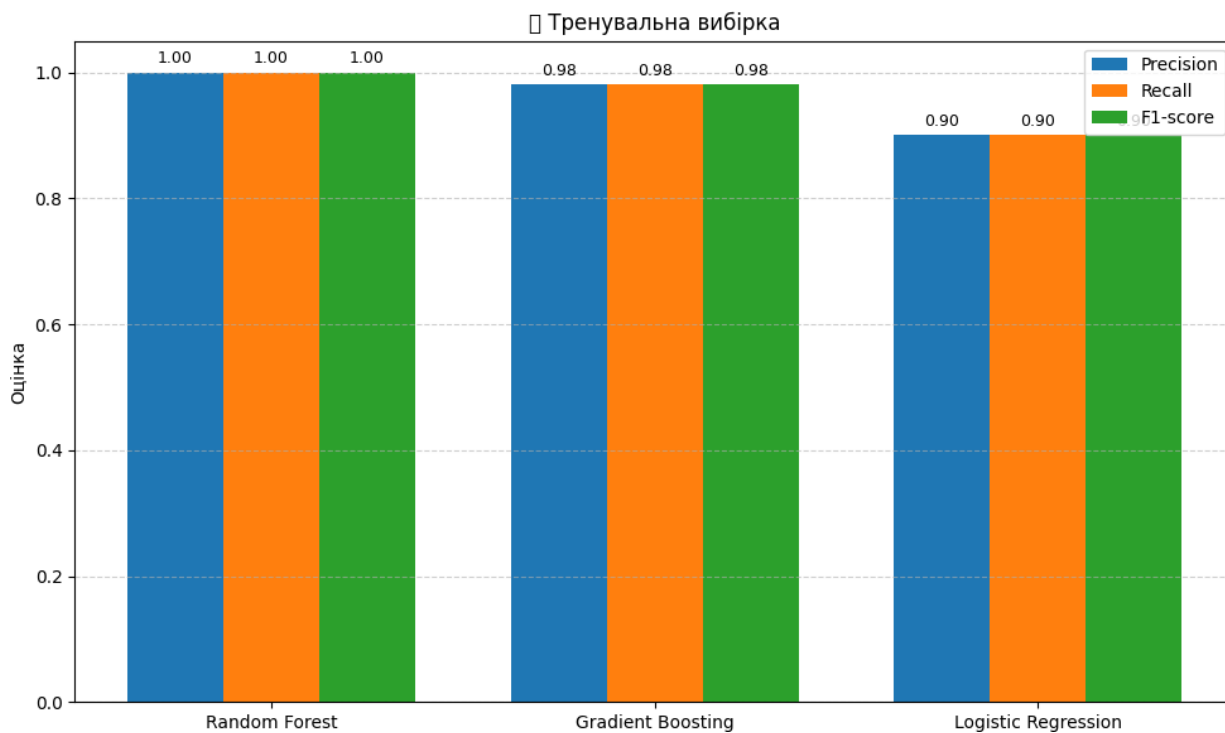


Рис. 9. Результати тренувальної вибірки.

На графіку зображено порівняння трьох моделей машинного навчання — Random Forest, Gradient Boosting та Logistic Regression — за трьома основними метриками (точність — *precision*, повнота — *recall*, та F1-метрика) на тренувальній вибірці.

Модель Random Forest демонструє ідеальні значення за всіма метриками — 1.00, що свідчить про повне запам'ятовування тренувальних даних. Це ознака потенційного перенавчання, коли модель занадто добре відображає навчальні приклади, але може мати труднощі з узагальненням на нові дані.

Gradient Boosting показує трохи нижчі, але все ще дуже високі результати — 0.98 за кожною з трьох метрик. Це свідчить про більш стабільну поведінку, де модель добре навчається, але не настільки «запам'ятовує» дані, як Random Forest. Такий результат може бути оптимальнішим з погляду узагальнення.

Нарешті, Logistic Regression має найнижчі значення серед моделей — 0.90 для всіх метрик. Це очікувано, оскільки логістична регресія є лінійною моделлю, яка менш потужна в складних нелінійних задачах, однак її простота робить її стійкою до перенавчання та добре інтерпретованою.

У підсумку, на тренувальних даних найвищу якість показує Random Forest, проте Gradient Boosting пропонує баланс між точністю і потенційною здатністю до узагальнення, а Logistic Regression демонструє стабільні, хоча й нижчі показники, властиві більш простій моделі.

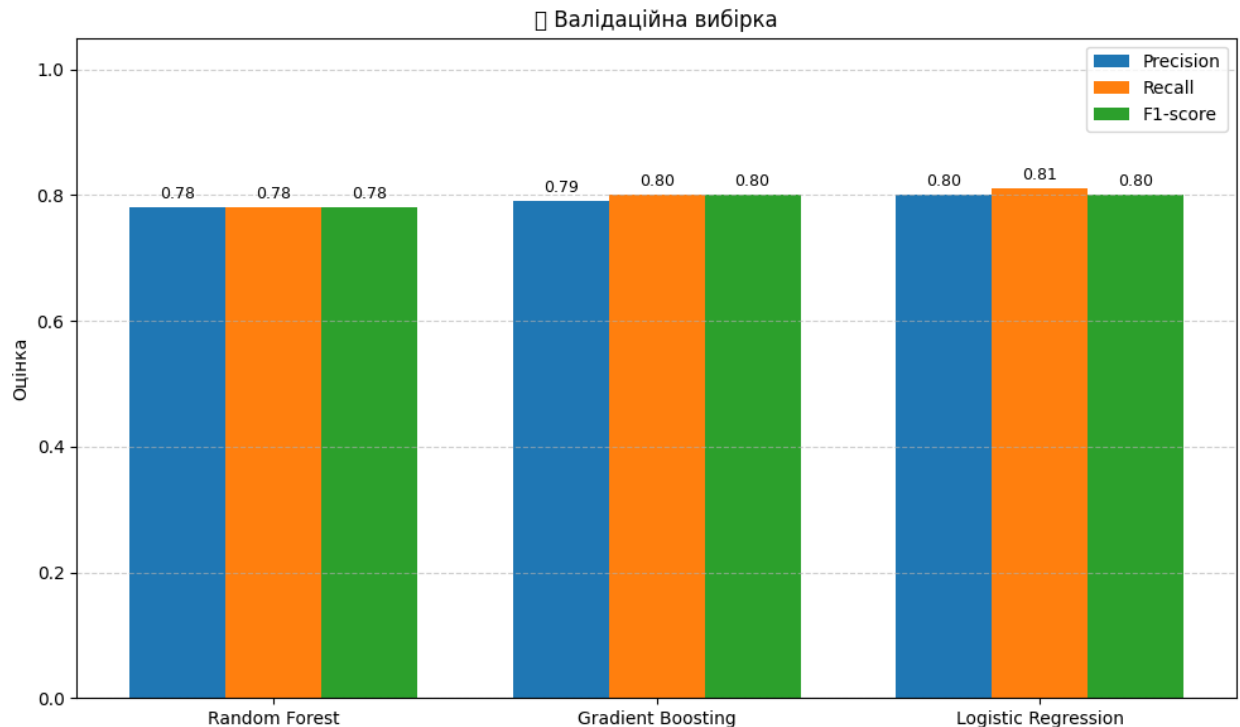


Рис. 10. Результати валідаційної вибірки.

На графіку зображено порівняння ефективності трьох моделей — Random Forest, Gradient Boosting та Logistic Regression — за основними метриками (точність, повнота, F1-метрика) на валідаційній вибірці, яка використовується для оцінки здатності моделі до узагальнення на ще не бачені дані.

Згідно з візуалізацією, усі три моделі показали дуже подібні результати, з невеликими відхиленнями, що свідчить про приблизно однакову здатність до узагальнення:

- Random Forest продемонстрував рівень точності, повноти та F1-метрики — 0.78, що трохи нижче за інші, можливо, через певну схильність до перенавчання, яку ми спостерігали на тренувальній вибірці.
- Gradient Boosting досяг значень 0.79–0.80, демонструючи кращу збалансованість між навчанням і узагальненням, з невеликою перевагою в порівнянні з Random Forest.
- Logistic Regression дещо неочікувано показала найвищу повноту — 0.81, а також однакову точність та F1-метрику, як у Gradient Boosting — 0.80. Це свідчить про її здатність досить добре передбачати всі класи навіть попри лінійну природу моделі.

Таким чином, на валідаційній вибірці усі моделі показали схожу і стабільну продуктивність, а Gradient Boosting та Logistic Regression виглядають найкраще з точки зору узагальнення. Водночас, Random Forest хоч і був ідеальним на тренуванні, на валідації дещо поступається.

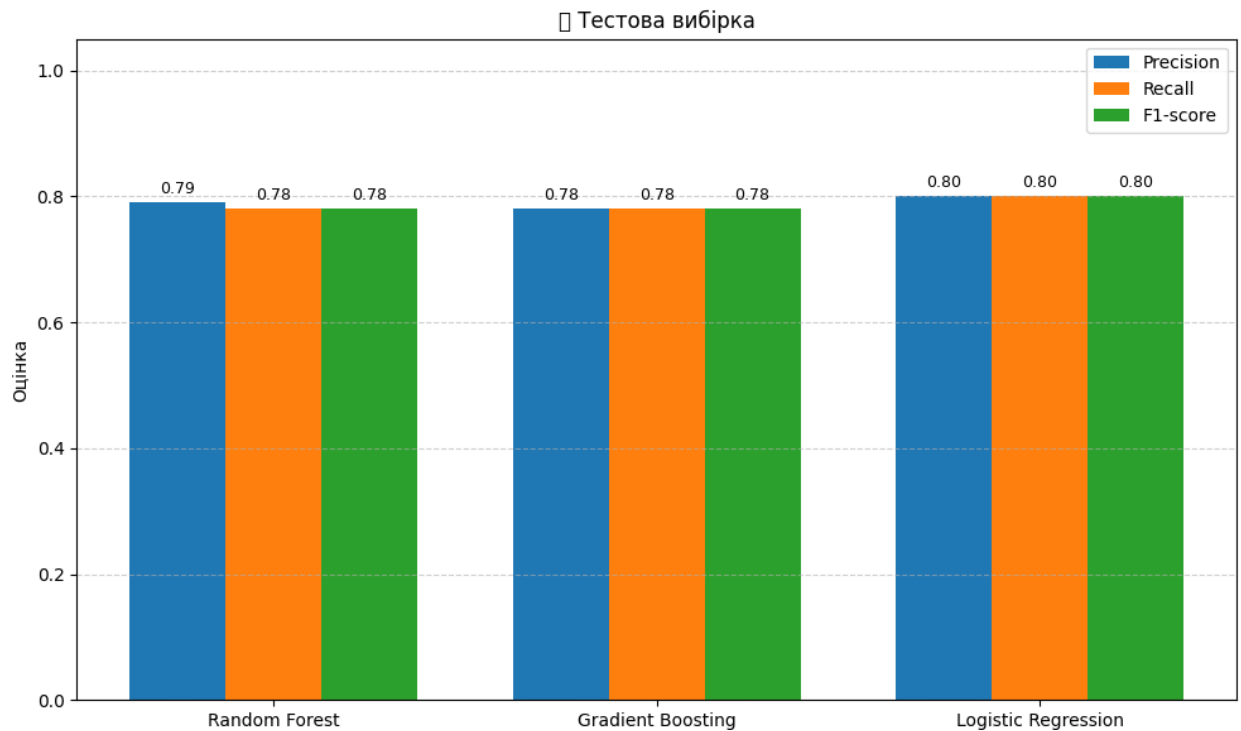


Рис. 11. Результати тестової вибірки.

На графіку представлено результати трьох моделей — Random Forest, Gradient Boosting та Logistic Regression — на тестовій вибірці, яка є фінальним незалежним набором для перевірки якості узагальнення моделей. Тестова оцінка є найоб'єктивнішою, оскільки ці дані не використовувались ні в навчанні, ні у валідації.

Згідно з результатами, усі три моделі показали дуже близькі результати, з незначними відмінностями:

- Random Forest досяг значень точності 0.79, повноти та F1-метрики — 0.78. Це дещо нижче, ніж на тренуванні, але узгоджується з валідаційними результатами, що свідчить про стабільну роботу, хоча й із певним спадом у продуктивності.
- Gradient Boosting має однакові показники — 0.78 для всіх трьох метрик, що майже ідентично Random Forest. Це підтверджує його

здатність до узагальнення та помірну стійкість, хоч він і дещо поступається очікуванням після сильнішої валідаційної оцінки.

- Logistic Regression, попри свою простоту, знову демонструє найвищу узгодженість — всі метрики становлять 0.80. Це найкращий результат серед трьох моделей на тесті, що свідчить про її стабільну роботу без явного перенавчання та здатність витримати нові, не бачені приклади.

Загалом, усі моделі показали задовільну якість узагальнення, однак Logistic Regression неочікувано виявилася найстабільнішою на тестовому наборі. Вона забезпечує найкраще поєднання точності, повноти й F1-метрики, перевершивши складніші моделі на етапі фінального оцінювання.

## **Висновок**

У даній роботі було проведено комплексне дослідження інвестиційного ринку в контексті застосування штучного інтелекту, а також здійснено спробу моделювання інвестиційного ризику компаній що працюють у сфері ІІІ. У першій частині було детально охарактеризовано структуру глобального інвестиційного ринку та його ключові ризики. Друга частина розкрила специфіку застосування інструментів штучного інтелекту у трейдингу, акцентуючи увагу на глобальних гравцях, технологіях та ефективності використання AI-рішень у сфері фінансів.

У третій частині було здійснено побудову моделей класифікації інвестиційного ризику з використанням методів машинного навчання, зокрема Random Forest, Gradient Boosting і Logistic Regression. Отримані результати свідчать про те, що класифікаційні моделі здатні з високою точністю прогнозувати рівень ризику компаній на основі доступних фінансових та ринкових показників. Проте побудова повноцінної регресійної моделі не була

реалізована через обмежену кількість даних - їх недостатній обсяг і великий вплив однієї фічі.

Отже, незважаючи на реалізацію лише класифікаційного підходу, результати цієї роботи демонструють високу перспективність застосування штучного інтелекту для аналізу та оцінювання інвестиційних ризиків компаній ШІ-сегменту. У разі розширення датасету та доступу до повнішої інформації, дослідження може бути успішно продовжено, зокрема у напрямку побудови регресійної моделі ризику, що дозволить точніше визначати не лише клас, а й кількісну оцінку потенційних втрат або рівня ризику.

## Джерела

1. I. Financial Performance of Companies from S&P 500 **Kaggle**. URL: <https://www.kaggle.com/datasets/ilyaryabov/financial-performance-of-companies-from-sp500> (дата звернення: 10.04.2025).
2. [https://www.researchgate.net/publication/385842876\\_FONDOVIJ\\_RI\\_NOK\\_UKRAINI\\_STAN\\_AKTUALNI\\_PROBLEMI\\_TA\\_PERSPEKTIVI](https://www.researchgate.net/publication/385842876_FONDOVIJ_RI_NOK_UKRAINI_STAN_AKTUALNI_PROBLEMI_TA_PERSPEKTIVI) (дата звернення: 16.06.2025).
3. **PRO CONSULTING** <https://pro-consulting.ua/ua/pressroom/rynok-obligacij-v-ukraine-osobennosti-tendencii-i-perspektivy> (дата звернення: 14.06.2025).
4. **Cryptology** [https://cryptology.school/blog/foreks-shho-take-birza-forex-i-iak-torguvati-na-fx?utm\\_source=](https://cryptology.school/blog/foreks-shho-take-birza-forex-i-iak-torguvati-na-fx?utm_source=) (дата звернення: 14.06.2025).
5. [https://analizua.com/hotovi-roboty/200-rynok-aktsii-v-ukraini-suchasnyi-stand-stand-perspektyvy-rozvytku-dyplomna-robota?utm\\_source=](https://analizua.com/hotovi-roboty/200-rynok-aktsii-v-ukraini-suchasnyi-stand-stand-perspektyvy-rozvytku-dyplomna-robota?utm_source=) (дата звернення: 10.04.2025).
6. [www.nssmc.gov.ua/finhramotnist-prostoiu-movoiu-pro-tsinni-papery-shcho-take-deryvativni-instrumenty/?utm\\_source=](http://www.nssmc.gov.ua/finhramotnist-prostoiu-movoiu-pro-tsinni-papery-shcho-take-deryvativni-instrumenty/?utm_source=) (дата звернення: 11.04.2025).
7. Бутко М.П., Нагайчук О.О. (2021). Інвестування в нерухомість: сучасні тенденції та перспективи розвитку. Науковий вісник ХНУ імені В.Н. Каразіна, серія «Економічна», №100, с. 95–101. [karazin.ua](http://karazin.ua) (дата звернення: 11.04.2025).
8. <https://www.chasopys.law.npu.kiev.ua/archive/14/10.pdf> (дата звернення: 11.04.2025).
9. [www.easterneurope-ebm.in.ua/journal/43\\_2024/15.pdf?utm\\_source=](http://www.easterneurope-ebm.in.ua/journal/43_2024/15.pdf?utm_source=) (дата звернення: 10.05.2025).
10. [https://en.wikipedia.org/wiki/2008\\_financial\\_crisis](https://en.wikipedia.org/wiki/2008_financial_crisis) (дата звернення: 12.06.2025)
11. Kaminskyi A. Portfolio management, Kyiv, Znannia, 2015 p (дата звернення: 01.06.2025)

12. Cluster Method Applying to Covid-19 Event Study for the Largest USA Banks Cluster Method Applying to Covid-19 Event Study for the Largest USA Banks Kaminskyi, A. , Nehrey, M.
13. [https://www.businessinsider.com/wall-street-goldman-jpmorgan-bridgewater-using-ai-2023-12?utm\\_source](https://www.businessinsider.com/wall-street-goldman-jpmorgan-bridgewater-using-ai-2023-12?utm_source) (дата звернення: 17.05.2025).
14. [https://bank.gov.ua/admin\\_uploads/article/FSR\\_2023-H1.pdf](https://bank.gov.ua/admin_uploads/article/FSR_2023-H1.pdf) (дата звернення: 17.05.2025).
15. Терехова М. JPMorgan takes AI use to the next level [Електронний ресурс] // **Business Insider**. URL: <https://www.businessinsider.com/jpmorgan-takes-ai-use-to-the-next-level-2017-8> (дата звернення: 17.05.2025).
16. **RBC Capital Markets**. RBC launches Aiden – a new AI-powered electronic trading platform [Електронний ресурс] // **RBC News Release**. URL: <https://www.rbc.com/newsroom/news/article.html?article=125350> (дата звернення: 21.05.2025).
17. Басак С. Bridgewater now has \$2 bn fund run by machine learning [Електронний ресурс] // **Bloomberg (via Bridgewater.com)**. URL: <https://www.bridgewater.com/in-the-news/bloomberg-on-bridgewaters-latest-innovations-in-ai-ml> (дата звернення: 21.05.2025).
18. **Two Sigma Investments**, <https://www.twosigma.com> (дата звернення: 21.05.2025).
19. [quartr.com/insights/edge/renaissance-technologies-and-the-medallion-fund?utm\\_source](https://quartr.com/insights/edge/renaissance-technologies-and-the-medallion-fund?utm_source) (дата звернення: 21.05.2025).
20. **Citadel LLC** <https://www.citadelsecurities.com/news/how-we-use-ai-to-augment-trading/> (дата звернення: 28.05.2025).
21. **Nomura Securities** [https://www.nomuraholdings.com/innovation/?utm\\_source](https://www.nomuraholdings.com/innovation/?utm_source) (дата звернення: 28.05.2025).
22. <https://www.xtxmarkets.com> (дата звернення: 28.05.2025).
23. ML <https://wepub.org/index.php/TCSISR/article/view/2986> (дата звернення: 28.05.2025).

24. DL <https://arxiv.org/pdf/2103.09750> (дата звернення: 28.05.2025).
25. NLP <https://ela.kpi.ua/items/333a9853-832f-4fce-abe1-6ad74506e023> (дата звернення: 01.06.2025).
26. RL [rbcboREALIS.com/applications/aiden/?utm\\_source](https://rbcboREALIS.com/applications/aiden/?utm_source) (дата звернення: 01.06.2025).
27. **Generative AI** <https://www.emagia.com/uk/blog/giagpt-generative-ai-for-finance/> (дата звернення: 08.06.2025).
28. **Federative Learning** [https://lucinity.com/blog/federated-learning-in-fincrime-how-financial-institutions-can-fight-crime-without-sensitive-data-sharing?utm\\_source](https://lucinity.com/blog/federated-learning-in-fincrime-how-financial-institutions-can-fight-crime-without-sensitive-data-sharing?utm_source) (дата звернення: 08.06.2025).
29. **AI in Trading** <https://cointelegraph.com/news/ai-role-trading-institutions-jpmorgan> (дата звернення: 08.06.2025).
30. **9/10 фондів** [https://www.hedgeweek.com/nine-out-ten-hedge-funds-use-ai-2023-says-survey/?utm\\_source=chatgpt.com](https://www.hedgeweek.com/nine-out-ten-hedge-funds-use-ai-2023-says-survey/?utm_source=chatgpt.com) (дата звернення: 09.06.2025).
31. [https://www.researchgate.net/figure/EurekaHedges-AI-Hedge-Fund-Index-Performance-Source-EurekaHedge\\_fig3\\_378816475](https://www.researchgate.net/figure/EurekaHedges-AI-Hedge-Fund-Index-Performance-Source-EurekaHedge_fig3_378816475) (дата звернення: 09.06.2025).
32. <https://hai.stanford.edu/ai-index/2023-ai-index-report> (дата звернення: 11.06.2025).
33. <https://blog.colobridge.net/uk/2024/05/artificial-intelligence-and-machine-learning-ua/> (дата звернення: 11.06.2025).
34. [https://uk.wikipedia.org/wiki/Біржовий\\_інвестиційний\\_фонд](https://uk.wikipedia.org/wiki/Біржовий_інвестиційний_фонд) (дата звернення: 11.06.2025).
35. Конкуренція за добу штучного інтелекта. Автори Марко Янсїті, Карім Лакані. (дата звернення: 08.06.2025).
36. Financial Relations between Hong Kong and the Mainland. License CC BY-NC 4.0. Автора Jia Hu

## Додаток

Дане посилання на Google Colab. Розрахунки і візуалізації були зроблені в ньому.

[https://colab.research.google.com/drive/1ErYwwAnM\\_CtJ0rRqOkYo2EVyiJo\\_0WIF?usp=sharing](https://colab.research.google.com/drive/1ErYwwAnM_CtJ0rRqOkYo2EVyiJo_0WIF?usp=sharing)