

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

**САВКА Андрій Петрович**

**Метод управління портфелем проєктів з використанням  
засобів штучного інтелекту / A Method of Project Portfolio  
Management Using Artificial Intelligence Tools**

спеціальність: 122 - Комп'ютерні науки  
освітньо-професійна програма – Управління проєктами

Кваліфікаційна робота

Виконав студент групи КНУПм-21  
А.П. Савка

---

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

---

Кваліфікаційну роботу  
допущено до захисту:  
«\_\_\_» \_\_\_\_\_ 20\_\_\_ р.  
В.о. завідувача кафедри  
\_\_\_\_\_ Н.В. Дзюбановська

**ТЕРНОПІЛЬ – 2025**

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

ЗАТВЕРДЖУЮ  
В.о. завідувача кафедри  
Н.М. Васильків  
« \_\_\_\_ » \_\_\_\_\_ 20\_\_ р.

**ЗАВДАННЯ  
НА КВАЛІФІКАЦІЙНУ РОБОТУ СТУДЕНТУ  
САВІЦІ Андрію Петровичу**

---

(прізвище, ім'я, по батькові)

1. Тема кваліфікаційної роботи

**Метод управління портфелем проєктів з використанням засобів штучного інтелекту / A Method of Project Portfolio Management Using Artificial Intelligence Tools**

керівник роботи к.т.н., доцент П.Є. Биковий

затверджені наказами по університету від 20 грудня 2024 року № 938 та від 14 жовтня 2025 року № 724.

2. Строк подання студентом закінченої кваліфікаційної роботи 1 грудня 2025 р.

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

4. Основні питання, які потрібно розробити

– проаналізувати предметну область управління проектами в кіноіндустрії;

– дослідити існуючі методи та засоби управління портфелем проєктів, визначити їхні обмеження;

– зробити постановку задачі дослідження;

– розробити концепцію управління портфелем кінопроєктів на основі даних, що поєднує прогнозування результатів із ризик-орієнтованим відбором проєктів;

– реалізувати підхід до прогнозування касових зборів із використанням методів штучного інтелекту;

– побудувати оптимізаційну модель формування портфеля з урахуванням ризику;

– провести експериментальну перевірку запропонованих рішень.

5. Перелік графічного матеріалу у роботі

– схема концепції управління портфелем кінопроєктів на основі даних;

– графіки з результатами експериментальних досліджень.

## 6. Консультанти розділів кваліфікаційної роботи

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

7. Дата видачі завдання 20 грудня 2024 р.

**КАЛЕНДАРНИЙ ПЛАН**

№ з/п	Назва етапів кваліфікаційної роботи	Строк виконання етапів кваліфікаційної роботи	Примітка
1	Затвердження теми кваліфікаційної роботи, ознайомлення з літературними джерелами та складання плану роботи.	до 01.01. 2025 р.	
2	Написання 1 розділу кваліфікаційної роботи	до 01.03. 2025 р.	
3	Написання 2 розділу кваліфікаційної роботи	до 20.05.2025 р.	
4	Написання 3 розділу кваліфікаційної роботи	до 28.10. 2025 р.	
5	Представлення попереднього варіанту кваліфікаційної роботи, перевірка та внесення змін керівником	до 11.11.2025 р.	
6	Опрацювання зауважень та представлення завершеного варіанту кваліфікаційної роботи. Підготовка супроводжуючих документів.	до 25.11.2025 р.	
7	Перевірка кваліфікаційної роботи на оригінальність тексту.	до 01.12.2025 р.	
8	Оформлення кваліфікаційної роботи та отримання допуску до захисту	до 04.12.2025 р.	
9	Подання кваліфікаційної роботи до захисту на засіданні атестаційної комісії.	до 14.12. 2025 р.	

Студент \_\_\_\_\_ Савка А.П.  
підписКерівник роботи \_\_\_\_\_ к.т.н., доцент П.Є. Биковий  
підпис

## РЕЗЮМЕ

Кваліфікаційна робота на тему «Метод управління портфелем проєктів з використанням засобів штучного інтелекту» на здобуття освітнього ступеня «Магістр» зі спеціальності 122 «Комп'ютерні науки» освітньої програми «Управління проєктами» написана обсягом в 82 сторінки і містить 10 ілюстрацій, 3 таблиці, 3 додатки та 41 використане джерело.

Метою кваліфікаційної роботи є розроблення комплексного підходу до управління портфелем кінопроєктів на основі даних, який поєднує ймовірнісне прогнозування касових зборів із сценарним аналізом та ризик-орієнтованою оптимізацією для формування портфеля проєктів.

Методи досліджень: методи інтелектуального аналізу даних, теорії прийняття рішень та математичної оптимізації, методи машинного навчання, математичного моделювання.

Результати дослідження: полягають у вдосконаленні методу управління портфелем кінопроєктів шляхом інтеграції ймовірнісного прогнозування касових зборів із сценарним поданням результатів та ризик-орієнтованою оптимізацією портфеля на основі CVaR, що забезпечує формування стійкіших портфельних рішень в умовах високої невизначеності та асиметричності розподілу доходів.

Результати роботи можуть успішно застосовуватися для підтримки прийняття рішень щодо інвестування та дистрибуції у кіноіндустрії під час формування портфеля кінопроєктів за обмеженого бюджету та високої невизначеності касових зборів.

Ключові слова: ПОРТФЕЛЬ ПРОЄКТІВ, УПРАВЛІННЯ ПОРТФЕЛЕМ, ЙМОВІРНІСНЕ ПРОГНОЗУВАННЯ, СЦЕНАРНИЙ АНАЛІЗ, ОПТИМІЗАЦІЯ ПОРТФЕЛЯ, РИЗИК-ОРІЄНТОВАНА ОПТИМІЗАЦІЯ, CVAR, ПІДТРИМКА ПРИЙНЯТТЯ РІШЕНЬ, ШТУЧНИЙ ІНТЕЛЕКТ, МАШИННЕ НАВЧАННЯ, АНАЛІЗ ДАНИХ.

## ABSTRACT

Qualification work on the topic «A Method of Project Portfolio Management Using Artificial Intelligence Tools» for Master's degree on speciality 122 «Computer Science» educational and professional program «Project Management» is written on 82 pages and it contains 10 figures, 3 tables, 3 annexes and 41 sources.

The purpose of this qualification work is to develop a comprehensive data-driven approach to managing a portfolio of film projects, which combines probabilistic box-office forecasting with scenario analysis and risk-oriented optimization to form a project portfolio.

Research methods: intelligent data analysis methods, decision theory and mathematical optimization methods, machine learning methods, and mathematical modeling.

Research results: consist in improving the method for managing a portfolio of film projects by integrating probabilistic box-office forecasting with a scenario-based representation of outcomes and risk-oriented portfolio optimization based on CVaR, which enables the formation of more robust portfolio decisions under high uncertainty and an asymmetric distribution of revenues.

The results of the work can be successfully applied to support decision-making on investment and distribution in the film industry when forming a portfolio of film projects under a limited budget and high uncertainty of box-office revenues.

Keywords: PROJECT PORTFOLIO, PORTFOLIO MANAGEMENT, PROBABILISTIC FORECASTING, SCENARIO ANALYSIS, PORTFOLIO OPTIMIZATION, RISK-ORIENTED OPTIMIZATION, CVAR, DECISION SUPPORT, ARTIFICIAL INTELLIGENCE, MACHINE LEARNING, DATA ANALYSIS.

## ЗМІСТ

Вступ.....	7
1 Аналіз відомих методів управління портфелем проєктів .....	10
1.1 Аналіз предметної області.....	10
1.2 Аналіз методів та засобів управління портфелем проєктів .....	12
1.3 Постановка задачі дослідження.....	20
Висновки до розділу 1 .....	22
2 Управління портфелем проєктів з використанням засобів штучного інтелекту.....	23
2.1 Концепція управління портфелем проєктів в кіноіндустрії на основі даних.....	23
2.2 Прогнозування касових зборів .....	26
2.3 Формування портфеля за допомогою оптимізаційної моделі з урахуванням ризику.....	32
Висновки до розділу 2 .....	36
3 Результати експериментальних досліджень запропонованих рішень.....	37
3.1 Експериментальні дані та їх підготовка .....	37
3.2 Результати ймовірнісного прогнозування.....	40
3.3 Результати формування портфеля на основі сценаріїв та CVaR-підходу..	43
3.4 Програмна реалізація запропонованих рішень .....	45
Висновки до розділу 3 .....	47
Висновки .....	49
Список використаних джерел.....	52
Додаток А Копії публікацій .....	56
Додаток Б Результати експериментальних досліджень .....	73
Додаток В Лістинги програмного коду.....	76

## ВСТУП

**Актуальність роботи.** Сучасна кіноіндустрія функціонує в умовах високої конкуренції за увагу аудиторії, швидких змін інформаційного середовища та суттєвої невизначеності фінансових результатів. Для дистриб'юторів та інвесторів рішення щодо фінансування й просування фільмів мають стратегічний характер, оскільки потребують розподілу обмежених ресурсів між багатьма альтернативними кінопроектами, кожен з яких характеризується власним профілем витрат, потенційної дохідності та ризику. Особливістю ринку є нерівномірність касових зборів: окремі релізи формують надвисокі доходи, тоді як значна частина проектів демонструє помірні або низькі результати. За таких умов традиційні детерміновані підходи до оцінювання ефективності, що спираються на точкові прогнози та середні значення, не забезпечують достатньої надійності портфельних рішень, оскільки не враховують ймовірність несприятливих сценаріїв і не дають інструментів контролю “хвостових” втрат.

Управління портфелем проектів у кіноіндустрії, по суті, є задачею прийняття рішень за умов ризику, де необхідно одночасно забезпечувати досягнення фінансових цілей, дотримуватися бюджетних обмежень і мінімізувати наслідки можливих провалів. На практиці портфельні рішення часто формуються на основі експертних оцінок, попереднього досвіду та ринкових індикаторів. Однак суб'єктивність експертних суджень, обмеженість історичних аналогів і динамічність зовнішнього середовища знижують якість таких рішень. У зв'язку з цим актуальним є розвиток підходів управління портфелем на основі даних, де ключову роль відіграють методи штучного інтелекту та математичної оптимізації.

Перспективним напрямом є поєднання ймовірнісного прогнозування доходів з ризик-орієнтованими оптимізаційними моделями. Ймовірнісний прогноз дозволяє відобразити невизначеність касових зборів через інтервали та сценарії можливих реалізацій, а ризик-орієнтована оптимізація – сформувати портфель, що є стійким не лише в середньому, але й у несприятливих умовах.

**Мета і завдання дослідження.** Мета роботи полягає у розробленні комплексного підходу до управління портфелем кінопроектів на основі даних, який поєднує ймовірнісне прогнозування касових зборів із сценарним аналізом та ризик-орієнтованою оптимізацією для формування портфеля проектів.

Завдання дослідження:

- проаналізувати предметну область управління проектами в кіноіндустрії;
- дослідити існуючі методи та засоби управління портфелем проектів, визначити їхні обмеження;
- зробити постановку задачі дослідження;
- розробити концепцію управління портфелем кінопроектів на основі даних, що поєднує прогнозування результатів із ризик-орієтованим відбором проектів;
- реалізувати підхід до прогнозування касових зборів із використанням методів штучного інтелекту;
- побудувати оптимізаційну модель формування портфеля з урахуванням ризику;
- провести експериментальну перевірку запропонованих рішень.

**Об’єкт дослідження** – процес управління портфелем кінопроектів у кіноіндустрії в умовах невизначеності та ризику фінансових результатів.

**Предмет дослідження** – методи й моделі управління портфелем кінопроектів на основі даних, зокрема ймовірнісне прогнозування касових зборів, формування сценаріїв та ризик-орієтована оптимізація портфеля.

**Методи дослідження** включають аналіз і синтез наукових джерел та практик управління портфелем проектів; системний підхід до побудови концепції управління портфелем кінопроектів на основі даних; методи статистичного аналізу даних і описової статистики для дослідження властивостей вибірки; методи попередньої обробки даних і кодування ознак; методи машинного навчання для ймовірнісного прогнозування касових зборів; сценарний аналіз для формування множини можливих реалізацій доходів; методи математичного моделювання та оптимізації для відбору портфеля за

бюджетних обмежень із контролем ризику на основі CVaR; методи експериментального дослідження та порівняльного аналізу для оцінювання якості прогнозування, ефективності портфельних рішень і стійкості до несприятливих сценаріїв; методи програмної інженерії для реалізації відтворюваного конвеєра обчислень і формування звітних матеріалів.

**Наукова новизна одержаних результатів** полягає у вдосконаленні методу управління портфелем кінопроектів шляхом інтеграції ймовірнісного прогнозування касових зборів із сценарним поданням результатів та ризик-орієнтованою оптимізацією портфеля на основі CVaR, що забезпечує формування стійкіших портфельних рішень в умовах високої невизначеності та асиметричності розподілу доходів.

**Практичне значення отриманих результатів** полягає у можливості застосування розробленого підходу як інструментарію підтримки прийняття рішень для дистриб'юторів та інвесторів у кіноіндустрії під час формування портфеля кінопроектів за умов обмеженого бюджету та високої невизначеності касових зборів. Запропонований метод забезпечує отримання ймовірнісних прогнозів доходів, формування сценаріїв можливих результатів та вибір портфеля з контролем ризику негативних сценаріїв, що підвищує стійкість фінансового результату портфеля та зменшує ймовірність суттєвих втрат.

**Публікації та апробація КР.** Результати кваліфікаційної роботи апробовані та опубліковані у матеріалах (додаток А):

- II Всеукраїнської науково-практичної конференції «Інтелектуальні комп'ютерні системи та мережі», 25 листопада 2025 р., Тернопіль, Україна;
- 2nd International Scientific and Practical Conference «Progressive Approaches in Science and Engineering», November 26-28, 2025. Copenhagen, Denmark.

Кваліфікаційна робота складається із вступу, трьох розділів, висновків, списку використаних джерел та додатків.

# 1 АНАЛІЗ ВІДОМИХ МЕТОДІВ УПРАВЛІННЯ ПРОРТФЕЛЕМ ПРОЄКТІВ

## 1.1 Аналіз предметної області

Кіноіндустрія є одним із найвпливовіших та економічно значущих секторів світової креативної економіки. Як зазначають дослідники, культура та мистецтво є життєво важливими компонентами розвитку будь-якого суспільства, роблячи суттєвий внесок у прогрес нації [21]. Глобальні касові збори, що сягають десятків мільярдів доларів щорічно, та мільйони робочих місць по всьому світу підкреслюють її масштаб [19]. Проте за блискучим фасадом криється бізнес-середовище, що характеризується надзвичайно високим рівнем ризику та невизначеності.

Ключовим викликом для всіх учасників ринку, від продюсерів до дистриб'юторів, є прогнозування комерційного успіху кінопроєкту. На відміну від традиційних галузей, де успіх продукту можна з певною точністю розрахувати на основі виробничих витрат та ринкового попиту, успіх фільму залежить від складної та важко прогнозованої взаємодії багатьох чинників. До них належать жанр, бюджет, маркетингова кампанія, сезонність прокату, конкурентне середовище та, що найголовніше, реакція аудиторії, яка часто є ірраціональною.

Однією з фундаментальних характеристик предметної області є асиметричний характер фінансових результатів. Дослідження показують, що доходи в кіноіндустрії розподіляються за принципом "переможець отримує все". Незначна кількість фільмів-блокбастерів генерує левову частку прибутків усієї галузі, тоді як переважна більшість проєктів ледве покривають витрати або є збитковими [11]. Ця властивість робить інвестиції в кіно надзвичайно ризикованими і ставить перед інвесторами та дистриб'юторами складне завдання – ідентифікувати потенційні "хіти" серед сотень доступних проєктів.

У спробах знизити рівень невизначеності стейкхолдери індустрії спираються на аналіз певних атрибутів проєкту. Одним із найвагоміших факторів, що історично вважається запорукою успіху, є залучення

знаменитостей – акторів, режисерів та сценаристів. "Зіркова сила" розглядається як інструмент маркетингу, що привертає увагу аудиторії та слугує своєрідним знаком якості, знижуючи для глядача ризик розчарування [18]. Присутність відомого імені в титрах може суттєво впливати на рішення споживачів [5, 14]. Однак кількісна оцінка цього впливу завжди була складною задачею, що традиційно покладалася на інтуїцію та суб'єктивний досвід експертів, що неминуче несе в собі ризик упередженості [34].

Саме тому центральним процесом в управлінні інвестиціями в кіноіндустрії є портфельний менеджмент. Для дистриб'юторської компанії успіх залежить не стільки від одного вдалого проєкту, скільки від формування збалансованого портфеля, де потенційні надприбутки від блокбастерів могли б компенсувати збитки від менш вдалих фільмів. Процес управління портфелем, як його визначають у класичній фінансовій теорії, передбачає вибір та управління сукупністю активів для досягнення конкретних інвестиційних цілей [13]. В контексті кіноіндустрії це означає ухвалення обґрунтованих рішень щодо того, в які саме кінопроєкти інвестувати, збалансовуючи потенційні доходи та ризики.

Існуючі наукові дослідження в межах медійного ланцюга постачання фокусувалися на окремих аспектах, таких як оптимізація рекламних розкладів [2] чи розробка рекомендаційних систем [33]. Проте, як зазначається у вступі до цієї роботи, комплексна задача оптимізації інвестиційного портфеля дистриб'ютора залишалася недостатньо дослідженою. Традиційні підходи, засновані виключно на експертних оцінках, не здатні впоратися з великими обсягами даних та складністю взаємозв'язків між факторами.

Таким чином, аналіз предметної області виявляє гостру потребу в розробці нових, більш досконалих методів управління портфелем у кіноіндустрії. Такі методи мають спиратися на об'єктивний аналіз даних, враховувати асиметричний характер ринку, дозволяти кількісно оцінювати неформалізовані чинники (як-от популярність знаменитостей) та інтегрувати стратегічні вподобання особи, що ухвалює рішення. Саме на вирішення цієї комплексної задачі і спрямоване дане дослідження.

## 1.2 Аналіз методів та засобів управління портфелем проєктів

Управління портфелем проєктів є дисципліною високого рівня, що інтегрує процеси управління проєктами зі стратегічним плануванням організації. Якщо управління проєктами відповідає на питання "Як нам правильно виконувати проєкти?", то PPM відповідає на більш фундаментальне питання: "Які проєкти нам слід виконувати?". Згідно зі стандартом Інституту управління проєктами, портфель визначається як набір проєктів, програм, підпортфельів та операційної діяльності, керованих як група для досягнення стратегічних цілей [25]. Отже, ключова мета PPM полягає в максимізації цінності портфеля та забезпеченні його збалансованості шляхом систематичного процесу відбору, пріоритезації та моніторингу проєктів.

Аналіз існуючих методів є критично важливим, оскільки він дозволяє зрозуміти їх еволюцію від простих фінансових розрахунків до складних, інтегрованих, керованих даними систем, і, як наслідок, обґрунтувати вибір методології, що є найбільш адекватною для вирішення завдань у такій специфічній галузі, як кіноіндустрія.

### 1.2.1 Класичні методи відбору проєктів: фінансові та стратегічні підходи

Історично перші методи PPM виростили з корпоративних фінансів та якісного стратегічного аналізу. Вони заклали основу для дисципліни, проте їхні обмеження стали стимулом для подальшого розвитку.

#### 1.2.1.1 Фінансові моделі

Ця група методів розглядає проєкти як інвестиційні інструменти, які повинні генерувати фінансову віддачу. Їхня головна перевага – об'єктивність та порівнянність результатів, виражених у грошових одиницях.

Чиста приведена вартість (Net Present Value, NPV): найбільш теоретично обґрунтований метод, що розраховує різницю між дисконтованою вартістю майбутніх грошових потоків від проєкту та сумою початкових інвестицій.

Проект вважається прийнятним, якщо його  $NPV > 0$ . При виборі між альтернативними проектами перевага надається тому, що має вищий  $NPV$ .

Внутрішня норма дохідності (Internal Rate of Return, IRR): показує ставку дисконтування, за якої  $NPV$  проекту дорівнює нулю. Фактично, IRR є очікуваною дохідністю проекту. Рішення приймається шляхом порівняння IRR із вартістю капіталу (WACC) компанії.

Період окупності (Payback Period): дуже простий метод, що визначає час, необхідний для того, щоб кумулятивні грошові потоки від проекту зрівнялися з початковими інвестиціями.

Обмеження фінансових методів у контексті кіноіндустрії: незважаючи на їхню фундаментальність [7], ці методи мають серйозні обмеження при застосуванні до кінопроектів. Їхня точність повністю залежить від якості прогнозування майбутніх касових зборів, що, як відомо, є надзвичайно складним завданням, влучно описаним афоризмом "ніхто нічого не знає". Крім того, ці методи не враховують нефінансові аспекти, такі як створення інтелектуальної власності (потенціал для франшизи), вплив на бренд студії, синергію з іншими проектами або художню цінність.

#### 1.2.1.2 Якісні (стратегічні) моделі

Ці моделі були розроблені для подолання обмежень суто фінансового підходу шляхом включення стратегічних чинників.

Моделі скорингу (Scoring Models): є одними з найпоширеніших у практиці RRM. Процес включає кілька етапів: 1) визначення набору критеріїв (наприклад, відповідність стратегії, ринковий потенціал, технологічний ризик, наявність ресурсів); 2) присвоєння ваги кожному критерію; 3) оцінка кожного проекту за кожним критерієм; 4) розрахунок загального балу як зваженої суми оцінок. Проекти з найвищим балом отримують найвищий пріоритет.

Портфельні матриці (напр., діаграми "бульбашок"): це візуальні інструменти, що допомагають збалансувати портфель. Проекти зображуються у вигляді "бульбашок" на двовимірній сітці. Осями можуть бути "ймовірність

успіху – потенційна винагорода", "стратегічна важливість – складність реалізації" тощо. Розмір бульбашки може представляти бюджет проекту.

Обмеження якісних методів: хоча ці методи краще враховують стратегію [10], їхній головний недолік полягає у високому ступені суб'єктивності. Вибір критеріїв, визначення їх ваги та виставлення оцінок залежать від експертів, що може призводити до упередженості, внутрішніх "політичних ігор" та маніпуляцій. Саме цю проблему суб'єктивності частково вирішує запропонований у даній роботі підхід із залученням LLM.

### 1.2.2 Сучасні кількісні та гібридні підходи

Розвиток обчислювальної техніки та аналітичних наук стимулював появу більш досконалих методів, здатних обробляти значно складніші моделі.

#### 1.2.2.1 Методи математичної оптимізації

Ці методи формують задачу вибору проектів як задачу математичного програмування, що дозволяє знайти математично доведене оптимальне рішення.

Цілочисельне та бінарне програмування: цей клас моделей ідеально підходить для задач відбору, де для кожного проекту потрібно прийняти бінарне рішення: "так" (змінна дорівнює 1) або "ні" (змінна дорівнює 0). Цільова функція, як правило, максимізує сумарну цінність (напр., NPV) портфеля, а обмеження відображають ліміти ресурсів (бюджет, персонал тощо).

Багатокритеріальна оптимізація (Multi-Objective Optimization): це потужне розширення класичної оптимізації, яке визнає, що організації майже завжди переслідують кілька цілей одночасно, і ці цілі часто є суперечливими. Наприклад, компанія може прагнути максимізувати прибуток, мінімізувати ризик та максимізувати частку ринку. Замість одного оптимального рішення, ці методи генерують набір компромісних, так званих Парето-оптимальних рішень [12]. Це дає особі, що приймає рішення, можливість обрати той баланс між критеріями, який найкраще відповідає поточній стратегії.

Обмеження оптимізаційних методів: їхня ефективність повністю залежить від якості вхідних даних. Якщо прогнози прибутку або оцінки стратегічної

цінності є неточними, навіть найскладніша оптимізаційна модель дасть нерелевантний результат ("сміття на вході – сміття на виході").

#### 1.2.2.2 Методи, що базуються на аналізі даних та машинному навчанні

Поява та стрімкий розвиток методів машинного навчання ознаменували собою парадигматичний зсув в управлінні портфелем проєктів, особливо в галузях з високим рівнем невизначеності. Якщо класичні підходи покладалися на історичні середні показники, фінансові моделі з жорсткими припущеннями та суб'єктивні експертні оцінки, то ML-методи пропонують принципово інший підхід: виявлення складних, нелінійних та прихованих закономірностей у великих масивах історичних даних для побудови точних та об'єктивних прогнозів. У контексті кіноіндустрії, де успіх проєкту визначається десятками взаємопов'язаних факторів, застосування машинного навчання є не просто вдосконаленням, а необхідністю для переходу від інтуїтивного до керованого даними (data-driven) ухвалення рішень.

Серцевиною застосування ML в управлінні портфелем є прогнозна аналітика. Її мета – надати кількісні, об'єктивні оцінки ключових показників майбутніх проєктів, які згодом слугуватимуть вхідними даними для моделей пріоритезації та оптимізації. У даному дослідженні прогнозна аналітика реалізується через два основні класи моделей.

##### А. Регресійні моделі для прогнозування кількісних показників

Регресійний аналіз – це клас алгоритмів ML, завданням яких є моделювання залежності між набором вхідних ознак (незалежних змінних) та безперервною вихідною змінною (залежною змінною). У контексті даної роботи:

Вхідні ознаки (X): характеристики кінопроєкту, такі як бюджет, жанр, рейтинг МРАА, місяць виходу, оцінки на IMDb, а також інноваційні, згенеровані за допомогою LLM, показники популярності режисера, сценариста та головного актора.

Вихідна змінна (Y): ключовий показник комерційного успіху – касові збори (box office revenue).

У дослідженні застосовується цілий спектр регресійних алгоритмів, від простих до складних, для вибору найбільш ефективної моделі:

Лінійна регресія (та її регуляризовані версії Ridge, LASSO): базові моделі, що припускають лінійну залежність між ознаками та результатом. Вони є швидкими та легко інтерпретованими, але рідко здатні вловити складні взаємозв'язки в реальних даних.

Моделі на основі дерев рішень (Decision Tree, Random Forest): ці алгоритми не роблять припущень про лінійність. "Випадковий ліс" (Random Forest) є ансамблевим методом, що будує велику кількість незалежних дерев рішень і усереднює їхні прогнози. Це робить його надзвичайно потужним, стійким до перенавчання та здатним моделювати складні нелінійні залежності, що є типовим для кіноринку.

Моделі на основі градієнтного бустингу (Gradient Boosting, XGBoost, LightGBM, CatBoost): найсучасніші та найефективніші алгоритми для роботи з табличними даними. На відміну від Random Forest, вони будують дерева послідовно, де кожне наступне дерево намагається виправити помилки попереднього. Це дозволяє досягати найвищої точності прогнозування, що й підтверджується результатами даного дослідження.

#### Б. Класифікаційні моделі для роботи з асиметричними даними

Однією з ключових проблем даних у кіноіндустрії є їхня крайня асиметричність ("winner-take-all"). Пряме застосування регресійних моделей до такого набору даних може призвести до того, що модель буде добре прогнозувати "середні" проєкти, але систематично помилятиметься на крайніх значеннях – як на касових провалах, так і на блокбастерах.

Для вирішення цієї проблеми в даній роботі застосовується двохетапний підхід, де першим кроком є класифікація. Класифікація – це задача ML, де модель прогнозує не числове значення, а приналежність об'єкта до певного класу. У нашому випадку фільми класифікуються на три категорії за рівнем потенційного доходу (напр., "низький дохід", "середній дохід", "високий дохід / блокбастер"). Це дозволяє:

1. Структурувати проблему: розбити одну складну задачу прогнозування на кілька простіших.

2. Підвищити точність: на другому етапі для кожної з цих більш однорідних груп даних будується своя, окрема регресійна модель. Модель, навчена лише на даних про блокбастери, буде значно точніше прогнозувати дохід нового потенційного блокбастера, ніж модель, навчена на всіх фільмах одразу.

Одним із ключових наукових внесків даної роботи є розв'язання давньої проблеми кількісної оцінки "зіркової сили" (star power). Традиційно цей фактор оцінювався суб'єктивно. Запропонований підхід полягає у використанні великої мовної моделі (напр., Chat GPT) як "синтетичного експерта" або "агрегатора знань".

Процес полягає у формулюванні точних запитів (промптів) до LLM, де вказується ім'я знаменитості (актора, режисера) та конкретний рік (рік виходу фільму), і запитується оцінка її рівня популярності/впливовості за заданою шкалою (напр., від 0 до 10). Цей підхід має низку фундаментальних переваг:

1. Об'єктивність та відтворюваність: LLM, на відміну від людини-експерта, позбавлена особистих симпатій, втоми чи настрою. Її оцінки є стабільними та відтворюваними.

2. Доступ до величезного масиву знань: модель навчалася на терабайтах текстової інформації з Інтернету (статті, новини, рецензії, блоги), що дозволяє їй синтезувати узагальнене уявлення про публічне сприйняття особистості в певний момент часу.

3. Перетворення якісної інформації в кількісну: цей процес дозволяє перевести абстрактне поняття "популярність" у конкретну числову ознаку, яку можна безпосередньо використовувати в математичних моделях машинного навчання, значно підвищуючи їхню прогнозну силу, що й було доведено в розділі аналізу чутливості.

Таким чином, методи, що базуються на аналізі даних та машинному навчанні, відіграють у даному дослідженні критично важливу, фундаментальну роль. Вони є тим "двигуном", який перетворює "сирі" історичні дані та

неструктуровану текстову інформацію на якісні, об'єктивні та надійні прогнози. Саме ці прогнозовані показники (клас проєкту, очікувані касові збори) та згенеровані ознаки (рівень популярності) слугують надійним фундаментом для наступних етапів методології – багатокритеріального аналізу (MADM) та математичної оптимізації. Без цього фундаменту будь-яка спроба оптимізації портфеля була б побудована на хистких припущеннях, що нівелювало б її практичну цінність.

### 1.2.2.3 Гібридні методи

Найбільш перспективними є гібридні підходи, що поєднують сильні сторони різних методів в єдину інтегровану методологію. Дослідження показують, що найбільш ефективні системи PPM комбінують кілька технік [3]. Типовий гібридний фреймворк може виглядати так:

1. Етап прогнозування: моделі машинного навчання використовуються для генерації об'єктивних прогнозів фінансових показників.

2. Етап оцінки: методи багатокритеріального аналізу рішень (MADM), такі як ANP, ANP, або, як у даній роботі, BBWM та WASPAS, використовуються для оцінки проєктів за якісними критеріями та розрахунку інтегрального показника пріоритетності.

3. Етап оптимізації: результати попередніх етапів слугують вхідними даними для багатокритеріальної оптимізаційної моделі, яка знаходить оптимально збалансований портфель.

Саме така гібридна архітектура лежить в основі методології, що представлена в роботі Alipour-Vaezi та Tsui [1], де запропоновано цілісну data-driven методологію управління портфелем кінопроєктів, орієнтовану на потреби дистриб'ютора/інвестора, який формує набір фільмів за умов високої невизначеності касових зборів. Автори підкреслюють, що управління портфелем у кіноіндустрії ускладнюється поєднанням двох чинників: (1) істотним впливом “зірковості” ключових осіб (режисера/сценариста/головної ролі) на реакцію аудиторії та (2) асиметричною структурою ринку, коли невелика частка проєктів

генерує основну частку доходів, а отже хвостові події (“раптовий провал” або “вибуховий успіх”) мають визначальний вплив на портфельний результат.

Методологічно підхід побудовано як послідовність взаємопов’язаних етапів: оцінювання показника відомості (fame score) для ключових “селебріті” за допомогою LLM, подальша класифікація фільмів на групи для зменшення ефекту асиметрії даних, прогнозування касових зборів у межах кожного класу, обчислення інтегральної “preferability” (переважності) проєктів для дистриб’ютора за багатокритеріальним підходом і, нарешті, розв’язання оптимізаційної задачі формування портфеля [1].

Особливістю є саме використання LLM як “експерта” для кількісної формалізації складно вимірюваної змінної “зірковості”: автори описують, що ChatGPT отримує на вхід ім’я та рік релізу і повертає оцінку за шкалою 0–10 для режисера/сценариста/головної ролі.

Для моделювання застосовано набір ознак, який охоплює контентні та ринкові характеристики (жанр, рейтинг, місяць релізу, IMDb-оцінки та кількість голосів, бюджет, тривалість, ознаку франшизи тощо) та три fame-scores, а також наведено типові процедури передобробки даних (робота з пропусками, аномаліями, нормалізація, one-hot кодування, SMOTE, корекція на інфляцію CPI, PCA) [1].

Щоб зменшити вплив дисбалансу та важкохвостості, автори спершу виконують класифікацію на три класи за діапазонами касових зборів (пороги задані експертно), після чого навчають регресійні моделі окремо в кожному класі. Показовим є те, що “верхній” клас містить дуже малу частку спостережень (порядку кількох відсотків), що ілюструє структурну асиметрію ринку та необхідність спеціальних процедур для хвоста розподілу.

Для прогнозування використовуються стандартні ML-регресори (зокрема лінійні моделі, дерева рішень, випадковий ліс, градієнтний бустинг, XGBoost) із оцінюванням за MAE/MSE/RMSE/MAPE/R<sup>2</sup> [1].

### 1.2.3 Програмні засоби управління портфелем проєктів

Практична реалізація описаних методів неможлива без відповідних програмних засобів. Їх можна класифікувати наступним чином:

1. Засоби загального призначення (напр., Microsoft Excel): доступні та гнучкі, але не призначені для складного аналізу, спільної роботи та управління великими обсягами даних.

2. Спеціалізоване ПЗ для PPM (напр., Microsoft Project Server, Planview, Broadcom Clarity, Jira Align): пропонують комплексні платформи для управління всім життєвим циклом портфеля.

3. Аналітичні та математичні платформи (напр., GAMS, MATLAB, SAS) та мови програмування (Python, R): єдині засоби, що дозволяють реалізовувати кастомні, складні гібридні моделі, що поєднують предиктивну аналітику та багатокритеріальну оптимізацію. Використання Python для машинного навчання та GAMS для оптимізації в даній роботі є прикладом застосування такого потужного інструментарію.

### 1.3 Постановка задачі дослідження

Проведений комплексний аналіз предметної області та існуючих методів управління портфелем проєктів виявив низку фундаментальних проблем, які перешкоджають ефективному ухваленню інвестиційних рішень у кіноіндустрії.

По-перше, класичні фінансові моделі виявляються недостатньо ефективними через принципову складність точного апріорного прогнозування грошових потоків у середовищі, де попит є мінливим, а результати окремих релізів часто мають характер «непередбачуваного успіху» або «раптового провалу».

По-друге, якісні та скорингові методи страждають від високого ступеня суб'єктивності та потенційної упередженості експертних оцінок, що знижує відтворюваність і прозорість портфельних рішень.

По-третє, значна частина наявних підходів не пропонує достатнього інструментарію для коректної роботи з екстремальною асиметричністю даних касових зборів, яка є іманентною властивістю ринку та призводить до домінування «хвостових» подій у загальному результаті портфеля.

По-четверте, багато моделей орієнтовані на один критерій (переважно фінансовий), не забезпечуючи узгодженого врахування стратегічних цілей, вимог до стійкості та ризик-орієнтації інвестора.

І нарешті, базові підходи до портфельної оптимізації у кіно переважно використовують детерміновані оцінки доходів і не містять явного механізму контролю хвоста розподілу результатів, що обґрунтовує необхідність переходу до ймовірнісного прогнозування, сценарного аналізу та ризик-орієнтованої оптимізації.

Ця сукупність невирішених питань формує науково-практичну прогалину та обумовлює необхідність створення нової інтегрованої методології. Таким чином, основною науково-практичною задачею дослідження є розроблення комплексного підходу до формування оптимального портфеля кінопроектів, який, на відміну від існуючих рішень, поєднує ймовірнісне прогнозування касових зборів, сценарне подання невизначеності та ризик-орієнтовану оптимізацію портфеля з контролем хвостових втрат.

Для вирішення цієї задачі, її необхідно декомпонувати на низку послідовних, логічно пов'язаних дослідницьких завдань:

- проаналізувати предметну область управління проектами в кіноіндустрії;
- дослідити існуючі методи та засоби управління портфелем проектів, визначити їхні обмеження;
- зробити постановку задачі дослідження;
- розробити концепцію управління портфелем кінопроектів на основі даних, що поєднує прогнозування результатів із ризик-орієнтованим відбором проектів;
- реалізувати підхід до прогнозування касових зборів із використанням методів штучного інтелекту;
- побудувати оптимізаційну модель формування портфеля з урахуванням ризику;
- провести експериментальну перевірку запропонованих рішень.

## Висновки до розділу 1

1. Виконано аналіз предметної області управління портфелем кінопроектів та узагальнено ключові фактори, що впливають на комерційний успіх фільмів і прийняття інвестиційних рішень у кіноіндустрії. Показано, що ринок характеризується високою невизначеністю попиту, значною конкуренцією за увагу аудиторії та екстремальною асиметричністю касових зборів, унаслідок чого навіть наявність детальної інформації про проєкт не гарантує точного апріорного прогнозування результатів.

2. Проаналізовано існуючі підходи до оцінювання та відбору кінопроектів: класичні фінансові моделі, експертно-скорингові методи, методи машинного навчання для прогнозування касових зборів, а також оптимізаційні моделі формування портфеля. Встановлено, що традиційні фінансові підходи обмежені детермінованістю та слабкою здатністю враховувати нестабільність ринку, тоді як експертні методи мають проблему суб'єктивності та низької відтворюваності оцінок.

3. Детально розглянуто сучасну data-driven методологію, яка поєднує прогнозування, ранжування альтернатив і оптимізаційне формування портфеля. Показано її практичну цінність, зокрема використання LLM для формалізації чинника “зірковості” та застосування комбінованих ML-моделей. Водночас обґрунтовано ключові обмеження підходу: детермінованість фінальної оптимізації, відсутність імовірнісного опису невизначеності, чутливість до експертно заданих порогів класифікації, а також відсутність прямого механізму контролю хвостових ризиків.

4. На основі виявлених проблем сформульовано постановку задачі дослідження: розробити інтегрований підхід до управління портфелем кінопроектів на основі даних, яка передбачає ймовірнісне прогнозування касових зборів, формування сценарного простору можливих реалізацій і побудову ризик-орієнтованої оптимізаційної моделі формування портфеля з контролем хвостових втрат при збереженні фінансової логіки “дохід–витрати–оподаткування” у сценарному вигляді.

## 2 УПРАВЛІННЯ ПОРТФЕЛЕМ ПРОЄКТІВ З ВИКОРИСТАННЯМ ЗАСОБІВ ШТУЧНОГО ІНТЕЛЕКТУ

### 2.1 Концепція управління портфелем проєктів в кіноіндустрії на основі даних

Управління портфелем кінопроєктів у кіноіндустрії є задачею прийняття рішень за умов істотної невизначеності, коли необхідно одночасно досягати стратегічних і фінансових цілей, узгоджувати бюджетні та ресурсні обмеження, а також контролювати ризик низьких касових зборів. Складність зумовлена багатофакторністю успіху фільму (жанр, сезонність, конкуренція, маркетинг, реакція аудиторії тощо) та тим, що ці чинники взаємодіють нелінійно й важко прогнозовано [19].

Водночас ринок має виражено асиметричний характер: невелика частка «хітів» генерує основну частину доходів, тоді як більшість проєктів мають значно нижчі результати, що підвищує «хвостові» ризики портфеля [13].

На практиці дистриб'ютор або інвестор ухвалює рішення щодо включення фільмів до портфеля, спираючись на історичний досвід, експертні оцінки та ринкові сигнали. Однак за сучасних умов (цифровізація каналів, зміни споживчої поведінки, вплив зовнішніх подій) така інтуїтивна логіка часто не забезпечує відтворюваності та прозорості рішень, що посилює потребу у формалізованій, керованій даними концепції портфельного управління [51].

Запропонована в даній роботі концепція ґрунтується на дисципліні управління портфелем проєктів (Project Portfolio Management), яка розглядає портфель як сукупність проєктів/програм/підпортфелів, що керуються як група для досягнення стратегічних цілей організації [42].

У контексті кіноіндустрії це узгоджується з логікою портфельного підходу у фінансовій теорії: успіх дистриб'ютора визначається не одиничним «вдалим» проєктом, а збалансованістю всієї сукупності інвестицій і керуванням ризиком на рівні портфеля [16, 50].

Критично важливим чинником, який традиційно впливає на рішення, є «зіркова сила» (акторів/режисерів/сценаристів): вона може підвищувати увагу

аудиторії, але її кількісна оцінка часто залишається суб'єктивною і схильною до упереджень експертів [7, 17, 29, 59].

Саме тому в межах концепції управління на основі даних пропонується поєднувати об'єктивні числові характеристики проєкту з формалізованим урахуванням якісних чинників (зокрема пов'язаних із репутацією та популярністю), щоб зменшити залежність від інтуїції та підвищити відтворюваність оцінювання.

Ключова ідея концепції полягає в тому, що для якісного формування портфеля недостатньо оцінювати лише «очікуваний» фінансовий результат кожного фільму. Потрібно враховувати ризик небажаних сценаріїв – ситуацій, коли касові збори суттєво нижчі за прогнозні очікування або не досягають заданих порогів. За асиметричних розподілів доходів [13] контроль таких сценаріїв стає центральним елементом портфельного рішення.

Тому штучний інтелект у даній концепції розглядається не як інструмент отримання одного «точкового» прогнозу, а як основа для побудови імовірнісного опису результату – такого, що відображає невизначеність і підтримує сценарне мислення під час ухвалення рішень.

Концептуальна модель управління портфелем включає послідовність взаємопов'язаних етапів (рисунок 2.1).

Перший етап – формування інформаційної бази. Збираються та узгоджуються дані, що описують кінопроєкти і контекст їх прокату: атрибути фільму (жанр, тривалість, рейтинг, мова, країна виробництва), параметри виробництва і дистрибуції, характеристики творчої команди, а також історичні результати аналогічних проєктів. Оскільки успіх фільму залежить від складної взаємодії чинників [19], важливо забезпечити якість даних (узгодженість форматів, коректність значень, роботу з пропусками, контроль викидів) та коректне відображення часової динаміки ринку.

Другий етап – імовірнісне прогнозування касових зборів. На відміну від класичного підходу, де результатом є одне число, у даній роботі прогнозування має формувати розподіл можливих значень касових зборів. Практично це доцільно реалізовувати через квантильні оцінки (наприклад, песимістичний,

медіанний та оптимістичний рівні), що дає змогу безпосередньо перейти до аналізу «кращих» і «гірших» сценаріїв – а отже, до контролю ризику, який посилюється асиметрією доходів у кіноіндустрії [13].

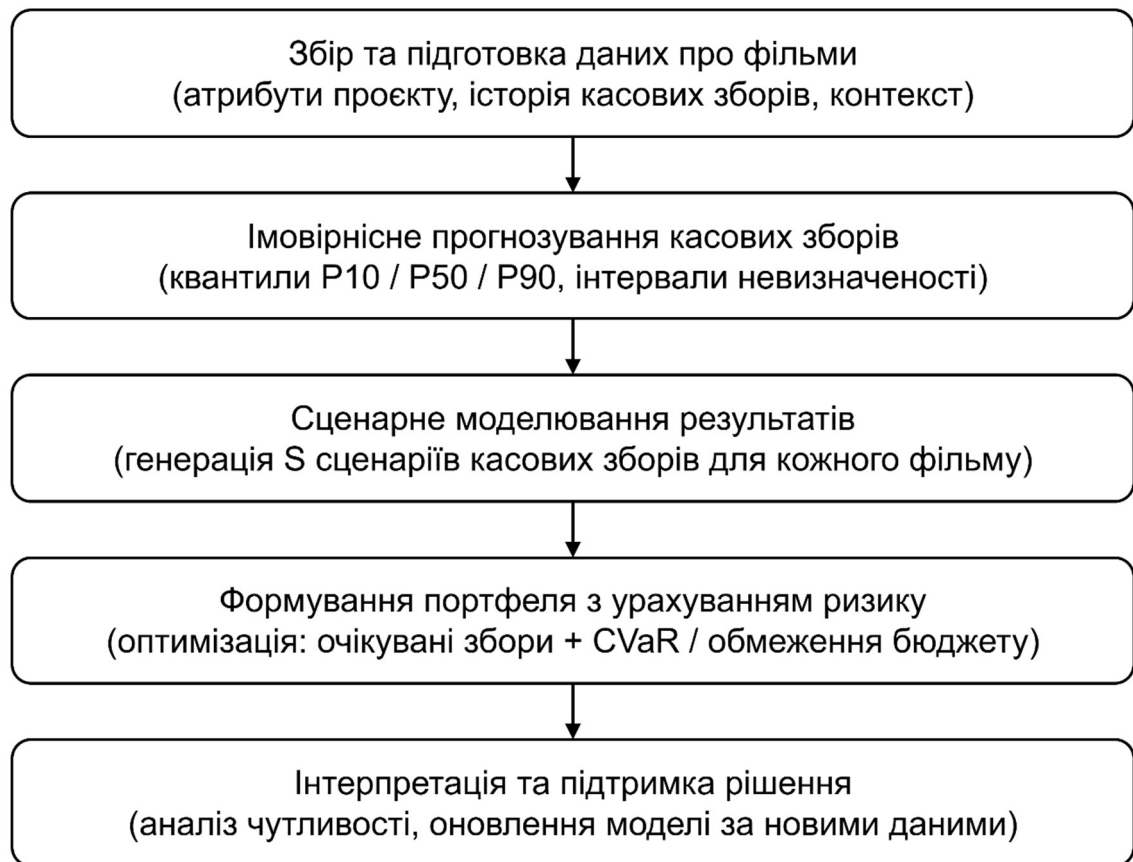


Рисунок 2.1 – Схема концепції управління портфелем кінопроектів на основі даних

Третій етап – перетворення прогнозу у сценарний набір. Імовірнісний опис результату трансформується у множину сценаріїв, придатних для оптимізаційного формування портфеля. Сценарії інтерпретуються як альтернативні стани ринку та реакції аудиторії, що узгоджується з багатофакторною природою комерційного успіху фільму [19].

Четвертий етап – формування портфеля на основі оптимізаційної моделі з урахуванням ризику. Портфель задається бінарними рішеннями щодо включення кожного фільму та обмеженнями (бюджет, ресурсні/стратегічні вимоги, потреба в диверсифікації тощо). Критерій вибору має відображати компроміс між очікуваною фінансовою результативністю портфеля (логіка портфельного

інвестування [16, 50]) та контролем ризику небажаних сценаріїв, що є принциповим за «перекошених» розподілів доходів [13].

П'ятий етап – інтерпретація рішення та зворотний зв'язок. Результатом є не лише перелік відібраних проєктів, а й пояснювані характеристики: очікувані показники портфеля, ризик-профіль, чутливість до бюджету/параметрів моделі. Після фактичного прокату отримані дані повертаються до інформаційної бази для оновлення моделей, що особливо важливо в умовах змін споживчої поведінки та зовнішніх шоків [51].

Таким чином запропонована концепція переводить формування портфеля у площину відтворюваного, кількісно обґрунтованого ризик-орієнтованого управління: невизначеність не ігнорується, а формалізується через імовірнісні прогнози (квантили) та сценарії.

## 2.2 Прогнозування касових зборів

Ключовою передумовою ефективного управління портфелем кінопроєктів є коректне прогнозування фінансового результату від кожного фільму. Проте для кіноіндустрії характерна висока невизначеність: розподіл касових зборів є асиметричним, із «важкими хвостами», коли невелика частка фільмів формує непропорційно великий внесок у загальний дохід. Це узгоджується з результатами емпіричних досліджень кіноринку, які обґрунтовують стабільну паретівську поведінку прибутків та «ефект суперзірки» [13]. У таких умовах портфельне рішення не повинно спиратися лише на одне прогнозне число (наприклад, середнє) [1], оскільки саме хвіст розподілу визначає ризик небажаних фінансових наслідків, а у фінансовому менеджменті це прямо пов'язано з управлінням ризиком і невизначеністю [9].

Тому у даній роботі прогнозування організовано так, щоб отримати не точковий прогноз, а квантильний (ймовірнісний) опис касових зборів, який далі трансформується у сценарії й безпосередньо використовується для CVaR-оптимізації портфеля [27, 28].

### 2.2.1 Постановка задачі прогнозування та вихідні величини

Нехай для кожного фільму  $i$  відомий вектор ознак  $x_i$  (жанр, сезон, бюджетні та маркетингові індикатори, характеристики акторського складу тощо), а також фактичні касові збори  $S_i$ . Задача прогнозування полягає у побудові функції

$$\hat{S}_i = f(x_i). \quad (2.1)$$

Проте, для ризик-орієнтованого портфельного відбору цього недостатньо. Потрібно оцінити умовний розподіл  $S_i | x_i$ , принаймні в ключових точках (квантилях).

Для цього вводиться квантильна функція:

$$Q_S(\tau | x_i) = \inf \{s: \mathbb{P}(S_i \leq s | x_i) \geq \tau\}, \tau \in (0,1). \quad (2.2)$$

У моделі прогноуються значення  $\hat{q}_{i,\tau} \approx Q_S(\tau | x_i)$ .

### 2.2.2 Дворівнева схема «класифікація – регресія» як основа стійкого прогнозу

У роботі [1] доведено, що попередня класифікація суттєво підвищує якість регресійного прогнозу, а відмова від класифікації призводить до різкого падіння точності (зокрема за  $R^2$ ). З огляду на це зберігається дворівнева схема:

1. Класифікація фільмів на класи  $c \in \{1, \dots, C\}$  за подібністю закономірностей формування зборів (класи з різною природою ризику та «масштабом» доходів).

2. Квантильна регресія у межах кожного класу: для фільму з класом  $c$  прогноуються  $\hat{q}_{i,\tau}^{(c)}$ .

Такий підхід також узгоджується з ідеєю портфельного менеджменту: рішення приймається не поодинці, а в умовах обмежень і компромісу між вигодою та ризиком [12].

### 2.2.3 Квантильне прогнозування касових зборів

У дослідженні було застосовано такі алгоритми машинного навчання: Лінійна регресія [31], Гребенева регресія [20], LASSO-регресія [26], Регресійне дерево рішень [36], Регресія випадкового лісу [22, 29], Регресія на основі градієнтного бустингу [37] та XGBoost-регресія [16, 23].

Ефективність зазначених регресійних моделей машинного навчання, побудованих на основі зібраного набору даних, оцінювалася за такими показниками: середня абсолютна помилка (MAE), середньоквадратична помилка (MSE), коренева середньоквадратична помилка (RMSE), середня абсолютна відсоткова помилка (MAPE) та коефіцієнт детермінації ( $R^2$ ). Математично ці показники описуються наступними рівняннями [8, 35].

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|, \quad (2.3)$$

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2, \quad (2.4)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}, \quad (2.5)$$

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right|, \quad (2.6)$$

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y}_i)^2}, \quad (2.7)$$

де  $y_i$ ,  $\hat{y}_i$  та  $\bar{y}_i$  – фактичне значення  $i$ -го спостереження, прогнозоване значення  $i$ -го спостереження та середнє значення всіх фактичних спостережень відповідно;

$n$  – загальна кількість спостережень.

Для кожного рівня  $\tau$  модель навчається мінімізувати квантильну (pinball) функцію втрат:

$$\mathcal{L}_\tau(S_i, \hat{q}_{i,\tau}) = \begin{cases} \tau(S_i - \hat{q}_{i,\tau}), S_i \geq \hat{q}_{i,\tau}, \\ (1 - \tau)(\hat{q}_{i,\tau} - S_i), S_i < \hat{q}_{i,\tau}. \end{cases} \quad (2.8)$$

Набір квантилів (щоб коректно працювати з хвостом для CVaR):

$$\tau = \{0.05, 0.10, 0.25, 0.50, 0.75, 0.90, 0.95\}. \quad (2.9)$$

Пара (0.10, 0.90) дає базовий інтервал невизначеності; додаткові 0.05/0.95 стабілізують оцінку хвостів (що критично для CVaR), а 0.25/0.75 роблять форму розподілу керованою для генерації сценаріїв.

#### 2.2.4 Калібрування квантилів і контроль монотонності

Оскільки сценарії та CVaR залежать від правильності хвоста розподілу, квантильні прогнози повинні бути калібровані. Калібрування оцінюється двома групами показників.

1. Квантильна точність. Для кожного  $\tau$  на валідаційній вибірці обчислюється:

$$\overline{\mathcal{L}}_\tau = \frac{1}{n} \sum_{i=1}^n \mathcal{L}_\tau(S_i, \hat{q}_{i,\tau}). \quad (2.10)$$

2. Покриття квантилів та інтервалів. Для квантиля  $\tau$  емпірична частота попадання:

$$\widehat{Cov}(\tau) = \frac{1}{n} \sum_{i=1}^n 1\{S_i \leq \hat{q}_{i,\tau}\}. \quad (2.11)$$

Ідеально  $\widehat{Cov}(\tau) \approx \tau$ . Для інтервалу  $[\tau_L, \tau_U]$ :

$$(\widehat{Cov}(\tau_L, \tau_U) = \frac{1}{n} \sum_{i=1}^n 1\{\hat{q}_{i,\tau_L} \leq S_i \leq \hat{q}_{i,\tau_U}\}. \quad (2.12)$$

Наприклад, для [0.10,0.90] очікуване покриття близько 0.80.

Додатково вводиться контроль перетинів квантилів, оскільки для коректної генерації сценаріїв має виконуватися:

$$\hat{q}_{i,\tau_1} \leq \hat{q}_{i,\tau_2}, \forall \tau_1 < \tau_2 \quad (2.13)$$

Якщо перетини виникають, застосовується монотонна післяобробка (наприклад, ізотонічна регресія або «перестановка» квантилів у зростаючий порядок без зміни рівнів), щоб зберегти коректність ймовірнісної інтерпретації.

### 2.2.5 Формування сценаріїв касових зборів на основі квантилів

Після калібрування квантилів формується сценарний набір можливих реалізацій касових зборів, який потрібен для CVaR.

Нехай  $S$  – кількість сценаріїв. Для кожного фільму  $i$  генерується набір випадкових чисел  $u_s \sim U(0,1)$ ,  $s = 1, \dots, S$ .

$$S_i^{(s)} = \hat{Q}_i(u_s), s = 1, \dots, S, \quad (2.14)$$

де  $\hat{Q}_i(\cdot)$  – апроксимація квантильної функції на основі значень  $\hat{q}_{i,\tau}$  для  $\tau \in \mathcal{T}$ . Практично  $\hat{Q}_i(u)$  будується як кусочно-лінійна інтерполяція між сусідніми квантилями.

Кількість сценаріїв:

- базовий рівень для стабільного CVaR:  $S = 1000$ ;
- якщо обчислювальні ресурси обмежені:  $S = 500$ ;
- якщо потрібно особливо стабільно оцінити хвіст ( $\alpha = 0.95$ ):  $S = 2000$ .

Такий сценарний підхід відповідає практиці імітаційного аналізу в задачах фінансового та портфельного планування, де ризик оцінюється на множині можливих реалізацій випадкового доходу [9], а якість узагальнення забезпечується коректною валідацією (зокрема K-fold підходами) [18].

### 2.2.6 Перехід від сценарних зборів до фінансового результату

У роботі [1] фінансовий результат портфеля описувався через різницю «дохід–витрати» та подальше оподаткування. У запропонованому варіанті ця логіка не змінюється, змінюється лише те, що дохід стає сценарним.

Для заданого портфеля  $x_i \in \{0,1\}$  прибуток до оподаткування у сценарії  $s$  визначається:

$$P^{(s)}(x) = \sum_{i=1}^N (S_i^{(s)} - C_i)x_i. \quad (2.16)$$

Далі прибуток після оподаткування обчислюється за схемою, де для кожного сценарію отримується:

$$Z_1^{(s)}(x) = TaxAdjust(P^{(s)}(x)). \quad (2.17)$$

де  $Z_1^{(s)}(x)$  – фінансовий результат портфеля після оподаткування у сценарії  $s$ .

Таким чином, формується емпіричний розподіл фінансового результату  $Z_1^{(s)}(x)$  для будь-якого кандидата-портфеля, що і є потрібною основою для CVaR-оптимізації.

Оскільки портфель формується з урахуванням ризику в найгірших сценаріях, вводиться сценарна величина втрат (ризик-орієнтована трансформація результату). Узгодженою постановкою для дистриб'ютора є трактування ризику як небажаного зниження фінансового результату відносно цільового рівня  $T$  (наприклад, мінімально прийнятний прибуток після оподаткування):

$$L^{(s)}(x) = \max \{0, T - Z_1^{(s)}(x)\}. \quad (2.18)$$

Тоді CVaR на рівні  $\alpha$  (наприклад,  $\alpha = 0.90$  або  $\alpha = 0.95$ ) оцінює середній розмір втрат у найгірших  $1 - \alpha$  частині сценаріїв і використовується як ризиковий критерій/обмеження:

$$CVaR_{\alpha}(L(x)) = \mathbb{E}[L(x) \mid L(x) \geq VaR_{\alpha}(L(x))]. \quad (2.19)$$

Таким чином, результатом даного підрозділу є калібрований імовірнісний опис касових зборів для кожного фільму у вигляді квантилів та сформованого на їх основі сценарного набору значень. Отримані сценарії дозволяють оцінювати не лише очікуваний фінансовий ефект, але й поведінку результату в несприятливих випадках, що є критично важливим для управління ризиком у портфельних рішеннях. Надалі цей сценарний опис використовується як вхід до оптимізаційної постановки, де портфель формується з урахуванням обмежень та мінімізації хвостового ризику за показником CVaR.

### 2.3 Формування портфеля за допомогою оптимізаційної моделі з урахуванням ризику

У даному підрозділі задача формування портфеля фільмів розглядається як задача математичної оптимізації за умов обмеженого бюджету та невизначеності касових зборів. На відміну від підходів, які спираються лише на точкові прогнози, зокрема і у [1], у даній роботі оптимізація виконується на сценарному описі фінансового результату, сформованому в підрозділі 2.2 на основі квантильного прогнозування та калібрування. Це дозволяє одночасно максимізувати очікуваний фінансовий ефект портфеля та контролювати ризик небажаних сценаріїв через показник CVaR, який широко застосовується в задачах ризик-менеджменту та оптимізації [9].

#### 2.3.1 Множина проєктів, змінні рішення та бюджетне обмеження

Нехай задано множину кандидатних фільмів  $i = 1, \dots, N$ . Для кожного фільму відома його вартість (витрати/інвестиції)  $C_i$  та інші параметри, які

використовуються в прогнозуванні касових зборів. Портфель описується бінарними змінними:

$$x_i = \begin{cases} 1, \text{ якщо фільм } i \text{ включено до портфеля,} \\ 0, \text{ інакше} \end{cases} \quad (2.20)$$

Нехай  $B$  – доступний бюджет. Тоді бюджетне обмеження має вигляд:

$$\sum_{i=1}^N C_i x_i \leq B. \quad (2.21)$$

За потреби модель може бути доповнена іншими технологічними обмеженнями (наприклад, верхньою межею на кількість проєктів у портфелі), однак базовим і визначальним є саме обмеження бюджету.

### 2.3.2 Сценарний прибуток до оподаткування

У роботі [1] фінансовий результат портфеля до оподаткування формувався як різниця між доходом (касовими зборами) та витратами по включених проєктах. У запропонованій моделі ця логіка залишається, але застосовується для кожного сценарію.

Сценарний прибуток портфеля до оподаткування:

$$P^{(s)}(x) = \sum_{i=1}^N (S_i^{(s)} - C_i) x_i, \quad s = 1, \dots, S. \quad (2.22)$$

Очікуваний прибуток до оподаткування оцінюється як середнє за сценаріями:

$$\mathbb{E}[P(x)] \approx \frac{1}{S} \sum_{s=1}^S P^{(s)}(x). \quad (2.23)$$

### 2.3.3 Сценарний фінансовий результат після оподаткування

Оскільки у роботі [1] фінансовий результат фіксується після оподаткування (за прийнятою схемою оподаткування), у запропонованому

підході оподаткування застосовується до кожного сценарного значення  $P^{(s)}(x)$ .

Позначимо:

$$Z_1^{(s)}(x) = TaxAdjust\left(P^{(s)}(x)\right), s = 1, \dots, S. \quad (2.24)$$

Очікуваний фінансовий результат після оподаткування:

$$\mathbb{E}[Z_1(x)] \approx \frac{1}{S} \sum_{s=1}^S Z_1^{(s)}(x). \quad (2.25)$$

### 2.3.4 Втрати та ризик: постановка для CVaR

Оскільки небажаними є сценарії з недостатнім фінансовим результатом, вводиться функція втрат як недобір до заданого порогу  $T$  (мінімально прийнятний фінансовий результат після оподаткування):

$$L^{(s)}(x) = \max\{0, T - Z_1^{(s)}(x)\}, s = 1, \dots, S. \quad (2.26)$$

Такий вибір є управлінськи інтерпретованим: якщо портфель у сценарії забезпечує  $Z_1^{(s)}(x) \geq T$ , то втрати відсутні; інакше втрати пропорційні розміру недобору.

### 2.3.5 Лінійна сценарна форма CVaR

Для практичної оптимізації використовується стандартна еквівалентна форма CVaR через змінну  $\eta$  та допоміжні змінні  $u_s$  [28, 27]:

$$CVaR_\alpha(L(x)) = \eta + \frac{1}{(1-\alpha)S} \sum_{s=1}^S u_s, \quad (2.27)$$

за умов:

$$u_s \geq L^{(s)}(x) - \eta, u_s \geq 0, s = 1, \dots, S. \quad (2.28)$$

Щоб уникнути  $\max \{\cdot\}$  у (2.26) і мати чисту лінійну постановку, вводимо змінні недобору  $d_s \geq 0$ :

$$d_s \geq T - Z_1^{(s)}(x), d_s \geq 0, s = 1, \dots, S, \quad (2.29)$$

тоді  $L^{(s)}(x)$  замінюється на  $d_s$ , а (2.28) записується як:

$$u_s \geq d_s - \eta, u_s \geq 0. \quad (2.30)$$

У результаті CVaR повністю задається лінійно через  $\eta$ ,  $u_s$ ,  $d_s$  та  $Z_1^{(s)}(x)$ .

### 2.3.6 Оптимізаційна постановка задачі формування портфеля

Далі можливі дві рівносильні за змістом постановки.

Постановка А: максимізація очікуваного результату зі штрафом за ризик

$$\max_{x, \eta, u, d} \frac{1}{S} \sum_{s=1}^S Z_1^{(s)}(x) - \lambda \left( \eta + \frac{1}{(1-\alpha)S} \sum_{s=1}^S u_s \right), \quad (2.31)$$

за умов:

$$\sum_{i=1}^N C_i x_i \leq B, \quad (2.32)$$

$$d_s \geq T - Z_1^{(s)}, d_s \geq 0, s = 1, \dots, S, \quad (2.33)$$

$$u_s \geq d_s - \eta, u_s \geq 0, s = 1, \dots, S, \quad (2.34)$$

$$x_i \in \{0,1\}, i = 1, \dots, N. \quad (2.35)$$

Параметр  $\lambda \geq 0$  задає рівень ризик-орієнтації: при великих  $\lambda$  модель сильніше уникає портфелів, які провалюються у “поганих” сценаріях.

Постановка Б: максимізація очікуваного результату за умови обмеження CVaR

$$\max_{x,\eta,u,d} \frac{1}{S} \sum_{s=1}^S Z_1^{(s)}(x), \quad (2.36)$$

за умов (2.32)–(2.35) та:

$$\eta + \frac{1}{(1-\alpha)S} \sum_{s=1}^S u_s \leq \rho, \quad (2.37)$$

де  $\rho$  – допустимий рівень хвостового ризику (максимально прийнятний середній недобір у найгірших сценаріях).

Саме поєднання каліброваного сценарного прогнозу (підрозділ 2.2) та ризик-орієнтованої оптимізації за CVaR формує методологічне ядро даної роботи та забезпечує її відмінність від підходів, що опираються лише на точкові прогнози або евристичний відбір.

## Висновки до розділу 2

1. Сформовано концепцію управління портфелем кінопроектів на основі даних, у якій процес відбору фільмів формалізовано як послідовність етапів: формування якісної інформаційної бази, побудова прогнозної моделі та підготовка даних для подальшого оптимізаційного вибору портфеля в умовах невизначеності.

2. Розроблено підхід до прогнозування касових зборів у ймовірнісній постановці: отримано квантильні оцінки, сформовано сценарії можливих реалізацій результатів та виконано їх калібрування, що забезпечує адекватне відображення невизначеності та створює необхідну основу для ризик-орієнтованого портфельного рішення.

3. Побудовано оптимізаційну модель формування портфеля з урахуванням бюджетних обмежень і ризику: фінансовий результат збережено за логікою «дохід–витрати» з подальшим оподаткуванням, але узагальнено на множину сценаріїв, а контроль небажаних сценаріїв реалізовано через критерій CVaR, що підвищує стійкість портфеля до негативних коливань ринку.

## 3 РЕЗУЛЬТАТИ ЕКСПЕРИМЕНТАЛЬНИХ ДОСЛІДЖЕНЬ ЗАПРОПОНОВАНИХ РІШЕНЬ

### 3.1 Експериментальні дані та їх підготовка

У цьому розділі наведено результати експериментальних досліджень і практичні аспекти реалізації запропонованого підходу до управління портфелем кінопроектів. Акцент зроблено на відтворюваності експериментів, аналізі даних, якості ймовірного прогнозування та демонстрації того, як сформовані сценарії використовуються для ризик-орієнтованого вибору портфеля.

Для оцінювання підходу використано відкритий табличний датасет з описом повнометражних фільмів і їхніх касових зборів. Набір містить 6027 записів і 20 змінних (додаток Б, таблиця Б.1), що охоплюють атрибути проекту (жанр, рейтинг, параметри релізу), характеристики попиту (оцінка та голоси IMDb), індикатори «відомості» творчої команди, фінансові дані (бюджет) та низку макроекономічних показників. Цільовою змінною для прогнозування є Box Office (касові збори), додатковою міткою є Class (дискретний клас).

З огляду на виражену асиметрію розподілу касових зборів у подальших експериментах застосовано логарифмічне перетворення цільової змінної ( $\log(1+x)$ ), що стабілізує масштаб і зменшує вплив одиничних надвеликих значень. Категоріальні ознаки перетворено за допомогою one-hot кодування, числові – нормалізовано на рівні моделі. Для контролю узгодженості та відсутності артефактів виконано перевірку пропусків і типів, а також базову візуальну діагностику розподілів.

Візуальний аналіз підтвердив наявність сильного «довгого хвоста» для Box Office (рисунок 3.1), а також позитивний, але не лінійний зв'язок між бюджетом і касовими зборами (рисунок 3.2). Розподіл за місяцем релізу є нерівномірним (рисунок 3.3), що відображає сезонність і релізні «вікна». Мітка Class також є незбалансованою (рисунок 3.4), тому в подальших кроках основний акцент зроблено на ймовірнісному прогнозуванні Box Office, а не на дискретній класифікації.

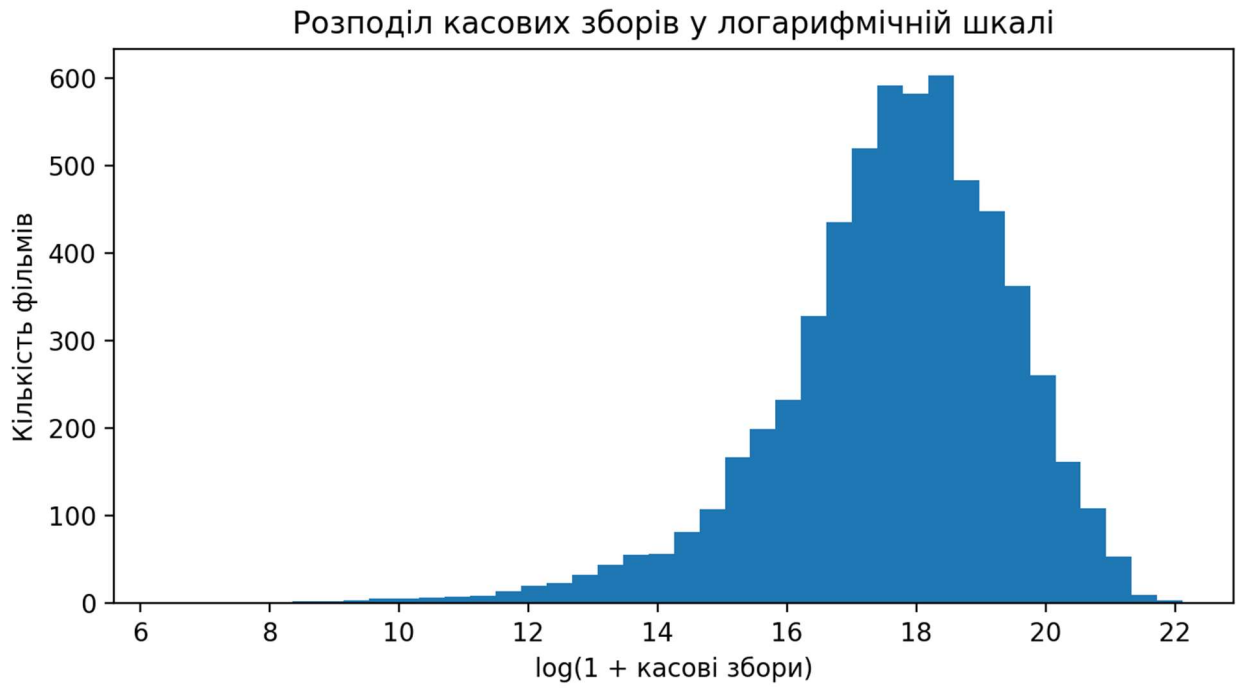


Рисунок 3.1 – Розподіл касових зборів у логарифмічній шкалі

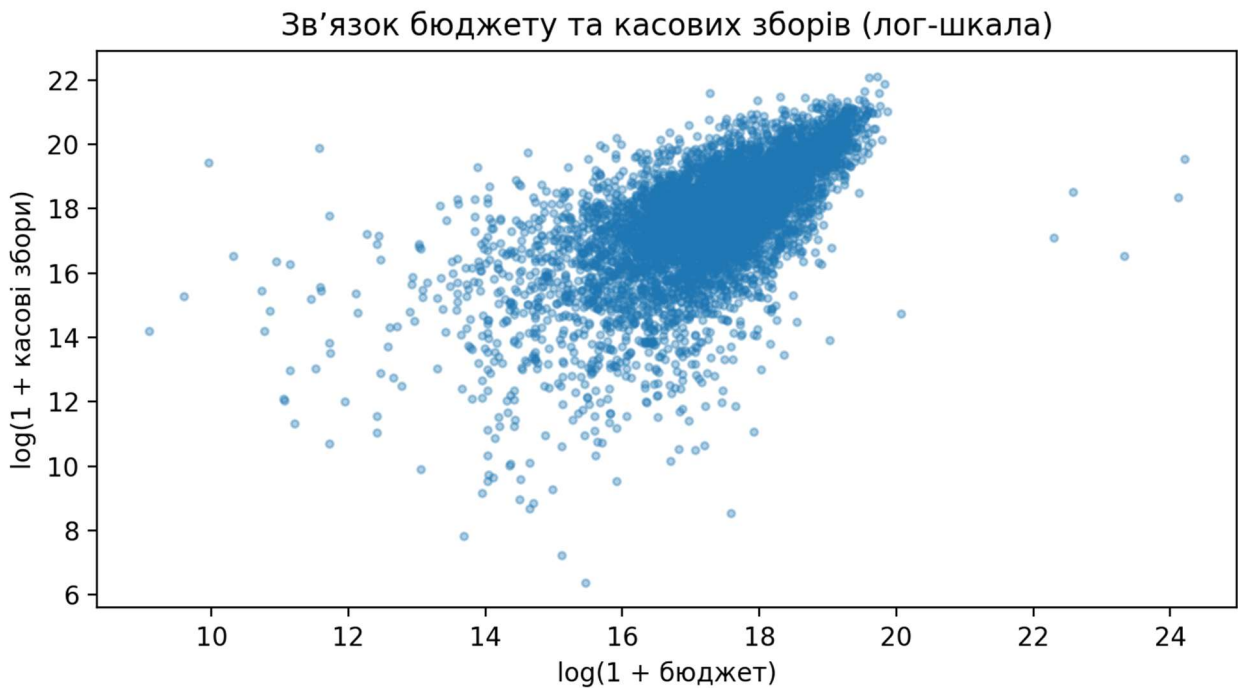


Рисунок 3.2 – Зв'язок бюджету та касових зборів (лог-шкала)

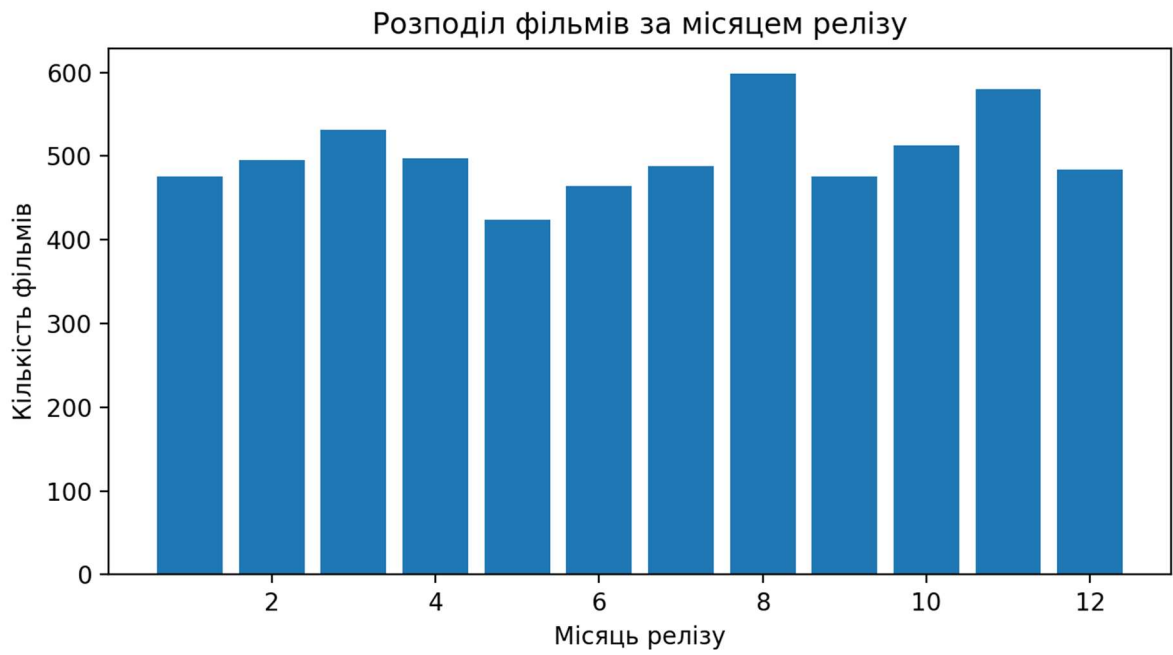


Рисунок 3.3 – Розподіл фільмів за місяцем релізу

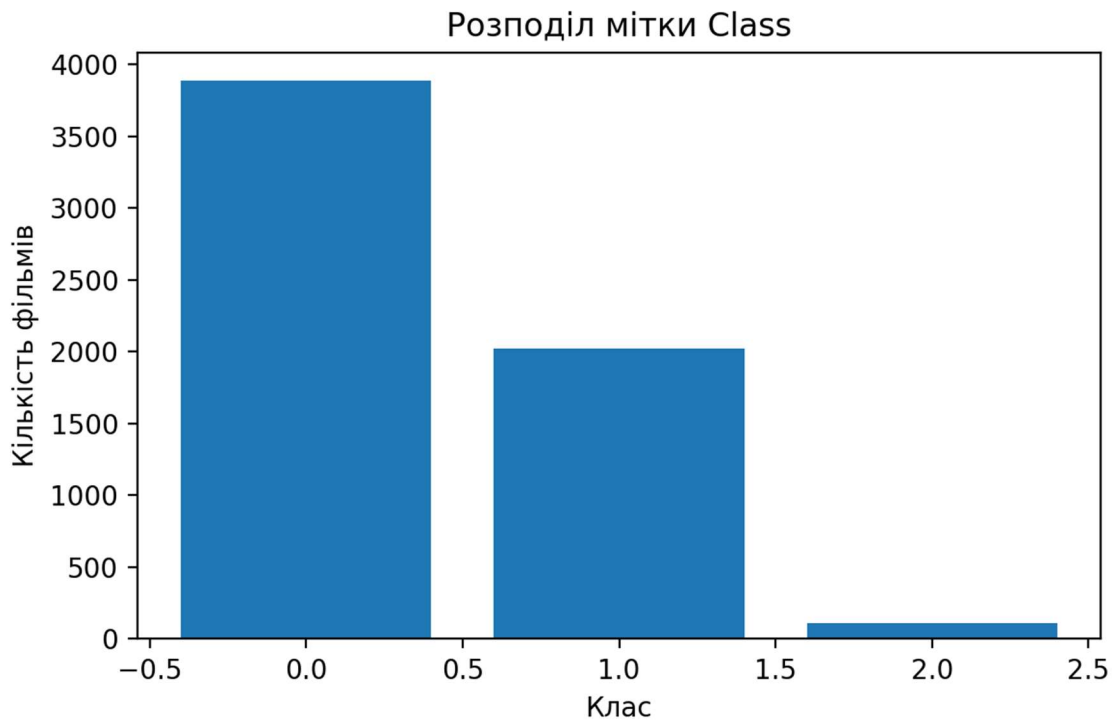


Рисунок 3.4 – Розподіл мітки Class

Статистика по ключових числових змінних представлена в таблиці 3.1.

Таблиця 3.1 – Статистика по ключових числових змінних

Показник	Середнє значення	Стандартне відхилення	Мінімальне значення	Медіана 50%	Максимальне значення
Release Year	2000.625	10.982	1980.0	2001.0	2020.0
IMDb Score	6.374	0.97	1.9	6.4	9.3
IMDb Vote Counts	105257.357	178038.143	51.0	44000.0	2400000.0
Director's Fame Score	3.992	1.96	0.0	4.0	10.0
Writer's Fame Score	3.706	2.069	0.0	3.0	10.0
Star's Fame Score	6.615	1.902	0.5	7.0	10.0
Run Time	107.866	18.076	63.0	105.0	271.0
Unemployment Rate	6.076	1.604	3.5	5.6	9.7
Inflation Rate	2.911	1.899	-0.4	2.7	13.5
GDP	11.011	5.219	2.86	10.59	21.44
Budget	67485590.2 47	604540395.0 99	8839.23 6	35502473. 85	3277250512 0.0
Box Office	145853454. 98	253531700.6	590.893	55233140. 52	4025084081. 0

### 3.2 Результати ймовірного прогнозування

На відміну від точкового прогнозу, для портфельної оптимізації критично важливо отримати опис невизначеності. Тому експерименти побудовано навколо квантильного прогнозування касових зборів: для кожного фільму оцінюються песимістичний, медіанний та оптимістичний рівні. На практиці це дозволяє

перейти від «одного числа» до інтерпретованих сценаріїв і зробити прогноз придатним для ризик-орієнтованого аналізу.

Моделі навчалися на тренувальній частині даних із виділенням окремої калібрувальної підвибірки. Калібрування застосовано для узгодження номінальних рівнів квантилів із фактичним покриттям: тобто, щоб твердження « $\tau=0.9$ » означало, що реальне значення з високою ймовірністю не перевищує прогнозований 0.9-квантиль. Для запобігання порушенню порядку квантилів використано постобробку, яка гарантує умову  $q_{10} \leq q_{50} \leq q_{90}$  (рисунок 3.5).

Результати квантильного прогнозування на тестовій вибірці подано в таблиці 3.2.

Таблиця 3.2 – Якість квантильного прогнозування на тестовій вибірці

$\tau$	Квантильна функція втрат	Емпіричне покриття
0.1	0.2	0.082
0.5	0.361	0.502
0.9	0.146	0.901

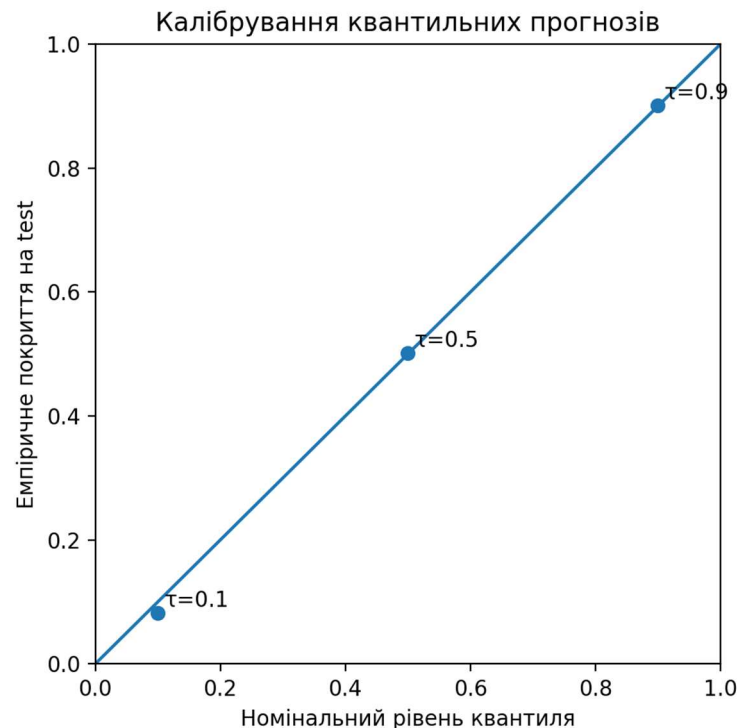


Рисунок 3.5 – Калібрування квантильних прогнозів

За результатами оцінювання найбільш стабільне покриття отримано для медіанного та верхнього квантилів ( $\tau=0.5$  і  $\tau=0.9$ ). Для нижнього квантиля ( $\tau=0.1$ ) спостерігається помірно недопокриття, що є типовим для «лівого хвоста» важко прогнозованих розподілів. Практичний висновок полягає в тому, що саме калібрування та контроль порядку квантилів є необхідними етапами, якщо прогноз використовується далі як джерело сценаріїв для ризик-орієнтованої оптимізації.

Приклад квантильного прогнозу для одного кінопроєкту подано на рисунку 3.6.

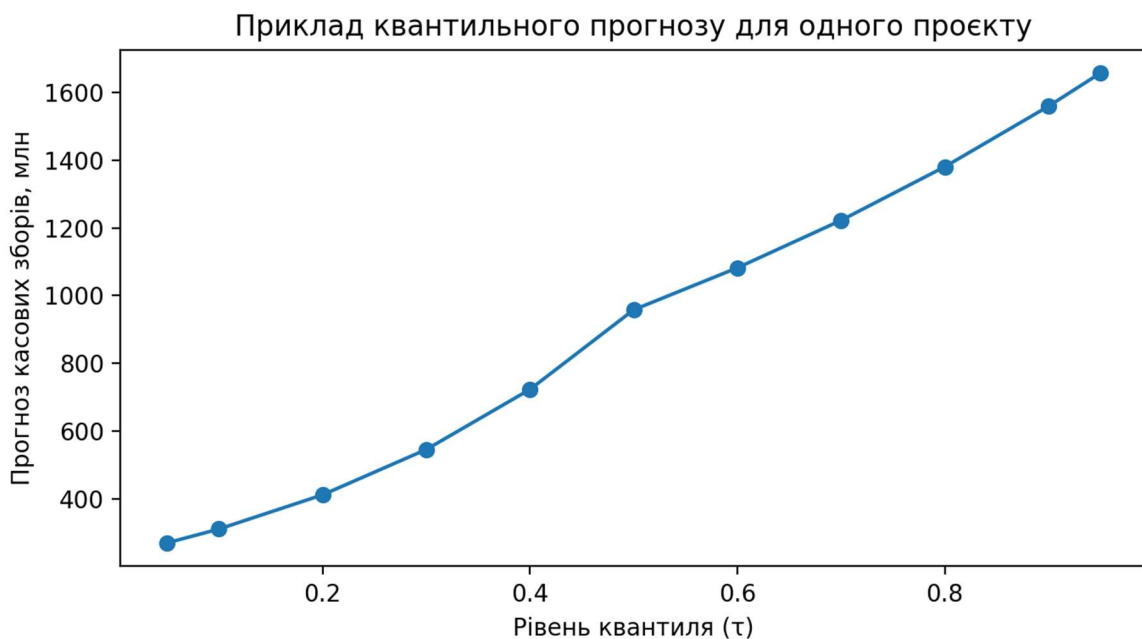


Рисунок 3.6 – Приклад квантильного прогнозу для одного кінопроєкту

Сценарний набір значень формується шляхом перетворення квантилів у дискретні «стани майбутнього» (наприклад,  $\tau=0.1$  як песимістичний сценарій,  $\tau=0.5$  як базовий,  $\tau=0.9$  як оптимістичний). У такому представленні кожен проєкт характеризується не одним очікуваним значенням, а вектором можливих результатів. Це безпосередньо «підводить» до наступного кроку – вибору портфеля, який має бути вигідним не лише в середньому, але й достатньо стійким у несприятливих сценаріях.

Сценарні значення касових зборів, сформовані з квантильного прогнозу представлено на рисунку 3.7.

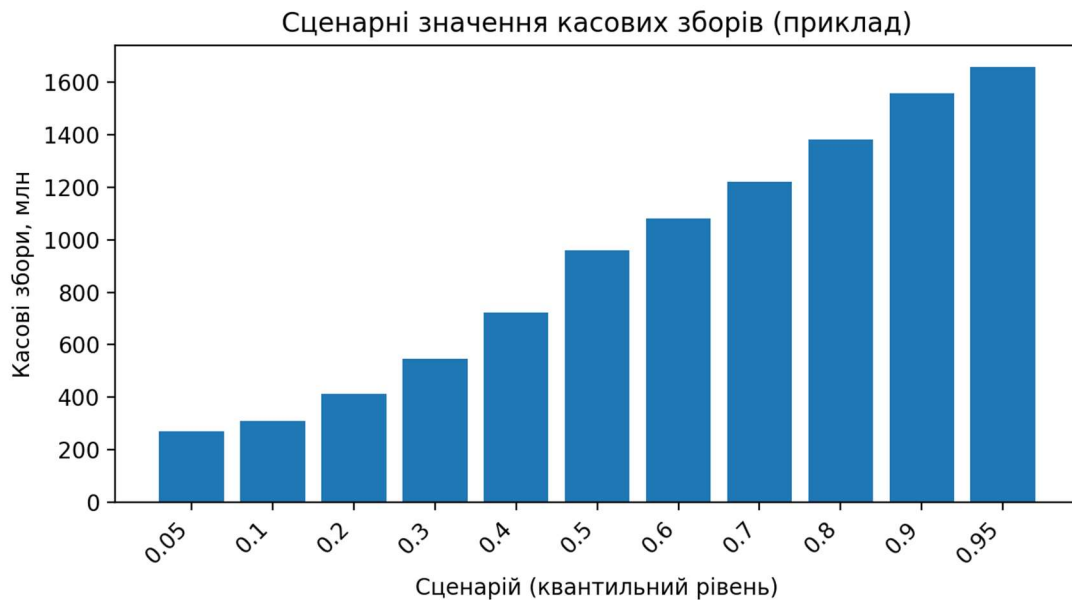


Рисунок 3.7 – Сценарні значення касових зборів, сформовані з квантильного прогнозу (приклад)

### 3.3 Результати формування портфеля на основі сценаріїв та CVaR-підходу

Після отримання сценаріїв касових зборів для кожного проєкту розраховується сценарний прибуток до оподаткування як різниця доходу та витрат: у кожному сценарії дохід відповідає прогнозованим касовим зборам, а витрати – бюджету проєкту. Далі портфель оцінюється як сукупний прибуток за всіма обраними проєктами в кожному сценарії.

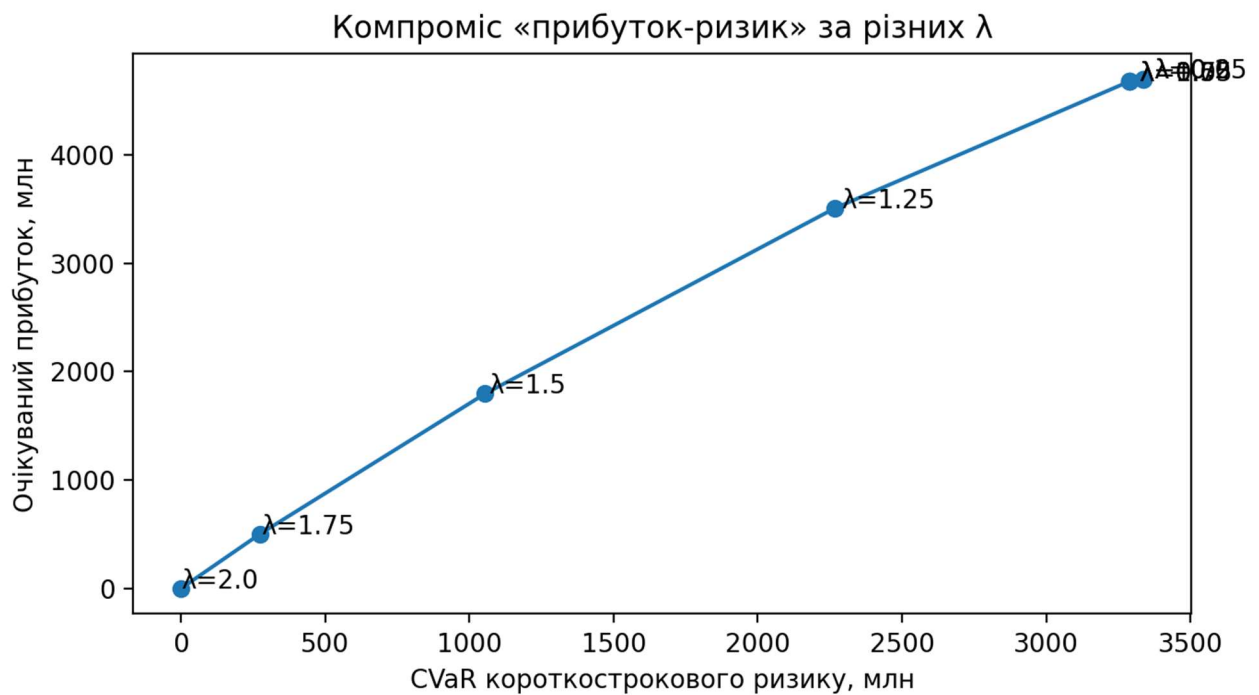
Для демонстрації ефекту ризик-орієнтації виконано експеримент на підмножині з 20 проєктів (відібраних за високим медіанним прибутком і реалістичними бюджетами). Задано обмеження загального бюджету (800 млн умовних одиниць у масштабі даних). Для різних значень параметра ризик-орієнтації  $\lambda$  сформовано портфель, що максимізує компроміс між очікуваним прибутком та показником CVaR короткострокового ризику (як середній розмір найгірших відхилень від очікуваного результату).

В таблиці 3.3 показано чутливість портфеля до параметра ризик-орієнтації  $\lambda$  (приклад на 20 проєктах).

На рисунку 3.8 відображено компроміс «очікуваний прибуток – CVaR-ризик» за різних  $\lambda$ .

Таблиця 3.3 – Чутливість портфеля до параметра ризик-орієнтації  $\lambda$   
(приклад на 20 проектах)

$\lambda$	Кількість проектів	Витрати (млн.)	Очік. прибуток (млн.)	Мін. прибуток (млн.)	CVaR короткостроковий ризик (млн.)
0.0	9.0	795.0	4699.7	1244.6	3337.2
0.25	9.0	795.0	4699.7	1244.6	3337.2
0.5	9.0	795.0	4677.9	1272.6	3287.4
0.75	9.0	795.0	4677.9	1272.6	3287.4
1.0	9.0	795.0	4677.9	1272.6	3287.4
1.25	8.0	773.4	3507.2	1143.6	2266.1
1.5	4.0	515.4	1797.7	686.3	1054.7
1.75	1.0	140.9	502.8	213.1	274.0
2.0	0.0	0.0	0.0	0.0	0.0



Рисунк 3.8 – Компромiс «очікуваний прибуток – CVaR-ризик» за різних  $\lambda$

Експеримент показує, що зі зростанням  $\lambda$  портфель поступово «стискається»: модель відмовляється від частини високодохідних, але більш нестабільних проєктів і переходить до рішень зі зменшеним сценарним розкидом (рисунок 3.9).

Таким чином, CVaR-орієнтація дозволяє керовано змінювати профіль ризику портфеля, що є принципово важливим для практики кіноіндустрії, де реальні результати часто визначаються рідкісними, але критичними несприятливими подіями.

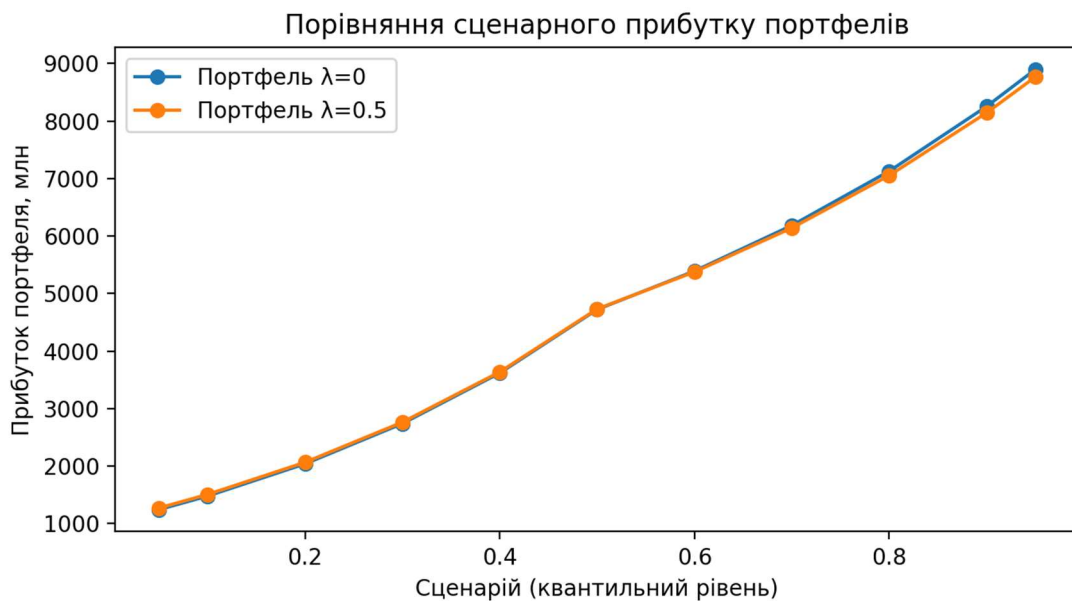


Рисунок 3.9 – Порівняння сценарного прибутку портфель ( $\lambda=0$  та  $\lambda=0.5$ )

Нижче наведено приклад складу портфель для двох режимів: без ризик-штрафу ( $\lambda=0$ ) (додаток Б, таблиця Б.2) та з помірною ризик-орієнтацією ( $\lambda=0.5$ ) (додаток Б, таблиця Б.3).

### 3.4 Програмна реалізація запропонованих рішень

Програмна реалізація запропонованих у роботі рішень побудована як відтворений конвеєр, що поєднує етапи підготовки даних, ймовірного прогнозування касових зборів, генерації сценаріїв та оптимізаційного формування портфеля з контролем ризику. Основна мета реалізації полягає в тому, щоб автоматизувати повний цикл експерименту та забезпечити можливість

багаторазових запусків із різними параметрами бюджету та ризик-орієнтації без ручного втручання в розрахунки.

У структурі програмного комплексу виділено два логічні рівні. Перший рівень реалізований мовою Python і відповідає за роботу з даними: завантаження вихідної таблиці, очищення й узгодження форматів, формування навчальної й тестової вибірок, підготовку ознак (зокрема кодування категоріальних змінних), навчання моделей ймовірнісного прогнозування та створення звітних артефактів. Результатом роботи цього рівня є набір файлів, що містять прогнози квантильні оцінки, показники калібрування, сценарні значення доходів для кожного фільму та ваги сценаріїв. Другий рівень реалізований у середовищі GAMS і використовується для постановки та розв'язання задачі портфельної оптимізації. На цей рівень надходять вже підготовлені сценарні дані та витрати, а виходом є оптимальний склад портфеля, агреговані показники ефективності та файли для порівняльного аналізу.

Ключовою особливістю реалізації є використання файлового інтерфейсу між модулями. Після етапу сценаризації Python формує стандартні артефакти: таблицю витрат проєктів, таблицю сценарних доходів, файл ваг сценаріїв та файл параметрів експерименту (наприклад, бюджет, рівень довіри, параметр ризик-орієнтації). Для інтеграції з GAMS ці дані експортуються у форматі include-файлів, що дозволяє підключати їх до оптимізаційної моделі без ручного копіювання або редагування. Такий підхід підвищує надійність, зменшує ризик помилок і робить конвеєр масштабованим.

Для забезпечення відтворюваності експериментів у реалізації фіксуються ключові параметри запуску. По-перше, задається значення випадкового зерна, яке використовується при розбитті даних та в алгоритмах, де присутні випадкові компоненти. По-друге, всі налаштування конвеєра зберігаються у конфігураційному файлі, що містить кількість сценаріїв, структуру сценарної сітки, рівень ризику та бюджет. По-третє, результати кожного запуску зберігаються у структурованій папці виводу, де окремо фіксуються таблиці прогнозів, таблиці сценаріїв, підсумкові КРІ портфеля та сформовані графіки. Це

дозволяє швидко перевіряти проміжні кроки та порівнювати результати між різними серіями експериментів.

Практична реалізація конвеєра включає окремі сценарії запуску для кожного етапу: модуль підготовки даних, модуль навчання квантильних моделей, модуль калібрування й сценаризації та модуль інтеграції з оптимізатором. Для повної автоматизації створено запускний скрипт, який послідовно виконує всі компоненти та формує фінальний набір звітів. Завдяки цьому користувач (або дослідник) може запускати експеримент “однією командою”, змінюючи лише конфігураційні параметри.

Результатом реалізації є програмний комплекс, який підтримує не лише отримання оптимального портфеля, але й аналітичний супровід управлінського рішення. Система формує підсумкові таблиці із складом портфеля, значеннями очікуваних показників та характеристиками стійкості у несприятливих сценаріях, а також створює графічні матеріали для інтерпретації компромісу між прибутковістю та ризиком.

Лістинги програмного коду подано в додатку В.

### Висновки до розділу 3

1. Виконано експериментальну перевірку запропонованого підходу на історичній вибірці кінопроектів та підтверджено, що дані кіноіндустрії характеризуються істотною асиметрією касових зборів, високою варіативністю результатів за подібних бюджетів і сезонними ефектами релізів. Це обґрунтовує необхідність переходу від точкових оцінок до ймовірнісного опису результату та сценарного аналізу як основи для портфельних рішень.

2. Отримані експериментальні результати показали, що використання ймовірнісного прогнозування з подальшим калібруванням дозволяє формувати сценарії касових зборів, які адекватно відображають невизначеність і забезпечують коректний аналіз несприятливих реалізацій. Сценарний опис доходів створює прямий зв'язок між прогнозуванням та оптимізацією й дозволяє

оцінювати портфель не за одним числом, а за поведінкою у множині можливих майбутніх ситуацій.

3. Реалізована портфельна оптимізація з контролем хвостового ризику забезпечує формування більш стійкого портфеля, зменшуючи чутливість фінансового результату до негативних сценаріїв, при збереженні прийняттого рівня очікуваної ефективності. Додаткові експерименти з аналізу чутливості підтвердили наявність керованого компромісу між прибутковістю та ризиком і продемонстрували можливість налаштування портфеля відповідно до стратегічних уподобань інвестора.

3. Програмна реалізація запропонованих рішень побудована як відтворюваний конвеєр, що автоматизує повний цикл експерименту та забезпечує прозорий обмін даними між Python-модулем і оптимізаційною моделлю в GAMS. Це підтверджує практичну придатність підходу та можливість його застосування як основи системи підтримки прийняття рішень для управління портфелем кінопроектів.

## ВИСНОВКИ

В результаті написання кваліфікаційної роботи отримано наступні результати:

1. Виконано аналіз предметної області управління портфелем кінопроектів та узагальнено ключові фактори, що впливають на комерційний успіх фільмів і прийняття інвестиційних рішень у кіноіндустрії. Показано, що ринок характеризується високою невизначеністю попиту, значною конкуренцією за увагу аудиторії та екстремальною асиметричністю касових зборів, унаслідок чого навіть наявність детальної інформації про проєкт не гарантує точного апріорного прогнозування результатів.

2. Проаналізовано існуючі підходи до оцінювання та відбору кінопроектів: класичні фінансові моделі, експертно-скорингові методи, методи машинного навчання для прогнозування касових зборів, а також оптимізаційні моделі формування портфеля. Встановлено, що традиційні фінансові підходи обмежені детермінованістю та слабкою здатністю враховувати нестабільність ринку, тоді як експертні методи мають проблему суб'єктивності та низької відтворюваності оцінок.

3. Детально розглянуто сучасну data-driven методологію, яка поєднує прогнозування, ранжування альтернатив і оптимізаційне формування портфеля. Показано її практичну цінність, зокрема використання LLM для формалізації чинника “зірковості” та застосування комбінованих ML-моделей. Водночас обґрунтовано ключові обмеження підходу: детермінованість фінальної оптимізації, відсутність імовірнісного опису невизначеності, чутливість до експертно заданих порогів класифікації, а також відсутність прямого механізму контролю хвостових ризиків.

4. На основі виявлених проблем сформульовано постановку задачі дослідження: розробити інтегрований підхід до управління портфелем кінопроектів на основі даних, яка передбачає ймовірнісне прогнозування касових зборів, формування сценарного простору можливих реалізацій і побудову ризик-орієнтованої оптимізаційної моделі формування портфеля з контролем

хвостових втрат при збереженні фінансової логіки “дохід–витрати–оподаткування” у сценарному вигляді.

5. Сформовано концепцію управління портфелем кінопроектів на основі даних, у якій процес відбору фільмів формалізовано як послідовність етапів: формування якісної інформаційної бази, побудова прогнозної моделі та підготовка даних для подальшого оптимізаційного вибору портфеля в умовах невизначеності.

6. Розроблено підхід до прогнозування касових зборів у ймовірнісній постановці: отримано квантильні оцінки, сформовано сценарії можливих реалізацій результатів та виконано їх калібрування, що забезпечує адекватне відображення невизначеності та створює необхідну основу для ризик-орієнтованого портфельного рішення.

7. Побудовано оптимізаційну модель формування портфеля з урахуванням бюджетних обмежень і ризику: фінансовий результат збережено за логікою «дохід–витрати» з подальшим оподаткуванням, але узагальнено на множині сценаріїв, а контроль небажаних сценаріїв реалізовано через критерій CVaR, що підвищує стійкість портфеля до негативних коливань ринку.

8. Виконано експериментальну перевірку запропонованого підходу на історичній вибірці кінопроектів та підтверджено, що дані кіноіндустрії характеризуються істотною асиметрією касових зборів, високою варіативністю результатів за подібних бюджетів і сезонними ефектами релізів. Це обґрунтовує необхідність переходу від точкових оцінок до ймовірнісного опису результату та сценарного аналізу як основи для портфельних рішень.

9. Отримані експериментальні результати показали, що використання ймовірнісного прогнозування з подальшим калібруванням дозволяє формувати сценарії касових зборів, які адекватно відображають невизначеність і забезпечують коректний аналіз несприятливих реалізацій. Сценарний опис доходів створює прямий зв'язок між прогнозуванням та оптимізацією й дозволяє оцінювати портфель не за одним числом, а за поведінкою у множині можливих майбутніх ситуацій.

10. Реалізована портфельна оптимізація з контролем хвостового ризику забезпечує формування більш стійкого портфеля, зменшуючи чутливість фінансового результату до негативних сценаріїв, при збереженні прийняттого рівня очікуваної ефективності. Додаткові експерименти з аналізу чутливості підтвердили наявність керованого компромісу між прибутковістю та ризиком і продемонстрували можливість налаштування портфеля відповідно до стратегічних уподобань інвестора.

11. Програмна реалізація запропонованих рішень побудована як відтворюваний конвеєр, що автоматизує повний цикл експерименту та забезпечує прозорий обмін даними між Python-модулем і оптимізаційною моделлю в GAMS. Це підтверджує практичну придатність підходу та можливість його застосування як основи системи підтримки прийняття рішень для управління портфелем кінопроектів.

## СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Alipour-Vaezi, M., Tsui, K.-L. Data-driven portfolio management for motion pictures industry: A new data-driven optimization methodology using a large language model as the expert. *Computers & Industrial Engineering*. 2024. Vol. 197. Article 110574.
2. Alipour-Vaezi, M., Tavakkoli-Moghaddam, R., Mohammadnazari, Z. Optimization of a television advertisement scheduling problem by multi-criteria decision making and dispatching rules. *Multimedia Tools and Applications*. 2022. Vol. 81(8). Pp. 11755–11772.
3. Archer, N. P., Ghasemzadeh, F. An integrated framework for project portfolio selection. *International Journal of Project Management*. 1999. Vol. 17(4). Pp. 207-216.
4. Baber, H., Fanea-Ivanovici, M. What drives people to crowdfund movies and web series? The mediating role of perceived trust and risk. *Economic Computation & Economic Cybernetics Studies & Research*. 2021. Vol. 55(3).
5. Barbas, S. *Movie Crazy: Stars, Fans, and the Cult of Celebrity*. Springer. 2016.
6. Batista, G. E., Prati, R. C., Monard, M. C. A study of the behavior of several methods for balancing machine learning training data. *ACM SIGKDD Explorations Newsletter*. 2004. Vol. 6(1). Pp. 20–29.
7. Brealey, R. A., Myers, S. C., Allen, F. *Principles of Corporate Finance* (13th ed.). McGraw-Hill Education. 2020.
8. Chicco, D., Warrens, M. J., Jurman, G. The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation. *PeerJ Computer Science*. 2021. Vol. 7. e623.
9. Chirici, G., Scotti, R., Montaghi, A., Barbati, A., Cartisano, R., Lopez, G., et al. Stochastic gradient boosting classification trees for forest fuel types mapping through airborne laser scanning and IRS LISS-III imagery. *International Journal of Applied Earth Observation and Geoinformation*. 2013. Vol. 25. Pp. 87–97.
10. Cooper, R. G., Edgett, S. J., Kleinschmidt, E. J. *Portfolio Management for*

New Products (2nd ed.). Basic Books. 2001.

11. De Vany, A. S., Walls, W. D. Motion picture profit, the stable Paretian hypothesis, and the curse of the superstar. *Journal of Economic Dynamics and Control*. 2004. Vol. 28(6). Pp. 1035–1057.

12. Ehrgott, M. *Multicriteria Optimization*. Springer. 2005.

13. Elton, E. J., Gruber, M. J., Brown, S. J., Goetzmann, W. N. *Modern Portfolio Theory and Investment Analysis*. John Wiley & Sons. 2009.

14. Fleck, N., Korchia, M., Le Roy, I. Celebrities in advertising: Looking for congruence or likability? *Psychology & Marketing*. 2012. Vol. 29(9). Pp. 651–662.

15. Fushiki, T. Estimation of prediction error by using K-fold cross-validation. *Statistics and Computing*. 2011. Vol. 21. Pp. 137–146.

16. Kjeldsberg, F., Munim, Z. H. Automated machine learning driven model for predicting platform supply vessel freight market. *Computers & Industrial Engineering*. 2024. Vol. 191. Article 110153.

17. Liu, Z., Zhang, Y., Abedin, M. Z., Wang, J., Yang, H., Gao, Y., Chen, Y. Profit-driven fusion framework based on bagging and boosting classifiers for potential purchaser prediction. *Journal of Retailing and Consumer Services*. 2024. Vol. 79. Article 103854.

18. Luo, L., Chen, X., Han, J., Park, C. W. Dilution and enhancement of celebrity brands through sequential movie releases. *Journal of Marketing Research*. 2010. Vol. 47(6). Pp. 1114–1128.

19. McClintock, P. 2019 Global Box Office Revenue Hit Record \$42.5B Despite 4 Percent Dip in U.S. *The Hollywood Reporter*. 2020.

20. McDonald, G. C. Ridge regression. *Wiley Interdisciplinary Reviews: Computational Statistics*. 2009. Vol. 1(1). Pp. 93–100.

21. Nurimbetov, R., Metyakubov, A. Advanced housing fund management system as a tool for improving delivery of municipal services on client satisfaction. *PalArch's Journal of Archaeology of Egypt/Egyptology*. 2020. Vol. 17(6). Pp. 3177–3188.

22. Ozen, F. Random forest regression for prediction of Covid-19 daily cases and deaths in Turkey. *Heliyon*. 2024. Vol. 10(4).

23. Panchal, P., Shingala, B., Thakor, S., Jain, P., Vaja, C. R., Joshi, A., et al. XGBoost regression analysis of dielectric properties of epoxy resin with inorganic hybrid nanofillers. *Journal of Macromolecular Science, Part B*. 2024. Pp. 1–17.
24. Peterson, L. E. K-nearest neighbor. *Scholarpedia*. 2009. Vol. 4(2). 1883.
25. Project Management Institute. *The Standard for Portfolio Management* (4th ed.). Project Management Institute. 2017.
26. Ranstam, J., Cook, J. LASSO regression. *British Journal of Surgery*. 2018. Vol. 105(10). Pp. 1348.
27. Rockafellar R. T., Uryasev S. Conditional Value-at-Risk for General Loss Distributions. *Journal of Banking & Finance*. 2002. Vol. 26(7). Pp. 1443–1471.
28. Rockafellar R. T., Uryasev S. Optimization of Conditional Value-at-Risk. *The Journal of Risk*. 2000. Vol. 2(3). Pp. 21–41.
29. Rodriguez-Galiano, V., Sanchez-Castillo, M., Chica-Olmo, M., Chica-Rivas, M. Machine learning predictive models for mineral prospectivity: An evaluation of neural networks, random forest, regression trees and support vector machines. *Ore Geology Reviews*. 2015. Vol. 71. Pp. 804–818.
30. Schein, A. I., Ungar, L. H. Active learning for logistic regression: An evaluation. *Machine Learning*. 2007. Vol. 68. Pp. 235–265.
31. Su, X., Yan, X., Tsai, C. L. Linear regression. *Wiley Interdisciplinary Reviews: Computational Statistics*. 2012. Vol. 4(3). Pp. 275–294.
32. Taha, A. A., Malebary, S. J. An intelligent approach to credit card fraud detection using an optimized light gradient boosting machine. *IEEE Access*. 2020. Vol. 8. Pp. 25579–25587.
33. Vellaichamy, V., Kalimuthu, V. Hybrid collaborative movie recommender system using clustering and bat optimization. *International Journal of Intelligent Engineering and Systems*. 2017. Vol. 10(1). Pp. 38–47.
34. Wilcock, J. A review of expert systems: their shortcomings and possible applications in archaeology. *Computer Applications in Archaeology*. 1985.
35. Willmott, C. J., Matsuura, K. Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. *Climate Research*. 2005. Vol. 30(1). Pp. 79–82.

36. Xu, M., Watanachaturaporn, P., Varshney, P. K., Arora, M. K. Decision tree regression for soft classification of remote sensing data. *Remote Sensing of Environment*. 2005. Vol. 97(3). Pp. 322–336.

37. Yang, F., Wang, D., Xu, F., Huang, Z., Tsui, K.-L. Lifespan prediction of lithium-ion batteries based on various extracted features and gradient boosting regression tree model. *Journal of Power Sources*. 2020. Vol. 476. Article 228654.

38. Савка А.П. Управління портфелем проєктів з використанням засобів штучного інтелекту. Збірник тез доповідей II Всеукраїнської науково-практичної конференції студентів, аспірантів та молодих вчених «Інтелектуальні комп'ютерні системи та мережі» (ІКСМ осінь 2025), м. Тернопіль, ЗУНУ, 20 травня 2025 р. Тернопіль, 2025. С. 17-20.

39. Липа А., Савка А. Методи машинного навчання та інтелектуального аналізу даних для прогнозування ризиків та управління портфелем проєктів. *Proceedings of the 2nd International Scientific and Practical Conference*. November 26-28, 2025. С. 305–309.

40. Островерхов В.М., Біловус Л.І., Возьний К.З., Луцишин О.О., Монастирський Г.Л., Надвичний С.А., Питель С.В., Шандрук С.К. Загальні методичні рекомендації з підготовки, оформлення, захисту та оцінювання кваліфікаційних робіт здобувачів вищої освіти першого (бакалаврського) і другого (магістерського) рівнів / Укладачі: Тернопіль: ЗУНУ, 2024. 83 с.

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

Додаток А  
Копії публікацій

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



**К**омп'ютерна  
**І**нженерія



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

**ІКСМ  
ОСІНЬ 2025**

**25 ЛИСТОПАДА 2025**



[KI.WUNU.EDU.UA/CONFERENCE/](http://KI.WUNU.EDU.UA/CONFERENCE/)

**ТЕРНОПІЛЬ  
2025**



## ПРОГРАМНИЙ КОМІТЕТ КОНФЕРЕНЦІЇ

Десятнюк О.М., ректор Західноукраїнського національного університету,  
д-р економічних наук, професор;

Дивак М.П., д-р технічних наук, професор, проректор з наукової роботи  
ЗУНУ;

Березький О.М., д-р технічних наук, професор, професор кафедри  
комп'ютерної інженерії Західноукраїнський національний університет;

Семанюк В.З., д-р економічних наук, професор, начальник науково-  
дослідної частини Західноукраїнського національного університету

Антощук С.Г., д.т.н, професор, Національний університет «Одеська  
політехніка»;

Баловсяк С.В., д.т.н, професор, Чернівецький національний університет

Бармак О.В., д.т.н., професор, Хмельницький національний університет

Батько Ю.М., к.т.н., доцент, Західноукраїнський національний університет;

Винокурова О.А., д.т.н., професор, Львівський національний університет  
імені Івана Франка

Возна Н.Я., д.т.н., професор, Західноукраїнський національний  
університет;

Говорущенко Т.О., д.т.н., професор, Хмельницький національний  
університет;

Дубчак Л.О., к.т.н., доцент, Західноукраїнський національний університет;

Дунець Р.Б., д.т.н., професор, НУ "Львівська політехніка";

Ізонін І.В., д.т.н., доцент, НУ "Львівська політехніка";

Комар М.П., д.т.н, професор, Західноукраїнський національний  
університет;

Литвиненко В.І., д.н.т, професор, Херсонський національний технічний  
університет ;

Лупенко С.А., д.т.н., професор, Опольський технологічний університет,  
Польща;

Ляцинський П.Б., доктор філософії з комп'ютерних наук, НУ "Львівська  
політехніка" ;

Мельник Г.М., к.т.н, доцент, Західноукраїнський національний  
університет;

Мельникова Н.І., д.т.н., професор, НУ "Львівська політехніка" ;

Пелешко Д.Д., д.т.н., професор, Львівський національний університет  
імені Івана Франка;

Піцун О.Й., к.т.н. доцент, Західноукраїнський національний університет;

Сельський П.Р., д.м.н., професор, Тернопільський національний  
медичний університет імені І. Я. Горбачевського ;

Субботін С.О., д.т.н., професор, Національний університет «Запорізька  
політехніка» ;

Теслюк В.М., д.т.н., професор, НУ "Львівська політехніка" ;

Тимченко Л.І., д.т.н., професор, Державний університет інфраструктури та технологій;  
 Цмоць І.Г., д.т.н., професор, НУ "Львівська політехніка" ;  
 Якименко І.З., к.т.н., доцент, Західноукраїнський національний університет;  
 Яровий А.А., д.т.н., професор, Вінницький національний технічний університет ;  
 Яцків В.В., д.т.н., професор, Західноукраїнський національний університет.

### **ОРГАНІЗАЦІЙНИЙ КОМІТЕТ КОНФЕРЕНЦІЇ**

Мельник Г.М. к.т.н., доцент, Західноукраїнський національний університет  
 Піцун О.Й. к.т.н., доцент, Західноукраїнський національний університет  
 Фаєрчук В.В. студент, Західноукраїнський національний університет  
 Галунька Б.В. студент, Західноукраїнський національний університет  
 Зінкевич О. В. студентка, Західноукраїнський національний університет  
 Кіт М. О. студентка, Західноукраїнський національний університет

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

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

III Всеукраїнська науково-практична конференція студентів, аспірантів та молодих вчених «Інтелектуальні комп'ютерні системи та мережі» (ІКСМ осінь 2025), м. Тернопіль, ЗУНУ, 25 листопада 2025 р. Тернопіль, 2025

**Адреса:**  
**Вул. О. Теліги, 8, корпус №6, Тернопіль**  
 URL: <https://ki.wunu.edu.ua/conference/>  
 e-mail: [kistudconference@gmail.com](mailto:kistudconference@gmail.com)

**Тези доповідей подаються за оригіналом рукопису**

## ЗМІСТ

<i>Березька К. М., Цимбалюк Л. В.</i> Цифрові засоби формування логічного мислення у процесі підготовки до ТЗНК .....	9
<i>Ковтуненко А.Р.</i> Мультимодальна висхідна сегментація об'єктів за текстовим запитом .....	11
<i>Андрухів Б.І., Воротній В.А.</i> Сучасні технології створення програмних засобів генерування звуків природніх мов .....	13
<i>Квітень Д.О.</i> Алгоритми класифікації режимів енергоспоживання для зниження пікових навантажень в розумному будинку .....	15
<i>Савка А.П.</i> Управління портфелем проєктів з використанням засобів штучного інтелекту .....	17
<i>Лука А.В.</i> Методи машинного навчання для прогнозування та управління ризиками в інфраструктурних проєктах .....	21
<i>Мороз Ю.П.</i> Нейромережева модель глибокого навчання для класифікації мережевих пакетів .....	24
<i>Шайнюк В.О.</i> Прогнозування транспортних потоків за допомогою Інтернету речей та машинного навчання .....	27
<i>Дзядик Б.-Д.Ю.</i> Інтеграція блокчейн-технології та штучного інтелекту для аналізу великих даних у середовищі Інтернету речей .....	30
<i>Сичов Р.С.</i> Модель машинного навчання для аналізу та прогнозування якості в процесах інтелектуального виробництва .....	33
<i>Каравець Р.О.</i> Аналіз настроїв в соціальних мережах на основі технологій великих даних .....	37
<i>Галин В.А.</i> Методи динамічного та статичного виявлення аномалій у великих даних .....	39
<i>Горяча І.В.</i> Автоматизований підхід до огляду літератури з використанням великих мовних моделей .....	43
<i>Киричук Д.О.</i> Дослідження ефективності застосування Slicing Aided Hyper Inference для виявлення малих об'єктів на зображеннях високої роздільної здатності .....	45
<i>Гуда Ю.Ю.</i> Застосування методів машинного навчання для прогнозування запахів на основі молекулярної структури .....	48
<i>Загрійчук В. І.</i> Аналіз способів автоматизації ділової комунікації в організаціях .....	50
<i>Панасюк Н.Р.</i> Метод та засоби відлагодження програмного забезпечення для інтелектуальних давачів наземної мобільної робототехнічної платформи .....	52
<i>Чайківська І.Р.</i> Модель та засоби оцінки дизайну ІТ-продуктів .....	55

Савка А.П.  
 магістрант 2 курсу ФКІТ ЗУНУ  
 Науковий керівник к.т.н., доцент Биковий П.Є., кафедра ІОСУ ЗУНУ

## УПРАВЛІННЯ ПОРТФЕЛЕМ ПРОЄКТІВ З ВИКОРИСТАННЯМ ЗАСОБІВ ШТУЧНОГО ІНТЕЛЕКТУ

Вступ. Складність екосистеми кіноіндустрії [1] створює значні виклики для ефективного портфельного менеджменту, який є вирішальним для забезпечення стабільної прибутковості та зростання. Портфельний менеджмент охоплює процеси вибору та управління сукупністю фінансових активів задля досягнення конкретних інвестиційних цілей [2]. Для дистриб'юторів кінопродукції це означає необхідність приймати обґрунтовані рішення щодо інвестування в ті чи інші кінопроекти, збалансовуючи потенційні прибутки та різноманітні ризики. Точні прогнози касових зборів є фундаментом для таких рішень, проте традиційні методи, що покладаються на експертні оцінки, часто нездатні врахувати неоднозначний вплив знаменитостей на успіх фільму.

Крім того, дані в кіноіндустрії мають асиметричний характер: незначна кількість фільмів генерує надвисокі доходи, тоді як більшість залишається неприбутковою, що суттєво ускладнює прогнозування. Для успішної роботи в такому динамічному середовищі необхідно застосовувати передові методології, здатні комплексно розв'язувати ці проблеми.

Чимало дослідників розглядали управлінські проблеми в медійному ланцюгу постачання, включно з кіноіндустрією, з метою оптимізації різноманітних процесів, як-от планування рекламних кампаній, прогнозування відвідуваності кінотеатрів та розробка систем рекомендацій фільмів на основі вподобань аудиторії [3]. Однак однією з ключових проблем, якій не приділялося належної уваги в попередніх наукових дослідженнях, є оптимізація інвестиційного портфеля дистриб'юторів кінопродукції.

З огляду на високу конкуренцію та динамічність кіноіндустрії, впровадження ефективних стратегій портфельного менеджменту є не просто бажаною, а життєво необхідною умовою для дистриб'юторів, що прагнуть максимізувати прибутки та водночас мінімізувати ризики. Це вимагає не лише фінансових розрахунків, а й глибокого розуміння та прогнозування глядацьких уподобань, на які дедалі більше впливають такі чинники, як цифрові технології, зміни у споживчій поведінці та зовнішні події, наприклад пандемія COVID-19 [4].

**Постановка задачі.** Проведений комплексний аналіз предметної області та існуючих методів управління портфелем проєктів виявив низку фундаментальних проблем, які перешкоджають ефективному ухваленню інвестиційних рішень у кіноіндустрії. По-перше, класичні фінансові моделі виявляються неефективними через неможливість точного апіорного прогнозування грошових потоків. По-друге, якісні та скорингові методи страждають від високого ступеня суб'єктивності та упередженості експертних оцінок. По-третє, існуючі підходи не пропонують дієвого інструментарію для роботи з екстремальною асиметричністю даних, що є іманентною властивістю ринку. Нарешті, більшість моделей зосереджені на одному критерії (зазвичай, фінансовому), ігноруючи необхідність врахування багатогранних стратегічних цілей та індивідуальних уподобань інвестора. Ця сукупність невирішених питань формує значну науково-практичну прогалину та обумовлює необхідність створення нової, інтегрованої методології.

Мета дослідження полягає у підвищенні ефективності процесу формування портфеля проєктів у кіноіндустрії через інтеграцію моделей машинного навчання та багатокритеріального аналізу для прийняття оптимальних управлінських рішень, що допоможе інвестиційним компаніям упевнено та успішно орієнтуватися в цьому складному середовищі. Об'єктом дослідження є процес формування та оптимізації портфеля кінопроектів для дистриб'юторських компаній. Предметом дослідження є методи інтелектуального аналізу даних та багатокритеріальної оптимізації для управління портфелем проєктів у кіноіндустрії.

**Основний матеріал.** Ключовим завданням для дистриб'юторів є забезпечення прибутковості проєктів, у які вони інвестують. Варто зазначити, що в більшості випадків,

особливо в кіноіндустрії, прогнозування рентабельності проекту стає складним завданням через взаємозв'язок численних чинників у цій сфері. Одним із чинників, що може впливати на прибутковість кінопроекту, є участь у ньому знаменитостей (наприклад, акторів, сценаристів та режисерів). Присутність відомих особистостей суттєво впливає на рішення споживачів. Безумовно, залучення кожної знаменитості до кінопроекту може слугувати додатковим стимулом для глядачів і, як наслідок, призвести до збільшення прибутку для дистриб'юторів.

Навіть якщо прибутковість проектів гарантована, дистриб'юторам все одно доводиться вирішувати, які з них варто включити до свого портфеля. Проте рентабельність – не єдиний ключовий чинник при формуванні портфеля в кіноіндустрії. Кожен дистриб'ютор має власні вподобання, що ґрунтуються на попередньому досвіді чи особистих очікуваннях.

З огляду на описану вище проблематику, виникають такі дослідницькі питання:

- Як можна спрогнозувати прибутковість кінопроекту?
- Як виміряти вплив знаменитостей на прибутковість кінопроекту?
- Як дистриб'ютору ефективно сформуванати портфель кінопроектів?
- Як врахувати вподобання дистриб'ютора в процесі ухвалення рішень?

Щоб надати відповіді на ці питання, в роботі запропоновано підхід оптимізації на основі даних для управління портфелем дистриб'юторів кінопродукції. При цьому було враховано кілька припущень, а саме:

- невизначеність відсутня, а всі параметри є дискретними величинами;
- протягом усього дослідження бюджет дистриб'ютора та характеристики проектів залишаються незмінними;
- єдиною мотивацією дистриб'ютора для інвестування в проект є його прибутковість.

Рисунок 1 ілюструє етапи запропонованого підходу.



Рисунок 1 – Схема підходу управління портфелем на основі даних в кіноіндустрії

Для формування оптимального портфеля для дистриб'юторів кінопродукції вкрай важливо знати, який прибуток принесе кожен проект. У цьому дослідженні за допомогою найбільш відповідної моделі машинного навчання буде спрогнозовано обсяг касових зборів для кожного кінопроекту. На рисунку 2 зображено кроки, необхідні для прогнозування касових зборів.

Як і в будь-якій іншій сфері, дистриб'ютори в кіноіндустрії мають власні вподобання щодо проектів, які вони готові включити до свого портфеля. У цьому дослідженні для розрахунку коефіцієнта пріоритетності кожного проекту для дистриб'юторів використовується комбінація методів BBWM та WASPAS. На рисунку 3 представлено етапи запропонованого підходу для обчислення цього показника.



Рисунок 2 – Етапи процесу прогнозування касових зборів

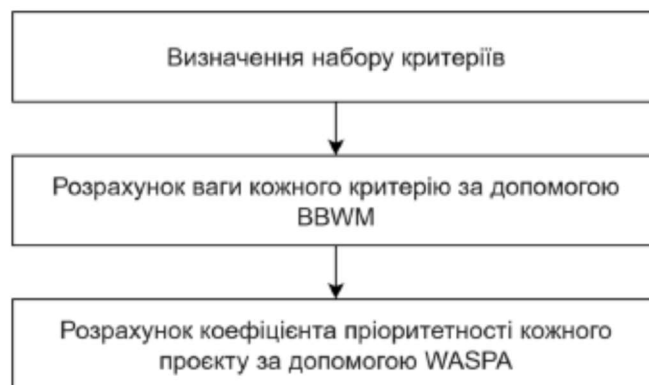


Рисунок 3 – Етапи розрахунку коефіцієнта пріоритетності кожного фільму за допомогою MADM

Спираючись на результати, отримані на попередніх етапах, а саме прогнозування касових зборів та визначення коефіцієнта пріоритетності кожного фільму для дистриб'ютора, розроблено математичну оптимізаційну задачу для формування оптимального портфеля.

$$\text{Max } Z_1 = (1 - T)R \quad (1)$$

$$\text{Max } Z_2 = \sum_{i=1}^n u_i x_i \quad (2)$$

За умов:

$$R = \sum_{i=1}^n (S_i - C_i) x_i \quad (3)$$

$$T = \sum_{b=1}^m t_b y_b \quad (4)$$

$$\sum_{b=1}^m t_b y_b = 1 \quad (5)$$

$$P_{b-1}y_b - M(1 - y_b) \leq R \leq P_b y_b + M(1 - y_b) \forall b \in \{1, 2, \dots, m\}, P_0 = 0 \quad (6)$$

$$\sum_{i=1}^n C_i x_i \leq B \quad (7)$$

$$x_i, y_j \in \{0, 1\} \forall i \in \{1, 2, \dots, n\}, b \in \{1, 2, \dots, m\} \quad (8)$$

Перша цільова функція цієї моделі полягає в максимізації сукупного прибутку дистриб'ютора. Ця мета математично виражена в рівнянні (1), де  $R$  – це прибуток компанії до оподаткування, а  $T$  – ставка податку. Друга цільова функція, проілюстрована в рівнянні (2), спрямована на максимізацію сукупної пріоритетності портфеля дистриб'ютора. Коефіцієнт пріоритетності кожного проекту ( $u_i$ ) визначається за допомогою гібридної техніки MADM.

Прибуток компанії ( $R$ ) розраховується в рівнянні (3), де  $S_i$  та  $C_i$  позначають доходи та витрати  $i$ -го кінопроекту відповідно. Тут  $x_i$  є бінарною змінною, що дорівнює 1, якщо  $i$ -й кінопроект включено до портфеля дистриб'ютора, і 0 в іншому випадку.

Модель враховує прогресивну систему оподаткування. Рівняння (4) обчислює ставку податку компанії на основі категорії прибутку, до якої вона потрапляє. У цьому контексті  $t_b$  позначає ставку податку для категорії  $y_b$  – бінарна змінна, що дорівнює 1, якщо рівень прибутку потрапляє в категорію  $b$ , і 0 в іншому випадку. Рівняння (5) забезпечує, що прибуток може належати лише до однієї податкової категорії.

Обмеження (6) визначає податкову категорію компанії. У цій нерівності  $M$  – це достатньо велике число, що гарантує, що коли прибуток перебуває в діапазоні від  $P_b$  до  $P_{b-1}$ , змінна  $y_b$  набуває значення один, а в іншому випадку – нуль.

Запропонована модель обмежується нерівністю (7), яка встановлює, що сукупна вартість відібраного портфеля не повинна перевищувати наявний бюджет дистриб'ютора, де  $C_i$  – вартість проекту  $i$ .

**Висновки.** Запропоновано дворівневий підхід: на першому етапі за допомогою алгоритмів класифікації всі проекти розподіляються на три класи за потенційним рівнем доходу, а на другому – для кожного класу будується індивідуальна регресійна модель для точного прогнозування числового значення. Запропоновано підхід оцінки пріоритетності проектів, який дозволяє формалізувати та врахувати суб'єктивні стратегічні переваги дистриб'ютора. Створено двокритеріальну оптимізаційну модель для формування збалансованого портфеля кінопроектів. Розроблена модель математичного програмування дозволяє відібрати такий набір проектів, який одночасно максимізує дві цільові функції: (1) сукупний очікуваний прибуток, розрахований на основі прогнозів моделі машинного навчання, та (2) сукупний коефіцієнт пріоритетності, визначений за допомогою MADM-моделі. Модель враховує реалістичні обмеження, зокрема загальний бюджет компанії та прогресивну систему оподаткування.

#### Список літератури

1. Nurimbetov, R., Metyakubov, A. Advanced housing fund management system as a tool for improving delivery of municipal services on client satisfaction. *PalArch's Journal of Archaeology of Egypt/Egyptology*. 2020. Vol. 17(6). Pp. 3177–3188.
2. Sheikhasadi, M., Hosseinpour, A., Alipour-Vaezi, M., Aghsami, A., Rabbani, M. Portfolio design for home healthcare devices production using a new data-driven optimization methodology. *Soft Computing*. 2024. Vol. 28(7). Pp. 5765–5784.
3. Alipour-Vaezi, M., Tavakkoli-Moghaddam, R., Mohammadnazari, Z. Optimization of a television advertisement scheduling problem by multi-criteria decision making and dispatching rules. *Multimedia Tools and Applications*. 2022. Vol. 81(8). Pp. 11755–11772.
4. Sokowati, M. E. The Economic Impact of the Film Industry on the Music Industry and the Challenges of the Digital Era. *Film Economy*. 2022. Vol. 1(1). Pp. 62–80.

isu-conference.com



COLLECTION OF SCIENTIFIC PAPERS



ISSUE  
№47

2<sup>ND</sup> INTERNATIONAL SCIENTIFIC  
AND PRACTICAL CONFERENCE

**PROGRESSIVE  
APPROACHES  
IN SCIENCE  
AND ENGINEERING**

NOVEMBER 26-28, 2025  
COPENHAGEN, DENMARK





2<sup>nd</sup> International Scientific and Practical Conference  
**«Progressive Approaches in Science and  
Engineering»**

Collection of Scientific Papers

November 26-28, 2025  
Copenhagen, Denmark

UDC 001(08)

*Progressive Approaches in Science and Engineering: Collection of Scientific Papers with Proceedings of the 2<sup>nd</sup> International Scientific and Practical Conference. International Scientific Unity. November 26-28, 2025. Copenhagen, Denmark. 697 p.*

ISBN 979-8-89704-979-0 (series)  
DOI 10.70286/ISU-26.11.2025

The conference is included in the Academic Research Index ReserchBib International catalog of scientific conferences.

The collection of scientific papers presents the materials of the participants of the 2<sup>nd</sup> International Scientific and Practical Conference "Progressive Approaches in Science and Engineering" (November 26-28, 2025. Copenhagen, Denmark).

The materials of the collection are presented in the author's edition and printed in the original language. The authors of the published materials bear full responsibility for the authenticity of the given facts, proper names, geographical names, quotations, economic and statistical data, industry terminology, and other information.

The materials of the conference are publicly available under the terms of the CC BY-NC 4.0 International license.

ISBN 979-8-89704-979-0



© Participants of the conference, 2025  
© Collection of Scientific Papers "International Scientific Unity", 2025  
Official site: <https://isu-conference.com/>

<b>Mamrosh V.S.</b> IMPROVED METHODOLOGY FOR DEFECT IDENTIFICATION IN MULTIPLAYER GAMES CASE STUDY OFF THE GRID.....	301
<b>Липа А., Савка А.</b> МЕТОДИ МАШИННОГО НАВЧАННЯ ТА ІНТЕЛЕКТУАЛЬНОГО АНАЛІЗУ ДАНИХ ДЛЯ ПРОГНОЗУВАННЯ РИЗИКІВ ТА УПРАВЛІННЯ ПОРТФЕЛЕМ ПРОЄКТІВ.....	305
<b>Галин В., Аравець Р., Сичов Р.</b> ІНТЕЛЕКТУАЛЬНІ МЕТОДИ АНАЛІЗУ ВЕЛИКИХ ДАНИХ: ВИЯВЛЕННЯ АНОМАЛІЙ, АНАЛІЗ НАСТРОЇВ ТА ПРОГНОЗУВАННЯ ЯКОСТІ В ІНТЕЛЕКТУАЛЬНОМУ ВИРОБНИЦТВІ.....	310
<b>Sharovalova S., Chyzh Ye.</b> ARCHITECTURAL APPROACHES TO IMPLEMENTING A ROLE- BASED ACCESS CONTROL (RBAC) MODEL FOR MODERN WEB PLATFORMS.....	314
<b>Дзядик Б., Мороз Ю., Шайнюк В.</b> ПІДХІД ДО АНАЛІЗУ МЕРЕЖЕВОГО ТРАФІКУ ТА ПРОГНОЗУВАННЯ ТРАНСПОРТНИХ ПОТОКІВ НА ОСНОВІ ІНТЕРНЕТ РЕЧЕЙ, БЛОКЧЕЙНУ Й ГЛИБОКОГО НАВЧАННЯ.....	316
<b>Шелег Я.П.</b> ОБМЕЖЕННЯ SAST ІНСТРУМЕНТІВ ПРИ ДЕТЕКЦІЇ КОНТЕКСТНО-ЗАЛЕЖНИХ ВРАЗЛИВОСТЕЙ ТА LLM- АЛЬТЕРНАТИВА.....	321
<b>Maiko D.R., Pohorilets V.M., Maiko T.S.</b> COMPARATIVE ANALYSIS OF CONTAINERIZATION AND VIRTUALIZATION TECHNOLOGIES IN CLOUD INFORMATION SYSTEMS DEPLOYMENT.....	323
<b>Юрченко В.О.</b> ПЕРЕВІРКА КОРЕКТНОСТІ ВІДПОВІДЕЙ АГЕНТІВ ШТУЧНОГО ІНТЕЛЕКТУ.....	327
<b>Sharovalova S., Huryn I.</b> PERSONALIZED RECOMMENDATIONS BASED ON THE PROCESSING OF TEXT DATA AND USER BEHAVIOR PATTERNS..	329
<b>Кім В.</b> ІЄРАРХІЯ КЕШІВ І ПОНЯТТЯ FALSE SHARING.....	333

---

## МЕТОДИ МАШИННОГО НАВЧАННЯ ТА ІНТЕЛЕКТУАЛЬНОГО АНАЛІЗУ ДАНИХ ДЛЯ ПРОГНОЗУВАННЯ РИЗИКІВ ТА УПРАВЛІННЯ ПОРТФЕЛЕМ ПРОЄКТІВ

Липа Андрій

здобувач вищої освіти

Савка Андрій

здобувач вищої освіти

Кафедра інформаційно-обчислювальних систем і управління  
Західноукраїнський національний університет, Україна

Сучасні інфраструктурні та креативні галузі функціонують у надзвичайно складному та динамічному середовищі, де рішення щодо інвестування й управління проектами пов'язані з високим рівнем невизначеності та взаємопов'язаними ризиками. Для інфраструктурних проектів така складність зумовлена багаторівневою взаємодією зацікавлених сторін, технічними та організаційними залежностями, а також системними ефектами, що можуть призводити до каскадних збоїв, перевитрат бюджету та зривів строків [1–4].

У кіноіндустрії проблема ускладнюється асиметрією розподілу доходів: невелика кількість фільмів генерує надвисокі касові збори, тоді як більшість залишається малоприбутковою або збитковою. Дистриб'ютори змушені формувати портфель кінопроектів в умовах високої невизначеності, залежності від глядацьких уподобань, впливу знаменитостей, цифрових платформ та зовнішніх шоків (наприклад, пандемій) [2, 6, 9].

Традиційні підходи до оцінки та управління ризиками в інфраструктурних проектах часто базуються на аналізі окремих загроз і не враховують кумулятивного ефекту їх взаємодії, що призводить до систематичної недооцінки реальних небезпек [1, 5]. Аналогічно, класичні фінансові та експертні методи в портфельному менеджменті кіноіндустрії не забезпечують достатньої точності прогнозів грошових потоків і не дають інструментів для роботи з екстремальною асиметрією даних [2, 6–9].

У цих умовах методи машинного навчання та інтелектуального аналізу даних відкривають нові можливості для виявлення прихованих закономірностей у великих історичних вибірках проектів, кількісної оцінки сукупного впливу взаємодіючих ризиків, а також побудови оптимальних портфелів інвестицій з урахуванням багатьох критеріїв.

Дослідження складності інфраструктурних проектів показують, що вона пов'язана як із внутрішніми характеристиками проекту (масштаб, тривалість, кількість дисциплін, структурна складність), так і з організаційними факторами (взаємозалежність команд, зацікавлених сторін, контрактних схем) [1–3]. Для опису та вимірювання складності застосовують індикатори, отримані за допомогою експертних опитувань, Delphi-процедур та формальних методик [3,

4]. Проте більшість робіт націлені або на вивчення окремих ризиків, або на аналіз системних ефектів, без єдиної інтегрованої моделі, що враховує їх спільний вплив [5].

У сфері портфельного менеджменту традиційні фінансові моделі (наприклад, варіаційно-дисперсійний підхід Марковіца) передбачають стаціонарні та відносно передбачувані грошові потоки, що слабо відповідає реаліям кіноіндустрії. Роботи, присвячені оптимізації портфеля продукції або послуг у суміжних доменах, використовують методи багатокритеріальної оптимізації, математичного програмування та MADM для врахування різних інтересів стейкхолдерів [2, 7, 8]. Однак у більшості випадків вони не оперують специфічною асиметрією дохідності, яка є типовою для кіноіндустрії, та не інтегрують у єдину схему прогнозування доходів на основі машинного навчання й багатокритеріальний вибір.

У медіа- та креативних індустріях досліджуються задачі прогнозування відвідуваності, планування рекламних кампаній, рекомендацій фільмів тощо [3, 9]. Застосування методів машинного навчання для прогнозування касових зборів фільмів, як правило, зосереджується на побудові регресійних або класифікаційних моделей, але не завжди поєднується з подальшою оптимізацією портфеля проєктів для інвестора (дистриб'ютора).

Таким чином, існує науково-практична прогалина, пов'язана:

- з відсутністю інтегрованих моделей, які кількісно оцінюють сукупний вплив взаємодії ризиків та системних ризиків на ефективність інфраструктурних проєктів із використанням алгоритмів машинного навчання;

- з нестачею методологій портфельного управління в кіноіндустрії, що поєднують прогнозування доходів ML-моделями та багатокритеріальний аналіз з урахуванням індивідуальних вподобань дистриб'ютора.

Мета дослідження – підвищити ефективність управління ризиками та портфелем проєктів шляхом розробки та обґрунтування моделей машинного навчання й багатокритеріальної оптимізації, які здатні кількісно оцінювати сукупний вплив взаємодії ризиків та системних ризиків на ключові показники ефективності (відхилення від бюджету й графіку), а також формувати оптимальний портфель проєктів, поєднуючи прогнози касових зборів (на основі ML-моделей) з індивідуальними пріоритетами дистриб'ютора (на основі MADM-процедур).

1. Модель машинного навчання для прогнозування продуктивності інфраструктурних проєктів. Розроблено прогнозну модель, яка використовує історичні дані про виконання інфраструктурних проєктів та пов'язані з ними ризики для оцінки відхилень від бюджету та графіку. На основі кореляційного аналізу початковий набір із 31 характеристики було скорочено до 17 найбільш інформативних предикторів; для зменшення впливу різних масштабів застосовано стандартизацію ознак.

Спочатку для побудови базового рівня було використано баєсівську регресію (BR), після чого розглядалися такі алгоритми:

- DT (дерево рішень) – як інтерпретована модель із можливістю виявлення ключових розгалужень;
- RF (випадковий ліс) – ансамбль дерев, здатний зменшувати дисперсію прогнозів;
- ANN (штучна нейронна мережа) – для моделювання складних нелінійних взаємозв'язків;
- AR (адаптивне підсилювання) – ансамблевий метод, орієнтований на послідовне покращення регресійних моделей;
- SVM (метод опорних векторів) – для побудови регресійної моделі з ядровими функціями.

Для кожного алгоритму проведено 7-кратну перехресну валідацію з підбором гіперпараметрів за метриками  $R^2$ , MSE та MAE. Зокрема, для RF оптимізувалися кількість дерев та глибина, для ANN – кількість прихованих шарів (один або два), кількість нейронів (64, 128), функція активації та параметри регуляризації. Для AR параметри кількості дерев, глибини та швидкості навчання обиралися на основі аналізу кривих продуктивності на тренувальних і тестових вибірках.

Особливу увагу приділено аналізу смуг невизначеності продуктивності моделей ( $\pm 1$  стандартне відхилення). Показано, що вузькі смуги невизначеності свідчать про вищу стабільність і здатність до узагальнення. Результати демонструють, що RF забезпечує кращий компроміс між точністю, стабільністю та обчислювальною ефективністю, порівняно з AR та іншими алгоритмами, завдяки вузьким смугам невизначеності та меншій чутливості до змін гіперпараметрів.

Запропонована модель дозволяє кількісно оцінювати сукупний ефект взаємодіючих ризиків та системних ризиків на результати проєкту. Це дає можливість менеджерам проєктів проводити сценарний аналіз, виявляти комбінації факторів, що найбільше впливають на відхилення, та вчасно коригувати планування.

2. Інтелектуальна підтримка управління портфелем кінопроєктів. Розроблено підхід до управління портфелем кінопроєктів, який поєднує прогностичні ML-моделі, багатокритеріальні методи та математичне програмування. Підхід є дворівневим. На першому етапі пропонується класифікація проєктів за рівнем очікуваних доходів (наприклад, низький, середній, високий). Для цього використовуються алгоритми класифікації, які враховують широкий спектр факторів: жанр, бюджет, участь знаменитостей (акторів, режисерів, сценаристів), виробничі характеристики, попередній досвід студії та ін. На другому етапі, для кожного класу рівня доходів, будується окрема регресійна модель (наприклад, градієнтний бустинг, RF чи інші), яка уточнює числове значення очікуваних касових зборів. Такий підхід дає змогу краще врахувати асиметричний розподіл доходів та відмінності між «блокбастерами» й малобюджетними фільмами.

Окрім суто фінансових показників, кожен дистриб'ютор має власні стратегічні уподобання (жанрові, регіональні, репутаційні, пов'язані з

співпрацю з певними студіями чи творчими командами). Для формалізації цих переваг застосовується комбінація методів Best–Worst Method (BBWM) та WASPAS. На основі експертних оцінок формуються ваги критеріїв, далі розраховується коефіцієнт пріоритетності кожного проєкту, що відображає, наскільки він відповідає стратегічним цілям дистриб'ютора.

На заключному етапі пропонується двокритеріальна оптимізаційна модель формування портфеля, яка одночасно максимізує:

- сукупний очікуваний прибуток (на основі прогнозів ML-моделі);
- сукупний коефіцієнт пріоритетності (на основі MADM-методів).

Модель враховує:

– бюджетні обмеження (сума інвестицій у відібрані проєкти не перевищує заданого бюджету);

– прогресивну шкалу оподаткування прибутку (за допомогою бінарних змінних і системи обмежень задається віднесення компанії до певного податкового діапазону);

– обмеження цілісності (кожен проєкт або включено до портфеля, або ні).

В результаті формується такий портфель кінопроєктів, який, з одного боку, максимізує очікуваний прибуток, а з іншого – відповідає суб'єктивним стратегічним пріоритетам дистриб'ютора, забезпечуючи збалансований компроміс між фінансовою дохідністю й довгостроковими цілями.

**Висновки.**

Показано, що інтелектуальні методи машинного навчання є ефективним інструментом для моделювання складних, нелінійних і взаємопов'язаних ризиків в інфраструктурних проєктах. Запропонована ML-модель із використанням ансамблевих алгоритмів (зокрема RF) дозволяє кількісно оцінювати сукупний вплив взаємодії ризиків та системних ризиків на відхилення від бюджету та графіку, що підвищує якість ухвалення управлінських рішень.

Розроблена в роботі для кіноіндустрії дворівнева ML-процедура (класифікація + регресія) дозволяє точніше прогнозувати касові збори в умовах асиметричного розподілу доходів і множини впливових факторів, включаючи участь знаменитостей та структурні характеристики проєктів.

Інтеграція методів MADM (BBWM, WASPAS) із результатами ML-прогнозів забезпечує формалізований облік стратегічних пріоритетів дистриб'ютора. Це дає можливість перейти від інтуїтивного чи суто фінансового вибору до багатокритеріального портфельного управління, що враховує як очікуваний прибуток, так і суб'єктивні переваги інвестора.

Запропонована двокритеріальна оптимізаційна модель портфеля кінопроєктів демонструє, як прогнозні ML-моделі можуть бути безпосередньо інтегровані в математичне програмування, утворюючи завершений ланцюг: «дані → прогноз → пріоритети → оптимальний портфель». Аналогічні підходи можуть бути адаптовані й для інших галузей, де портфель рішень формується в умовах невизначеності та складних залежностей.

Отже, обидва розглянуті напрями – прогнозування ризиків в інфраструктурних проєктах і портфельний менеджмент у кіноіндустрії –

демонструють, що поєднання машинного навчання та методів багатокритеріальної оптимізації формує потужний інструментарій для підтримки рішень у складних проектно-орієнтованих середовищах. Подальші дослідження можуть бути спрямовані на інтеграцію таких моделей у єдині цифрові платформи підтримки управління проектами та інвестиціями.

#### Список використаних джерел

1. De Toni, F. A., & Pessot, E. (2021). Investigating organisational learning to master project complexity: An embedded case study. *Journal of Business Research*, 129, 541–554.
2. Koo, H. J., & O'Connor, J. T. (2021). Complexity analysis of design deliverable defects on building projects. *Journal of Management in Engineering*, 37(3), 04021014.
3. Kermanshachi, S., Rouhanizadeh, B., & Dao, B. (2020). Application of Delphi method in identifying, ranking, and weighting project complexity indicators for construction projects. *Journal of Legal Affairs and Dispute Resolution in Engineering and Construction*, 12(1), 04519033.
4. Project Management Institute. (2021). *A guide to the project management body of knowledge (PMBOK® guide) (7th ed.)*. Newtown Square, PA: Project Management Institute.
5. Yassien, Y., Ezzeldin, M., Mohamed, M., & El-Dakhakhni, W. (2020). Air transportation infrastructure robustness assessment for proactive systemic risk management. *Journal of Management in Engineering*, 36(4), 04020037.
6. Nurimbetov, R., & Metyakubov, A. (2020). Advanced housing fund management system as a tool for improving delivery of municipal services on client satisfaction. *PalArch's Journal of Archaeology of Egypt/Egyptology*, 17(6), 3177–3188.
7. Sheikhasadi, M., Hosseinpour, A., Alipour-Vaezi, M., Aghsami, A., & Rabbani, M. (2024). Portfolio design for home healthcare devices production using a new data-driven optimization methodology. *Soft Computing*, 28(7), 5765–5784.
8. Alipour-Vaezi, M., Tavakkoli-Moghaddam, R., & Mohammadnazari, Z. (2022). Optimization of a television advertisement scheduling problem by multi-criteria decision making and dispatching rules. *Multimedia Tools and Applications*, 81(8), 11755–11772.
9. Sokowati, M. E. (2022). The economic impact of the film industry on the music industry and the challenges of the digital era. *Film Economy*, 1(1), 62–80.

## Додаток Б

## Результати експериментальних досліджень

Таблиця Б.1 – Склад змінних експериментального набору даних

Група	Змінна	Зміст	Тип даних
Ідентифікація	ID	Ідентифікатор запису	текст/число
Опис проєкту	Title	Назва фільму	текст
Опис проєкту	Franchise or Sequel	Ознака франшизи/сиквелу (0/1)	категорія
Опис проєкту	Rating	Віковий рейтинг (категорія)	категорія
Опис проєкту	Genre	Жанр (категорія)	категорія
Параметри релізу	Release Month	Місяць виходу	категорія/число
Параметри релізу	Release Year	Рік виходу	число
Якість/попит	IMDb Score	Оцінка IMDb	число
Якість/попит	IMDb Vote Counts	Кількість голосів IMDb	число
Креативна команда	Director's Fame Score	Індекс відомості режисера	число
Креативна команда	Writer's Fame Score	Індекс відомості сценариста	число
Креативна команда	Star's Fame Score	Індекс відомості головної зірки	число
Ринок	Domesticity	Ознака внутрішнього/міжнародного прокату	категорія
Технічні параметри	Run Time	Тривалість (хв)	число
Фінанси	Budget	Бюджет (USD)	число
Макроекономіка	Unemployment	Рівень безробіття	число
Макроекономіка	Inflation	Інфляція	число
Макроекономіка	GDP	ВВП	число
Ціль	Box Office	Касові збори (USD)	цільова змінна
Допоміжне	Class	Дискретна мітка класу	цільова змінна /категорія

Таблиця Б.2 – Приклад складу портфеля при  $\lambda=0$  (медіанні оцінки)

Портфель	Назва	Жанр	Рейти нг	Місяць виходу	Бюджет (млн.)	Q50 касових зборів (млн.)	Медіанний прибуток (млн.)
$\lambda=0$	Jurassic Park	1	3	6	127.9800016	958.1953956 523656	830.2153940 523656
$\lambda=0$	The Lost World: Jurassic Park	1	3	5	133.4596794	847.8397799 722169	714.3801005 722169
$\lambda=0$	Ocean's Eleven	6	3	12	140.8740161	761.1575940 280078	620.2835779 280077
$\lambda=0$	Star Wars: Episode VI - Return of the J...	1	2	5	95.75839141	704.6371909 832776	608.8787995 732775
$\lambda=0$	The Godfather: Part III	6	4	12	121.2839030 9999999	627.0683782 190757	505.7844751 1907564
$\lambda=0$	Twilight	7	3	10	50.43136356	490.4575051 395493	440.0261415 795493
$\lambda=0$	Home Alone	5	2	10	40.4279677	438.6853691 226136	398.2574014 226136
$\lambda=0$	Dead Poets Society	5	2	6	38.82268989	346.3803163 3866024	307.5576264 4866024
$\lambda=0$	Knives Out	5	3	10	45.91485643 9999994	342.7035392 302264	296.7886827 902264

Таблиця Б.3 – Приклад складу портфеля при  $\lambda=0.5$  (медіанні оцінки)

Портфель	Назва	Жанр	Рейти нг	Місяць виходу	Бюджет (млн.)	Q50 касових зборів (млн.)	Медіанний прибуток (млн.)
$\lambda=0.5$	Jurassic Park	1	3	6	127.9800016	958.1953956 523656	830.2153940 523656
$\lambda=0.5$	The Lost World: Jurassic Park	1	3	5	133.4596794	847.8397799 722169	714.3801005 722169
$\lambda=0.5$	Ocean's Eleven	6	3	12	140.8740161	761.1575940 280078	620.2835779 280077
$\lambda=0.5$	Star Wars: Episode VI - Return of the J...	1	2	5	95.75839141	704.6371909 832776	608.8787995 732775
$\lambda=0.5$	Rise of the Planet of the Apes	1	3	8	121.3297834 0000001	636.9834976 163669	515.6537142 163669
$\lambda=0.5$	Twilight	7	3	10	50.43136356	490.4575051 395493	440.0261415 795493
$\lambda=0.5$	Home Alone	5	2	10	40.4279677	438.6853691 226136	398.2574014 226136
$\lambda=0.5$	Dead Poets Society	5	2	6	38.82268989	346.3803163 3866024	307.5576264 4866024
$\lambda=0.5$	Knives Out	5	3	10	45.91485643 9999994	342.7035392 302264	296.7886827 902264

## Додаток В

## Лістинги програмного коду

Лістинг 3.1 – Підготовка даних і навчання квантильних моделей прогнозу касових зборів

```
import numpy as np
import pandas as pd

from sklearn.model_selection import train_test_split
from sklearn.compose import ColumnTransformer
from sklearn.preprocessing import OneHotEncoder
from sklearn.pipeline import Pipeline
from sklearn.ensemble import GradientBoostingRegressor

# 1) Завантаження датасету
df = pd.read_csv("data/mmc1.csv")

# 2) Цільова змінна (касові збори)
y = df["Box Office"].astype(float)

# 3) Ознаки (ID/Title використовуються лише для звітів; Class – службова)
X = df.drop(columns=["Box Office", "ID", "Title", "Class"])

# 4) Перелік категоріальних і числових ознак (приклад для цього датасету)
cat_cols = ["Franchise or Sequel", "Rating", "Genre", "Domesticity", "Release Month"]
num_cols = [c for c in X.columns if c not in cat_cols]

# 5) Розбиття на навчальну та тестову частини (для відтворюваності фіксуємо random_state)
X_train, X_test, y_train, y_test = train_test_split(
    X, y, test_size=0.25, random_state=42
)

# 6) Препроцесинг: One-Hot для категоріальних + passthrough для числових
preprocess = ColumnTransformer(
    transformers=[
        ("cat", OneHotEncoder(handle_unknown="ignore"), cat_cols),
        ("num", "passthrough", num_cols),
    ]
)

# 7) Навчання окремої моделі для кожного квантиля  $\tau$ 
taus = [0.1, 0.5, 0.9]
models = {}

# Для стабілізації навчання працюємо з  $\log(1+y)$ 
y_train_log = np.log1p(y_train)

for tau in taus:
    gbr = GradientBoostingRegressor(
```

```

    loss="quantile",
    alpha=tau,
    n_estimators=300,
    learning_rate=0.06,
    max_depth=3,
    subsample=0.9,
    random_state=42,
)

pipe = Pipeline([("pre", preprocess), ("model", gbr)])
pipe.fit(X_train, y_train_log)
models[tau] = pipe

# 8) Отримання квантильних прогнозів для всіх об'єктів (для сценаризації)
Q10 = np.expm1(models[0.1].predict(X))
Q50 = np.expm1(models[0.5].predict(X))
Q90 = np.expm1(models[0.9].predict(X))

# 9) Забезпечення монотонності квантилів (щоб уникнути перетинів)
Q10 = np.minimum(Q10, Q50)
Q90 = np.maximum(Q90, Q50)
# 10) Збереження прогнозів для подальших кроків (калібрування/сценарії/оптимізація)
out = pd.DataFrame({
    "ID": df["ID"],
    "Title": df["Title"],
    "Budget": df["Budget"].astype(float),
    "Q10": Q10,
    "Q50": Q50,
    "Q90": Q90
})
out.to_csv("outputs/quantile_predictions.csv", index=False)

```

### Лістинг 3.2 – Калібрування квантилів і генерація сценаріїв

```

import numpy as np
import pandas as pd

pred = pd.read_csv("outputs/quantile_predictions.csv")
df = pd.read_csv("data/mmc1.csv")

# 1) Фактичні касові збори (для калібрувальної перевірки)
y_true = df["Box Office"].astype(float).values

Q10 = pred["Q10"].values
Q50 = pred["Q50"].values
Q90 = pred["Q90"].values

# 2) Перевірка покриття квантилів (reliability)
# Фактична частка спостережень нижче прогнозованого квантиля
cov10 = np.mean(y_true <= Q10)
cov50 = np.mean(y_true <= Q50)
cov90 = np.mean(y_true <= Q90)

```

```

cal_report = pd.DataFrame({
    "tau": [0.1, 0.5, 0.9],
    "empirical_coverage": [cov10, cov50, cov90]
})
cal_report.to_csv("outputs/quantile_calibration_report.csv", index=False)

# 3) (Опційно) Калібрування квантилів.
# У дисертаційній реалізації це може бути ізотонічна корекція або лінійне масштабування.
# Тут показано простий шаблон: коригування може бути додано за потреби.

# 4) Формування сценарної сітки u_s.
# Щільніша сітка в зоні малих u дозволяє точніше оцінювати хвіст для CVaR.
S = 25
u = np.concatenate([np.linspace(0.01, 0.2, 10), np.linspace(0.25, 0.99, 15)])
p = np.ones(S) / S # базово рівноймовірні сценарії

def piecewise_quantile(u_val, q10, q50, q90):
    """Кусочно-лінійна апроксимація квантільної функції по трьох точках (0.1, 0.5, 0.9)."""
    if u_val <= 0.5:
        # Інтерполяція між (0.1, q10) і (0.5, q50)
        return q10 + (u_val - 0.1) * (q50 - q10) / (0.5 - 0.1)
    else:
        # Інтерполяція між (0.5, q50) і (0.9, q90)
        return q50 + (u_val - 0.5) * (q90 - q50) / (0.9 - 0.5)

# 5) Генерація матриці сценарних доходів S(i,s)
n = len(pred)
S_matrix = np.zeros((n, S), dtype=float)

for i in range(n):
    for s in range(S):
        S_matrix[i, s] = piecewise_quantile(u[s], Q10[i], Q50[i], Q90[i])

# 6) Збереження сценаріїв та ваг
# Формат long table зручний для імпорту у GAMS
rows = []
for i in range(n):
    for s in range(S):
        rows.append((i+1, s+1, S_matrix[i, s]))

scen = pd.DataFrame(rows, columns=["i", "s", "S_value"])
scen.to_csv("outputs/scenarios_S_i_s.csv", index=False)

weights = pd.DataFrame({"s": np.arange(1, S+1), "p": p})
weights.to_csv("outputs/scenario_weights.csv", index=False)

# 7) Витрати (бюджет) як C(i) – відповідно до логіки фінансового результату
costs = pd.DataFrame({"i": np.arange(1, n+1), "C": pred["Budget"].astype(float).values})
costs.to_csv("outputs/costs_C_i.csv", index=False)

```

### Лістинг 3.3 – Завантаження сценарних даних у GAMS та постановка CVaR-оптимізації

#### Sets

```
i "кінопроекти"
s "сценарії";
```

Alias (s,sp);

#### Parameters

```
Srev(i,s) "сценарні касові збори S(i,s)"
C(i)      "витрати/бюджет C(i)"
p(s)     "ваги сценаріїв p(s)"
B        "загальний бюджет"
alpha    "рівень довіри для CVaR"
lambda   "параметр ризик-аверсії";
```

#### Binary Variables

```
x(i)     "включення проекту i до портфеля";
```

#### Positive Variables

```
Loss(s)  "downside-втрати max(0, -P(s))"
u(s)     "допоміжні змінні для CVaR";
```

#### Variables

```
P(s)     "прибуток до оподаткування у сценарії s"
eta      "VaR-параметр"
CVaR     "умовна вартість під ризиком"
Obj      "цільова функція";
```

#### Equations

```
ProfitDef(s)  "визначення P(s)"
BudgetLimit   "бюджетне обмеження"
LossDef1(s)   "Loss(s) >= -P(s)"
LossDef2(s)   "Loss(s) >= 0"
CvarUdef(s)   "u(s) >= Loss(s) - eta"
CvarUpos(s)   "u(s) >= 0"
CvarDef       "визначення CVaR"
ObjDef        "цільова функція";
```

\* ---- Імпорт даних (приклад через CSV2GDX або таблиці) ----

\* У реальній реалізації тут додається блок з execute\_load або csv2gdx,

\* який завантажує множини i,s та параметри Srev(i,s), C(i), p(s).

\* ---- Параметри експерименту ----

B = 4e8;

alpha = 0.90;

lambda = 2.0;

\* ---- Сценарний прибуток до оподаткування:  $P(s) = \sum (S(i,s) - C(i)) * x(i)$  ----

ProfitDef(s)..

```
P(s) =E= sum(i, (Srev(i,s) - C(i)) * x(i));
```

```

* ---- Бюджетне обмеження ----
BudgetLimit..
    sum(i, C(i) * x(i)) =L= B;

* ---- Downside-втрати: Loss(s)=max(0, -P(s)) через лінійаризацію ----
LossDef1(s)..
    Loss(s) =G= -P(s);

LossDef2(s)..
    Loss(s) =G= 0;

* ---- Лінійне подання CVaR: u(s) >= Loss(s) - eta, u(s)>=0 ----
CvarUdef(s)..
    u(s) =G= Loss(s) - eta;

CvarUpos(s)..
    u(s) =G= 0;

CvarDef..
    CVaR =E= eta + (1/(1-alpha)) * sum(s, p(s) * u(s));

* ---- Цільова функція: max E[P] - lambda*CVaR ----
ObjDef..
    Obj =E= sum(s, p(s) * P(s)) - lambda * CVaR;

Model PortfolioCVaR /ProfitDef, BudgetLimit, LossDef1, LossDef2, CvarUdef, CvarUpos,
CvarDef, ObjDef/;

Solve PortfolioCVaR using MIP maximizing Obj;

* ---- Вивід результатів ----
Display Obj.l, CVaR.l, eta.l;
Display x.l, P.l, Loss.l;

```

### Лістинг 3.4 – Автоматизований запуск експерименту

```

# 1) готує квантильні прогнози
# 2) генерує сценарії та ваги
# 3) експортує дані для GAMS
# 4) запускає GAMS-модель
# 5) збирає результати та формує таблиці/графіки

import os
import subprocess
import json
import pandas as pd
import numpy as np

PROJECT_ROOT = "."
DATA_PATH = os.path.join(PROJECT_ROOT, "data", "mmc1.csv")
OUT_DIR = os.path.join(PROJECT_ROOT, "outputs")
GAMS_DIR = os.path.join(PROJECT_ROOT, "gams")

```

```
GAMS_MODEL = os.path.join(GAMS_DIR, "SM_5_V2.gms") # або ваш файл моделі
```

```
os.makedirs(OUT_DIR, exist_ok=True)
```

```
# -----
```

```
# 0) Параметри експерименту
```

```
# -----
```

```
config = {
    "S": 25,
    "u_grid": {"tail": [0.01, 0.2, 10], "mid": [0.25, 0.99, 15]},
    "B": 4e8,
    "alpha": 0.90,
    "lambda": 2.0,
    "random_state": 42
}
```

```
with open(os.path.join(OUT_DIR, "run_config.json"), "w", encoding="utf-8") as f:
    json.dump(config, f, ensure_ascii=False, indent=2)
```

```
# -----
```

```
# 1) Квантильні прогнози (виклик модуля/скрипта)
```

```
# -----
```

```
# У вашій структурі це може бути: python src/02_quantile_models.py
subprocess.check_call(["python", "src/02_quantile_models.py", "--data", DATA_PATH, "--out",
OUT_DIR])
```

```
# Очікуємо outputs/quantile_predictions.csv
```

```
q_path = os.path.join(OUT_DIR, "quantile_predictions.csv")
pred = pd.read_csv(q_path)
```

```
# -----
```

```
# 2) Генерація сценаріїв (виклик модуля/скрипта)
```

```
# -----
```

```
# Напр.: python src/04_scenarios.py --S 25 --out outputs
subprocess.check_call([
    "python", "src/04_scenarios.py",
    "--data", DATA_PATH,
    "--quantiles", q_path,
    "--S", str(config["S"]),
    "--out", OUT_DIR
])
```

```
# Очікуємо outputs/scenarios_S_i_s.csv, outputs/scenario_weights.csv, outputs/costs_C_i.csv
```

```
scen_path = os.path.join(OUT_DIR, "scenarios_S_i_s.csv")
w_path = os.path.join(OUT_DIR, "scenario_weights.csv")
c_path = os.path.join(OUT_DIR, "costs_C_i.csv")
```

```
# -----
```

```
# 3) Підготовка include-файлів для GAMS (простий CSV→INC)
```

```
# -----
```

```
scen = pd.read_csv(scen_path)
weights = pd.read_csv(w_path)
costs = pd.read_csv(c_path)
```

```
# Множини
```

```

i_max = int(scen["i"].max())
s_max = int(scen["s"].max())

with open(os.path.join(GAMS_DIR, "sets.inc"), "w", encoding="utf-8") as f:
    f.write(f"set i /1*{i_max}/;\n")
    f.write(f"set s /1*{s_max}/;\n")

# Параметри Srev(i,s)
# Формат: parameter Srev(i,s) / i.s value, ... /;
with open(os.path.join(GAMS_DIR, "Srev.inc"), "w", encoding="utf-8") as f:
    f.write("parameter Srev(i,s) ^\n")
    for _, r in scen.iterrows():
        f.write(f" {int(r['i'])} {int(r['s'])} {float(r['S_value']):.6g}\n")
    f.write(";\n")

# Параметри C(i)
with open(os.path.join(GAMS_DIR, "C.inc"), "w", encoding="utf-8") as f:
    f.write("parameter C(i) ^\n")
    for _, r in costs.iterrows():
        f.write(f" {int(r['i'])} {float(r['C']):.6g}\n")
    f.write(";\n")

# Ваги p(s)
with open(os.path.join(GAMS_DIR, "p.inc"), "w", encoding="utf-8") as f:
    f.write("parameter p(s) ^\n")
    for _, r in weights.iterrows():
        f.write(f" {int(r['s'])} {float(r['p']):.6g}\n")
    f.write(";\n")

# Параметри експерименту
with open(os.path.join(GAMS_DIR, "params.inc"), "w", encoding="utf-8") as f:
    f.write(f"B = {config['B']};\n")
    f.write(f"alpha = {config['alpha']};\n")
    f.write(f"lambda = {config['lambda']};\n")

# -----
# 4) Запуск GAMS
# -----
# Важливо: GAMS має бути доступний у PATH.
subprocess.check_call(["gams", GAMS_MODEL, "lo=2"], cwd=GAMS_DIR)

# -----
# 5) Збір результатів
# -----
# У реальній роботі зручно налаштувати в GAMS вивід у CSV (наприклад, x.csv).
# Тут показано шаблон читання такого файлу.
x_path = os.path.join(GAMS_DIR, "x_solution.csv")
if os.path.exists(x_path):
    xsol = pd.read_csv(x_path) # колонки: i, x
    selected = xsol[xsol["x"] > 0.5]["i"].astype(int).values
    pred_sel = pred.iloc[selected - 1][["ID", "Title", "Budget", "Q10", "Q50", "Q90"]]
    pred_sel.to_excel(os.path.join(OUT_DIR, "portfolio_selected.xlsx"), index=False)

print("Експеримент завершено. Результати збережено у папці outputs/.")

```