НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ
«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ імені ІГОРЯ СІКОРСЬКОГО»
Факультет менеджменту та маркетингу
Кафедра економічної кібернетики

ДО ЗАХИСТУ ДОПУЩЕНО
Завідувач кафедри
_____ Катерина БОЯРИНОВА
«____» червня 2025 р.

ДИПЛОМНА РОБОТА

на здобуття ступеня бакалавра
за освітньо–професійною програмою
«Економічна кібернетика»
спеціальності 051 «Економіка»
на тему: «Економіко-математичне моделювання маршрутизації фінансових потоків підприємства»

Виконала:
студентка IV курсу, групи УК–11
Неділько Яна Олексіївна _____
Керівник:
доцент кафедри економічної кібернетики, к.ф.-м.н., доцент
Лазаренко Ірина Сергіївна _____
Рецензент:
професор кафедри міжнародної економіки, д.т.н., професор
Гавриш Олег Анатолійович _____

Засвідчую, що у цій дипломній роботі немає запозичень з праць інших авторів без відповідних посилань.

Студентка _____

Київ – 2025

Національний технічний університет України

«Київський політехнічний інститут імені Ігоря Сікорського»

Факультет менеджменту та маркетингу

Кафедра економічної кібернетики

Рівень вищої освіти – перший (бакалаврський)

Спеціальність – 051 «Економіка»

Освітньо-професійна програма «Економічна кібернетика»

ЗАТВЕРДЖУЮ

Завідувач кафедри

_____ Катерина БОЯРИНОВА

«05» лютого 2025 р.

ЗАВДАННЯ

НА ДИПЛОМНУ РОБОТУ СТУДЕНТЦІ

Неділько Яні Олексіївні

1. Тема роботи: **«Економіко-математичне моделювання маршрутизації фінансових потоків підприємства»**
керівник роботи **Лазаренко Ірина Сергіївна**, к.ф.-м.н., доцент
затверджені наказом по університету від 26.05.2025 р. № 1747-с
2. **Термін подання студентом роботи:** 02.06.2025 р.
3. **Вихідні дані до роботи:** наукова та навчально-методична література, публікації вітчизняних та іноземних вчених, наукові статті, фінансово-економічні дані діяльності ТОВ «ТК Валенсія» а також наукові дослідження в галузі економіко-математичного моделювання та аналізу маршрутизації фінансових потоків підприємства.

4. Зміст пояснювальної записки:

а) теоретична частина:

− розкрити основні поняття маршрутизації фінансових потоків.

− проаналізувати перспективи розвитку обраної галузі дослідження.

− обґрунтувати вибір моделей оптимального розподілу маршрутизації фінансових потоків.

б) аналітична частина:

− надати економічний аналіз фінансових показників підприємства та проаналізувати маршрутизацію фінансових потоків підприємства.

− побудувати економіко-математичну модель оптимізації маршрутизації фінансових потоків.

− здійснити економіко-математичне моделювання маршрутизації фінансових потоків.

в) рекомендаційна частина:

− проаналізувати отримані результати моделювання та розробити рекомендації для забезпечення оптимального розподілу фінансових потоків підприємства.

− здійснити прогностичне моделювання очікуваного економічного ефекту від оптимізації маршрутизації фінансових потоків підприємства.

5. Перелік ілюстративного матеріалу

1. Класифікація фінансових потоків підприємства
2. Обсяг оброблених транзакцій платіжними системами за 2024 рік
3. Графічне відтворення моделі логістичної регресії
4. Відсоток успішних платежів у загальній кількості
5. Розподіл за класами можливих результатів моделювання
6. Значимі характеристики моделі
7. Динаміка ймовірності успіху логістична регресія
8. Динаміка ймовірності успіху дерева рішень
9. Динаміка ймовірності успіху випадковий ліс
10. Порівняння результатів моделювання
11. Прогноз економічного ефекту від оптимізації маршрутизації фінансових потоків підприємства
12. Відсотковий приріст ймовірності успішності по місяцях

6. Дата видачі завдання: «05» лютого 2025 р.

КАЛЕНДАРНИЙ ПЛАН

№ з/п	Назва етапів виконання дипломної роботи	Термін виконання етапів роботи	Примітка
1.	Збір необхідної інформації теоретичного, методичного та практичного змісту, вивчення та аналіз літературних джерел щодо маршрутизації фінансових потоків, моделей класифікації та маршрутизації	06.02.2025 – 23.02.2025	
2.	Розгляд теоретико-методичних засад дослідження маршрутизації фінансових потоків підприємства	24.02.2025 – 23.03.2025	
3.	Вибір підприємства, установи для дослідження та постановка економічної задачі	24.03.2025 – 30.03.2025	
4.	Аналіз фінансових показників підприємства та маршрутизації його фінансових потоків.	31.03.2025 – 13.04.2025	
5.	Побудова економіко-математичної моделі, методики розрахунків, вибір програмного забезпечення, розробка програмного коду	14.04.2025 – 27.04.2025	
6.	Економіко-математичне моделювання маршрутизації фінансових потоків підприємства	28.04.2025 – 11.05.2025	
7.	Аналіз отриманих результатів моделювання, порівняння використаних моделей	12.05.2025 – 18.05.2025	
8.	Прогностичне моделювання та формування рекомендацій щодо оптимізації маршрутизації фінансових потоків підприємства	19.05.2025 – 25.05.2025	
9.	Оформлення дипломної роботи першого (бакалаврського) рівня вищої освіти	26.05.2025 – 1.06.2025	

Студентка _____ Яна НЕДІЛЬКО

(підпис)

Керівник дипломної роботи _____ Ірина ЛАЗАРЕНКО

(підпис)

РЕФЕРАТ

Дипломна робота першого (бакалаврського) рівня вищої освіти на тему: «Економіко-математичне моделювання маршрутизації фінансових потоків підприємства» містить 86 сторінок, 9 таблиць, 22 рисунка, 1 додаток. Перелік посилань нараховує 32 найменування.

Метою роботи є розгляд теоретичних засад та економіко-математичне моделювання маршрутизації фінансових потоків підприємства.

Об'єктом дослідження є процес маршрутизації фінансових потоків підприємства.

Предметом дослідження є сукупність теоретичних аспектів, методів, моделей маршрутизації фінансових потоків підприємства.

Методами дослідження були узагальнення економічної інформації, економіко-математичне моделювання, графічний метод та інші аналітичні підходи.

Результати дипломної роботи:

Проведення оцінку та моделювання маршрутизації фінансових потоків підприємства. Було розглянуто два основні підходи до задачі моделювання маршрутизації фінансових потоків підприємства. За класифікаційним підходом було визначено чотири основні моделі та принципи їх оцінки. Методичні положення моделювання маршрутизації фінансових потоків були доповнені за рахунок використання математичних та економіко-математичних методів аналізу. Визначено основні методи оцінювання якості побудованих моделей.

Розроблено та застосовано моделі для маршрутизації фінансових потоків підприємства ТОВ «ТК Валенсія», яка дозволяє покращити маршрутизацію фінансових потоків. Розраховано прогнозований приріст прибутку підприємства за умов застосування моделі оптимальної маршрутизації фінансових потоків підприємства. Модель є динамічною, автоматичною та адаптується під умови повної або часткової невизначеності.

Рекомендації щодо використання результатів роботи.

Результати роботи можуть бути використані для покращення економічних результатів підприємства за допомогою моделювання оптимальної маршрутизації фінансових потоків. Запропонована економіко-математична модель дозволяє підприємствам динамічно приймати зважені рішення щодо оптимальної маршрутизації фінансових потоків підприємства. Модель є універсальною та не залежить від розміру підприємства чи галузі, в якій воно працює.

Ключові слова: маршрутизація, моделювання, фінансові потоки, динаміка розподілу фінансових потоків, прогнозування.

ABSTRACT

Thesis of the first (bachelor's) level of higher education on «Economic and Mathematical Modeling of Financial Flow Routing in an Enterprise» contains 86 pages, 9 tables, 22 figures, and 1 appendix, the list of links contains 32 item.

The aim of the thesis is to study the theoretical aspects and modeling of enterprise financial flow routing.

The object of the study is the process of routing of the financial flows within an enterprise.

The subject of the study is the complecs of theoretical aspects, methods, and models of enterprise financial flow routing.

Research methods included the generalization of economic information, economic and mathematical modeling, graphical methods, and other analytical approaches.

Thesis results:

An evaluation and modeling of enterprise financial flow routing were conducted. Two main approaches to modeling financial flow routing in enterprises were considered. Based on the classification approach, four primary models and their evaluation principles were identified. The methodological framework for modeling financial flow routing was expanded by incorporating mathematical and economic-mathematical methods of analysis. Key methods for assessing the quality of the constructed models were determined.

Models for financial flow routing were developed and applied to the enterprise LLC "TK Valencia", enabling improved routing of financial flows. A forecasted increase in the company's profit was calculated under the condition of applying the optimal financial flow routing model. The model is dynamic, automated, and adaptable to conditions of full or partial uncertainty.

Recommendations for the use of the thesis results:

The findings can be used to improve a company's economic performance through optimal financial flow routing modeling. The proposed economic and mathematical model allows enterprises to make informed, time-sensitive decisions regarding optimal financial flow routing. The model is universal and independent of the company's size or the industry in which it operates.

Keywords: routing, modeling, financial flows, financial flow distribution dynamics, forecasting.

ЗМІСТ

ВСТУП 8
1 ТЕОРЕТИКО-МЕТОДИЧНІ ЗАСАДИ ДОСЛІДЖЕННЯ МАРШРУТИЗАЦІЇ ФІНАНСОВИХ ПОТОКІВ 12
1.1 Сутність та основні поняття маршрутизації фінансових потоків 12
1.2 Аналіз перспектив розвитку моделювання маршрутизації фінансових потоків 19
1.3 Аналіз та обґрунтування вибору моделей оптимального розподілу маршрутизації фінансових потоків 24
Висновки до першого розділу 39
2 АНАЛІЗ ПЕРЕДУМОВ ТА ПОСТАВКА ЕКОНОМІЧНОГО ЗАВДАННЯ МОДЕЛЮВАННЯ МАРШРУТИЗАЦІЇ ФІНАНСОВИХ ПОТОКІВ 41
2.1 Аналіз економічних показників підприємства та організації маршрутизації фінансових потоків 41
2.2 Побудова економіко-математичне моделі оптимізації фінансових потоків підприємства 48
2.3 Реалізація економіко-математичного моделювання маршрутизації фінансових потоків підприємства 53
Висновки до другого розділу 59
3 ЗАСТОСУВАННЯ МОДЕЛЕЙ МАРШРУТИЗАЦІЇ ФІНАНСОВИХ ПОТОКІВ ПІДПРИЄМСТВА 61
3.1 Аналіз отриманих результатів моделювання маршрутизації фінансових потоків підприємства 61
3.2 Прогностичне моделювання очікуваного економічного ефекту від оптимізації маршрутизації фінансових потоків 69
Висновки до третього розділу 74
ВИСНОВКИ 76
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ 79
ДОДАТКИ 83

ВСТУП

Актуальність теми дослідження: в умовах воєнного стану, віялових відключень світла та можливих уражень перед підприємствами постали виклики забезпечення стабільності проведення онлайн платежів. Одним з способів мінімізації ризиків неможливості проведення платежів стало підключення більшої кількості надавачів послуг проведення платежів. Перед підприємствами постало питання маршрутизації платежів з метою досягнення максимальної стабільності при мінімізації відповідних затрат. Розробка автоматичної моделі маршрутизації платежів посприяє автоматизації та оптимізації процесу розподілу.

Метою роботи є розгляд теоретичних засад та економіко-математичне моделювання маршрутизації фінансових потоків підприємства.

Об'єктом дослідження є процес маршрутизації фінансових потоків підприємства.

Предметом дослідження є сукупність теоретичних аспектів, методів, моделей маршрутизації фінансових потоків підприємства.

Методи дослідження. У процесі дослідження застосовувалися як теоретичні, так і прикладні методи аналізу. Узагальнення економічної інформації дозволило систематизувати й інтерпретувати наявні дані, визначити основні закономірності та тенденції. Економіко-математичне моделювання використовувалося для формалізації економічних процесів і побудови математичних моделей, що дали змогу провести кількісну оцінку показників і проаналізувати взаємозв’язки між ними. Графічний метод забезпечив наочне представлення результатів, що сприяло кращому розумінню структурних змін і динаміки процесів. Прогностичні методи були використані для побудови сценаріїв майбутнього розвитку та оцінки потенційних економічних результатів. Крім того, застосовано низку аналітичних підходів, які поєднували елементи порівняльного аналізу, логічного узагальнення та візуалізації даних із використанням програмних засобів, зокрема Excel та Python.

Відповідно до поставленої мети було вирішено такі *завдання*:

– розкрити основні поняття маршрутизації фінансових потоків.

– проаналізувати перспективи розвитку обраної галузі дослідження.

– обґрунтувати вибір моделей оптимального розподілу маршрутизації фінансових потоків.

– надати економічний аналіз фінансових показників підприємства та проаналізувати маршрутизацію фінансових потоків підприємства.

– побудувати економіко-математичну модель оптимізації маршрутизації фінансових потоків.

– здійснити економіко-математичне моделювання маршрутизації фінансових потоків.

– проаналізувати отримані результати моделювання та розробити рекомендації для забезпечення оптимального розподілу фінансових потоків підприємства.

– здійснити прогностичне моделювання очікуваного економічного ефекту від оптимізації маршрутизації фінансових потоків підприємства.

База дослідження: ТОВ «ТК Валенсія»

Питання дослідження маршрутизації фінансових потоків та забезпечення їх ефективного розподілу висвітлювалися в наукових працях таких дослідників, як: Майборода О. В., який у своїй дисертації аналізував підходи до управління фінансовими потоками підприємства, виокремлюючи методи їх оптимізації з урахуванням внутрішніх і зовнішніх факторів [1]; Єрешко Ю. О., що досліджувала структуру та механізм управління грошовими потоками підприємств, акцентуючи увагу на системності та адаптивності фінансового управління [2]; Чайковський Я. І., який розглядав платіжні системи як важливу складову інфраструктури для ефективного руху фінансових ресурсів, з урахуванням сучасних технологічних рішень [3]; Гринько В. В., що вивчав застосування методів машинного навчання в економіці, зокрема у прогнозуванні й оптимізації фінансових потоків на основі аналізу великих

обсягів даних [26]; Прокопенко В. Ю., який аналізував інтелектуальні системи підтримки прийняття рішень як інструмент автоматизації фінансового управління та побудови адаптивних моделей маршрутизації фінансових потоків. Праці зазначених авторів формують наукову базу для розробки сучасних автоматизованих підходів до управління фінансовими потоками, що особливо актуально в умовах високої турбулентності середовища, зокрема під час воєнного стану, енергетичних криз і нестабільності платіжної інфраструктури.

Результати дипломної роботи:

Проведення оцінку та моделювання маршрутизації фінансових потоків підприємства. Було розглянуто два основні підходи до задачі моделювання маршрутизації фінансових потоків підприємства. За класифікаційним підходом було визначено чотири основні моделі та принципи їх оцінки. Методичні положення моделювання маршрутизації фінансових потоків були доповнені за рахунок використання математичних та економіко-математичних методів аналізу. Визначено основні методи оцінювання якості побудованих моделей та економічний ефект від оптимізації фінансових потоків підприємства.

Розроблено та застосовано моделі для маршрутизації фінансових потоків підприємства ТОВ «ТК Валенсія», яка дозволяє покращити маршрутизацію фінансових потоків. Розраховано прогнозований приріст прибутку підприємства за умов застосування моделі оптимальної маршрутизації фінансових потоків підприємства. Модель є динамічною, автоматичною та адаптується під умови повної або часткової невизначеності.

Рекомендації щодо використання результатів роботи.

Результати роботи можуть бути використані для покращення економічних результатів підприємства за допомогою моделювання оптимальної маршрутизації фінансових потоків. Запропонована економіко-математична модель дозволяє підприємствам динамічно приймати зважені рішення щодо оптимальної

маршрутизації фінансових потоків підприємства. Модель є універсальною та не залежить від розміру підприємства чи галузі, в якій воно працює.

Публікації: Неділько Я. О. Сучасні методи і моделі оптимізації online платежів. Моделювання та прогнозування економічних процесів: зб. тез доп. XVIII Міжнар. наук.-практ. конф. Київ: КПІ ім. Ігоря Сікорського, 2024. С. 88. URL: https://ecocyber.fmm.kpi.ua/wp-content/uploads/2025/05/mpep_XVIII_2024.pdf

1 ТЕОРЕТИКО-МЕТОДИЧНІ ЗАСАДИ ДОСЛІДЖЕННЯ МАРШРУТИЗАЦІЇ ФІНАНСОВИХ ПОТОКІВ

1.1 Сутність та основні поняття маршрутизації фінансових потоків

Поняття фінансових потоків на підприємстві можна розглядати як направлений рух фінансових ресурсів, у процесі якого здійснюється господарська діяльність, унаслідок чого виникають економічні відносини [1]. Основні фінансові потоки підприємства виникають під час економічної діяльності підприємства з постачальниками, клієнтами та державою.

Розглянемо класифікацію фінансових потоків підприємства (див. рис. 1.1).

За видами діяльності фінансові потоки класифікують за походженням від основної та неосновної, побічної, діяльності та за видами діяльності з дочірніми підприємствами, за окремими господарськими операціями та за структурними підрозділами [1].

За видами валюти фінансові потоки можуть бути класифікованими як розрахунки у національній валюті, вільноконвертованій валюті, наприклад долар США та Євро, та в інших валютах [1]. За методами обчислення обсягу розрізняють валовий, тобто загальний, та чистий фінансові потоки підприємства [1]. Чистий фінансовий потік визначається як різниця між надходженням та вибуттям коштів підприємства при виконанні ним економічної діяльності [1].

За рівнем передбачуваності класифікуються повністю передбачувані, частково передбачувані та форс мажорні, частково передбачувані в свою чергу можуть бути умовно поділені за ймовірністю виникнення на малоймовірні, ймовірна та з високим рівнем ймовірності [1].

За рівнем регулярності руху розділяють регулярні фінансові потоки, за часом, обсягами чи напрямком руху, та дискретні, тобто фінансові потоки котрі виникають

внаслідок одиничних, разових операцій [1]. За формою розрахунків розрізняють готівкові та безготівкові розрахунки [1].

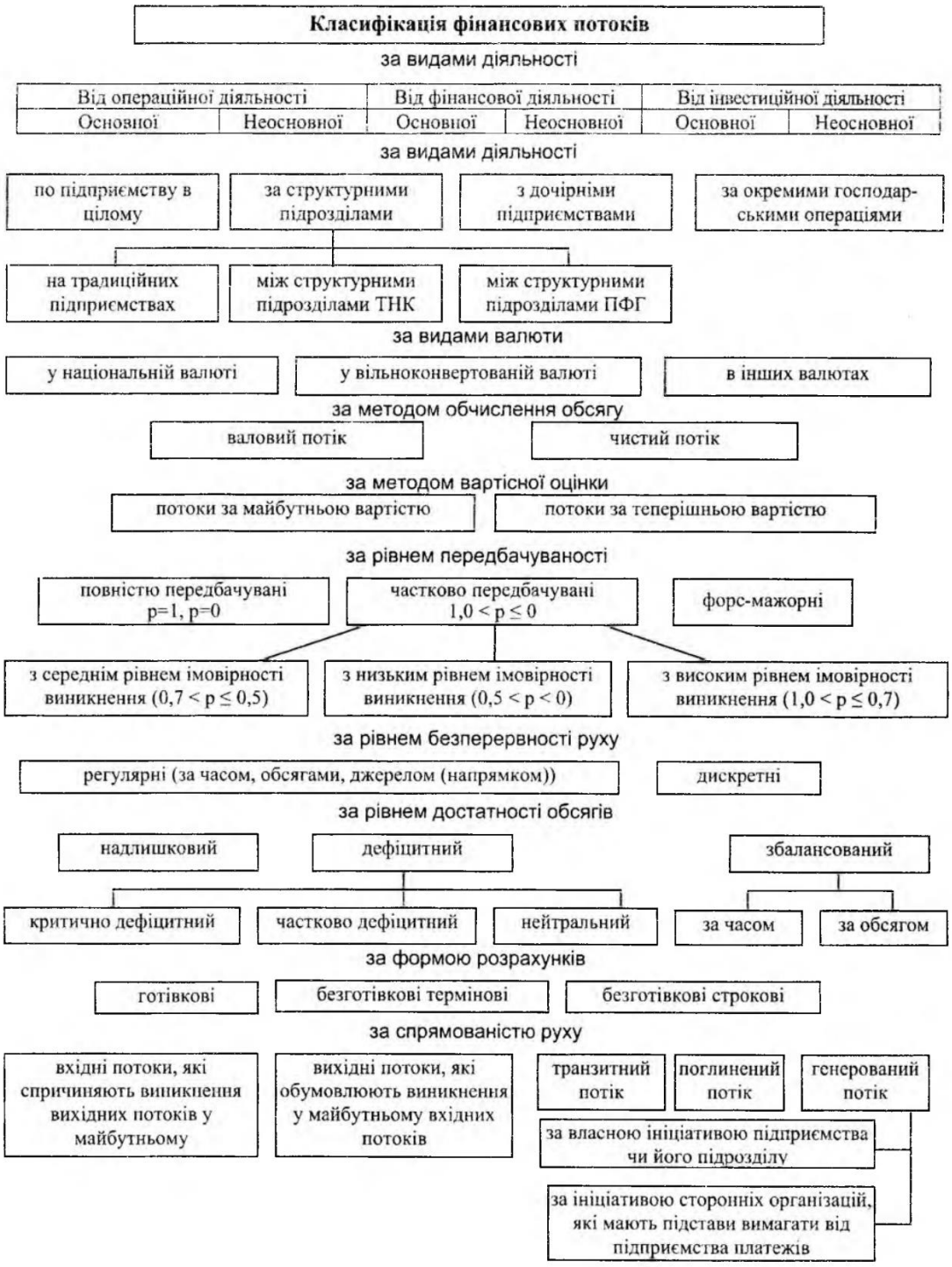

Рисунок 1.1 – Класифікація фінансових потоків підприємства

Джерело: [6]

Досягнення підприємством стабільного зростання забезпечується через розвиток і оптимізацію всіх ланок і сфер економічної діяльності суб'єкта господарювання [2]. Ключовим елементом економічної системи підприємства, при цьому, виступають грошові потоки, що забезпечують координацію потоків і залишків всіх видів ресурсів підприємства, обслуговуючи його діяльність [2]. Тому, саме через оптимізацію грошових потоків доцільно управляти економічним зростанням підприємства [2]. Сучасна вітчизняна наука характеризується наявністю двох підходів до проблеми оптимізації грошових потоків підприємства, розроблених І.О. Бланком та Л.О. Лігоненко і Г.В. Ситник [2]. Обидві торії наполягають на таких напрямках оптимізації, як збалансування і синхронізація надходжень і видатків грошових коштів[2].

У подальшому дослідженні буде розглянуто проблему моделювання та маршрутизації фінансових потоків підприємства саме на основі транзакцій, як специфічного виду фінансових потоків підприємства. Транзакція може бути визначена як банківська операція, що передбачає переказ коштів з одного рахунку на інший. Розглянемо класифікацію транзакцій як фінансового потоку підприємства:

- походять від основної операційної діяльності компанії
- є розрахунком з окремими господарськими організаціями, фізичними або юридичними особами
- обчислюються як валовий потік за теперішньою вартістю
- є частково передбачуваними за обсягами, часом та джерелами
- дискретні
- за формою розрахунку є безготівковими

В умовах розвитку електронної комерції та переформатування бізнесу на моделі онлайн роботи набуло важливості питання проведення платежів з мінімізацією витрат на проведення платежів, ризиків відхилення та комфортної роботи для користувачів.

У 1980х роках виникла нова форма бізнесу, така як платіжні системи, а саме системи котрі займаються обробкою та проведенням платежів у різних країнах світу

використовуючи різні методи оплати. Найбільш популярним методом оплати є платіж з банківської картки користувача, платіж за згенерованим qr кодом та метод синхронного банківського переказу.

Відповідно до означення, платіжна система – це сукупність визначених законом елементів, які служать для переказу грошей між суб'єктами економіки.

Основними суб'єктами платіжної системи є: платіжна організація, члени і учасники платіжної системи [3].

Найбільшими платіжними системами у світі прийнято вважати Visa, Mastercard, Advanced Cash, PayPal, Perfect Money, Payeer, Skrill. Кожна платіжна система має власні переваги та недоліки, котрі дозволяють підприємствам використовувати декілька платіжних систем для максимізації корисності [4].

Проте надавачі послуг проведення онлайн-платежів, стикаються з численними викликами, зокрема підвищення відсотку успішних транзакцій, мінімізація ризику шахрайських транзакцій, зниження відсотку відшкодування транзакцій (refund) та відкликаних платежів (chargedback), що є ключовими метриками у сфері електронних платежів.

Станом на 2024 рік ринок платіжних провайдерів оцінюється у 56.31 триліони доларів США [3]. Платіжні системи розширюють класичні платіжні методи, такі як карткові операції Visa, Mastercard, American Express та Maestro, шляхом розширенням на ринок криптовалюти, гаманців, альтернативних платежів та нових локацій.

Порівняємо дані про обсяг оброблених платежів деякими платіжними провайдерами за 2024 рік та розробимо інфографіку, див. рис. 1.2 на основі офіційних даних представлених на сайті провайдерів [5], [6], [7], [8], [9], [10].

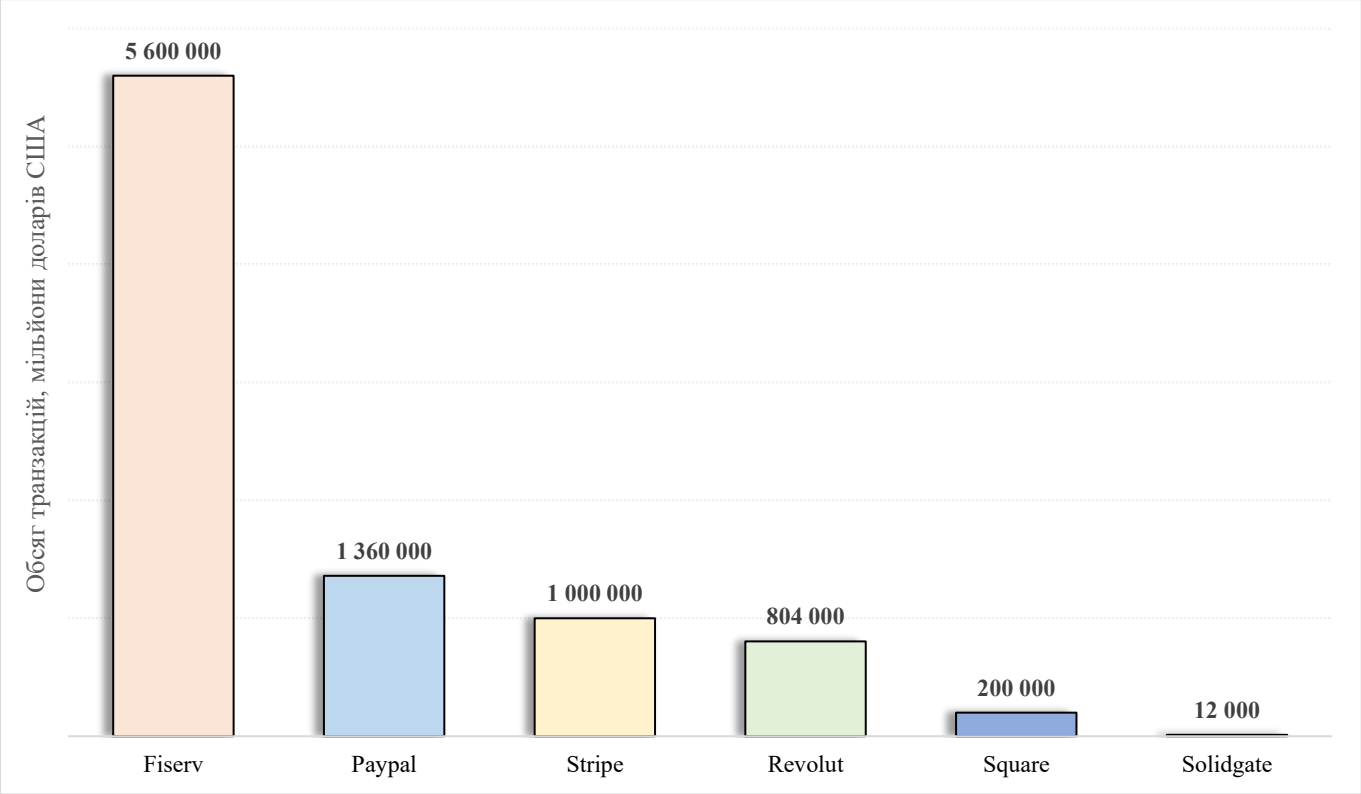

Рисунок 1.2 – Обсяг оброблених транзакцій платіжними системами за 2024 рік

Джерело: розроблено автором на основі [5], [6], [7], [8], [9], [10]

Розглянемо конкурентні переваги кожної з зазначених платіжних систем:

Fiserv – надає послуги одночасно як банківська система з випуску карток для розрахунку та проведення платежів для бізнесу [10]. Має власний університет для навчання фахівців компаній партнерів, обслуговує такі компанії, як Microsoft, Google та інші великі корпорації [5]. Тривалий час залишається експертом на ринку у сферах ризик-менеджменту та управління трафіком завдяки можливостям індивідуальних технічних налаштувань [5].

Paypal – є найвідомішою платформою для онлайн-платежів завдяки тривалому існуванню (заснована в 1998 році) [11]. Підтримує транзакції у 200+ країнах і більш ніж у 25 валютах [6].

Revolut – крім платежів, пропонує банківські послуги, включаючи заощадження, інвестиції та криптоактиви [6]. Тримає молоду аудиторію завдяки монетизації бонусів, використанні ігрових методів утримування аудиторії та

відсутності комісії при криптопереказах та зручному внутрішньому сервісі інвестування [6].

Square – Лідер у рішенні для точок продажу (POS) через зручні мобільні термінали [9]. Інтеграція з Cash App забезпечує зручний перехід між бізнес і споживчими послугами [9].

Solidgate – українська платіжна система, спрямована на менеджмент трафіку завдяки методам машинного навчання, налаштування логіки повторення платежів та кастомних правил для боротьби з шахрайством [10].

Stripe – високотехнологічна платіжна система котра налічує понад 200 партнерів для інтеграції [8]. Розвиває продуктом шляхом експертизи та пропозицій для покращення трафіку завдяки продуктам, спрямованих на зменшення рівню шахрайства Stripe Radar, керуванням процесом створення транзакцій для різних систем монетизації Stripe Billing, системі фінансової аналітики Revenue Recognition, автоматизації розрахунку та створення податків та декларацій Stripe Tax, та системі аналітики та репортингу Stripe Sigma [8].

Відповідно, усі провідні платіжні системи поєднує використання розумних систем роутинг транзакцій, платіжної логіки, ризик логіки та математико-аналітичного підходу до транзакцій заснованих на машинному навчанні.

Проте, платіжні системи є складною системою організації, котрі працюють у відповідності до регуляторних обмежень держави, відповідно до обмежень існуючих на рівні держави, міжнародних обмежень та внутрішніх правил підприємств-партнерів котрі довіряють проведення платежів в межах власної інфраструктури [3].

Учасниками платіжної системи є всі види підприємств, державні структури, громадські організації, населення, банківські установи та центральний банк [3]. Встановлення в межах платіжної системи для всіх її учасників визначених законом норм та правил поведінки ставить їх у рівні умови, захищає інтереси кожного з них [3]. Тому створення високоефективної, дійової та надійної платіжної системи є одним з ключових економічних завдань держави [3].

Не зважаючи на простоту задачі у задумці та підтримки вищих органів влади платіжні системи будують власну інфраструктуру на базі даних отриманих від підприємств-партнерів (мерчантів). Аналіз та математичне моделювання даних відіграє важливу роль у процесі обробки даних. Дані про користувача та транзакції оброблюються у момент надходження та візуалізуються у вигляді інтерактивних дешбордів, звітів та використання метрик. Розглянемо як найбільші платіжні системи працюють з маршрутизацією фінансових потоків (табл. 1.1).

Таблиця 1.1 – Підходи до маршрутизації фінансових потоків основних платіжних систем світу

Компанія	**Підхід до оптимізації маршрутизації фінансових потоків**
PayPal	PayPal активно працює над зниженням витрат на обробку платежів шляхом впровадження системи автоматичного вибору оптимального процесингового партнера. Це дозволяє гнучко перенаправляти транзакції через ті платформи, які надають найвигідніші умови в конкретний момент часу, враховуючи комісії, швидкість обробки та географічні обмеження.
Square	Компанія Square зосереджується на створенні локалізованих платіжних рішень, що дозволяють оптимізувати маршрути транзакцій. Завдяки інтеграції з локальними фінансовими інститутами, Square забезпечує швидшу та дешевшу обробку платежів, зменшуючи залежність від глобальних мереж і, відповідно, витрати.
Stripe	Stripe використовує складні алгоритми маршрутизації, які враховують не лише вартість обробки (Processing Cost), а й затримку транзакції (Latency). Система постійно аналізує великі обсяги даних для вибору найефективнішого маршруту, що дозволяє досягати високої швидкості й надійності обробки платежів.
Revolut	Revolut має в своєму розпорядженні широку глобальну мережу банківських і фінансових партнерів. Завдяки цьому компанія може обирати найкращі маршрути для міжнародних транзакцій, мінімізуючи витрати на конверсію валют та комісії. Гнучкість у виборі партнерів дозволяє підлаштовуватися під вимоги різних регіонів.
Solidgate	Solidgate пропонує індивідуальні рішення для бізнес-клієнтів, орієнтовані на оптимізацію міжнародних платежів. Компанія використовує модульний підхід, що дозволяє налаштовувати маршрутизацію відповідно до специфіки конкретного бізнесу, включаючи тип транзакції, локацію клієнтів та валюту операцій.
Fiserv	Fiserv застосовує адаптивні процеси для маршрутизації транзакцій, що дозволяє обирати найкращий шлях проходження платежів у режимі реального часу. Компанія використовує автоматизовані системи оцінки ефективності маршрутів на основі аналітики попередніх транзакцій, що забезпечує економію ресурсів і стабільність фінансових потоків.

Джерело: структуровано автором на основі [5], [6], [7], [8], [9], [10]

Перед підприємства в умовах конкуренції між платіжними системами та необхідністю налаштувань декількох платіжних систем на один сайт з метою мінімізації ціни на проведення платежів та максимізації прибутку постає питання оптимального розподілу транзакцій між платіжними системами використовуючи певні системи та технології.

Шаблон двофазного коміту (2PC) є класичним розподіленим транзакційним шаблоном, який використовується для координації транзакцій між декількома системами [11]. У цьому патерні існує координатор, що ініціює транзакцію, й інші учасники (мікросервіси, вузли системи), що отримують повідомлення від координатора та зберігають транзакцію [11]. Координатор визначає готовність кожного вузла зберегти транзакцію і, у випадку готовності усіх сервісів, відправляє друге повідомлення, аби вказати всім сервісам зберегти транзакцію [11]. Якщо якийсь сервіс не готовий зберегти транзакцію, координатор відправляє повідомлення про відмову від транзакції [11]. Цей патерн вирішує проблему консистентності даних в PC забезпечуючи те, що всі сервіси здійснюють або скасовують транзакцію у координований спосіб (за допомогою сервісу координатора) [11].

Існують системи розподілу транзакцій, які базуючись на системах штучного інтелекту та внутрішніх правил підприємства дозволяють налаштувати системи розподілу транзакцій для користувачів у момент ініціювання транзакції.

1.2 Аналіз перспектив розвитку моделювання маршрутизації фінансових потоків

Маршрутизація фінансових потоків є невід’ємною складовою ефективного управління фінансами підприємства, що забезпечує раціональний розподіл і використання наявних фінансових ресурсів. У сучасних умовах підвищення конкуренції та динамічності ринкового середовища завдання оптимізації грошових потоків стає особливо актуальним, оскільки від цього залежить фінансова стійкість,

платоспроможність і загальна прибутковість підприємства. Економіко-математичне моделювання маршрутизації фінансових потоків надає можливість здійснювати більш точний і своєчасний аналіз, прогнозування та оптимізацію цих процесів, що в свою чергу сприяє підвищенню ефективності фінансового менеджменту.

На розвиток господарської діяльності промислових підприємств значною мірою впливає недостатність обігових коштів і нераціональне їх використання, що призводить до втрати ринків збуту, виникнення кризи неплатежів, збиткової діяльності, що в кінцевому підсумку впливає на конкурентоспроможність та може спричинити банкрутство підприємства [1].

Задоволення попиту на гроші підприємства має ключове значення у забезпеченні його існування і ліквідності, обслуговуванні і здійсненні всіх видів господарської діяльності [2]. Наявність і дефіцитного, і надлишкового рівня грошових коштів підприємства зумовлює негативний вплив на діяльність підприємства: перший проявляється у створенні загрози неплатоспроможності підприємства – навіть прибуткове підприємство може опинитися в ситуації неспроможності розраховуватись за поточними зобов’язаннями через нестачу вільних грошових коштів, другий – у збільшенні тривалості фінансового циклу підприємства та втрачених можливостях інвестування надлишкових грошових коштів [2].

Одним з головних аспектів вирішення окреслених проблем є створення ефективної систем управління фінансовими потоками, яка дозволяє максимізувати результативність обігу фінансових ресурсів і забезпечити їх достатній рівень у необхідний момент часу [1]. Від розробки та впровадження ефективних методів управління фінансовими потоками залежить не тільки ритмічність забезпечення фінансовими ресурсами виробничого процесу, але і підвищення інвестиційно-інноваційної активності промислових підприємств, що дозволить досягти підприємству своїх стратегічних цілей [1].

На сучасному етапі розвитку науки та технологій для моделювання маршрутизації фінансових потоків застосовуються різноманітні підходи, серед яких

найбільш поширеними є методи лінійного програмування, теорії мереж, а також статистичні та регресійні моделі. Традиційні математичні методи дозволяють формалізувати процеси руху фінансових ресурсів, враховувати обмеження і знаходити оптимальні рішення щодо розподілу коштів між різними підрозділами або напрямками діяльності. Проте, незважаючи на широке використання, класичні методи мають певні обмеження, пов'язані з статичністю та обмеженими можливостями врахування динамічних змін ринкової ситуації та складних взаємозв'язків.

Філософія бізнесу змінюється під впливом цифровізації зазнає змін і поведінка учасників електронної комерції [25]. Недостатньо просто відстежувати існуючі наміри, треба мислити ширше – як створити усі необхідні умови для оптимальної та стабільної роботи бізнесу, які будуть сприяти в подальшому зростанню [25].

Стрімкий розвиток інформаційних технологій, зокрема у сфері штучного інтелекту, машинного навчання і аналітики великих даних, суттєво розширює горизонти моделювання фінансових потоків. Застосування алгоритмів штучного інтелекту дозволяє створювати більш адаптивні і самонавчальні моделі, які здатні враховувати велику кількість параметрів і факторів, оперативно реагувати на зміни зовнішнього середовища та підвищувати точність прогнозів. Водночас аналітика великих даних відкриває можливості для обробки і аналізу неструктурованої інформації, що надходить з різних джерел, включаючи ринки, соціальні медіа, макроекономічні показники, що суттєво розширює інформаційну базу для прийняття рішень.

Градієнтний бустинг на сьогодні вважається одним із найуспішніших фреймворків машинного навчання для розв'язання як задач регресії, так і класифікації [17]. Початкова формулювання цього методу, відоме як Gradient Boosting Machine (GBM), було запропоноване Фрідманом понад два десятиліття тому [17]. Однак величезний прогрес у обчислювальних потужностях і ефективності, який ми спостерігаємо останнім часом, призвів до створення новітніх, передових варіантів градієнтного підсилення [17].

Крім того, інтеграція хмарних технологій та Інтернету речей забезпечує доступ до оновленої інформації в режимі реального часу, що дає змогу підприємствам оперативно коригувати маршрути фінансових потоків відповідно до поточних потреб і зовнішніх умов. Усе це підвищує гнучкість фінансового менеджменту і дає змогу своєчасно мінімізувати ризики, пов'язані з затримками чи неправильним розподілом ресурсів. Таким чином, сучасні технології суттєво трансформують підходи до моделювання та оптимізації фінансових потоків.

У перспективі розвиток моделей маршрутизації фінансових потоків буде все більше орієнтуватися на використання гібридних підходів, що поєднують класичні математичні методи з сучасними технологіями штучного інтелекту. Гібридні системи здатні підвищувати якість прогнозування і робити маршрутизацію фінансових потоків більш точним і адаптивним інструментом управління.

Одним із найбільш перспективних напрямів є застосування глибинних нейронних мереж та інших алгоритмів машинного навчання для аналізу великих обсягів фінансових даних і виявлення прихованих залежностей, які неможливо зафіксувати традиційними методами. Завдяки цьому підприємства отримують можливість більш точно оцінювати ризики, виявляти оптимальні шляхи розподілу фінансових ресурсів, а також прогнозувати майбутні фінансові потоки з урахуванням складної зовнішньої кон'юнктури.

Машинне навчання дає компаніям уявлення про тенденції поведінки клієнтів і бізнес-операційні моделі, а також підтримує розроблення нових продуктів [14]. Багато провідних сучасних компаній, таких як Facebook, Google і Uber, роблять машинне навчання центральною частиною своєї діяльності [14]. Машинне навчання стало значним конкурентоспроможним фактором для багатьох компаній.

Водночас усе більшого значення набувають системи маршрутизації фінансових потоків, що працюють у режимі реального часу і адаптуються до швидких змін зовнішніх і внутрішніх умов діяльності підприємства. Такі системи дають змогу оперативно реагувати на коливання ринку, зміни у платіжній дисципліні клієнтів,

появу нових фінансових інструментів або зміну умов кредитування. Це особливо важливо для підприємств, які працюють у висококонкурентних та мінливих галузях економіки, де від швидкості прийняття рішень залежить безпосередньо фінансовий результат.

Машинне навчання тісно пов’язане з обчислювальною статистикою, зосереджуваною на прогнозуванні за допомогою комп’ютерів, але не все машинне навчання є статистичним навчанням [14]. Вивчення математичної оптимізації надає методи, теорію та сфери застосування в галузі машинного навчання [14]. Інтелектуальний аналіз даних є спорідненою галуззю дослідження, що базується на дослідницькому аналізуванні даних за допомогою навчання [14].

Ще одним важливим аспектом розвитку моделей маршрутизації фінансових потоків є підвищення уваги до питань кібербезпеки. З зростанням обсягів цифрових фінансових операцій і поширенням електронних платіжних систем зростає ризик кібератак, шахрайства та витоку конфіденційної інформації. Тому сучасні моделі повинні містити механізми контролю, верифікації та захисту фінансових операцій, що гарантує безпеку руху коштів та мінімізує можливі втрати.

Захист інформації забезпечується суб’єктами переказу грошей шляхом обов’язкового впровадження та використання відповідної системи захисту [3].

Система захисту інформації повинна забезпечувати:

- цілісність інформації, що передається в платіжній системі, та компонентів платіжної системи;
- конфіденційність інформації під час її обробки, передавання та зберігання в платіжній системі;
- неможливість відмови ініціатора від факту передавання та отримувачем від факту прийняття документа на переказ, документа за операціями із застосуванням засобів ідентифікації, документа на відкликання;

- забезпечення постійного та безперешкодного доступу до компонентів платіжної системи особам, які мають на це право або повноваження, визначені законодавством України, а також встановлені договором [3].

Високі фінансові витрати на розробку і впровадження сучасних моделей часто стають перепоною для підприємств, особливо в умовах економічної нестабільності. Необхідно враховувати також питання відповідності моделей нормативним вимогам, що постійно змінюються і потребують гнучких підходів до адаптації.

Отже, розвиток моделювання маршрутизації фінансових потоків має великі перспективи, пов'язані із впровадженням новітніх технологій і підвищенням якості управління фінансами підприємств. Проте для успішної реалізації потенціалу цих моделей необхідно вирішити низку технологічних, організаційних та фінансових проблем. Майбутні дослідження повинні бути спрямовані на розробку ефективних гібридних моделей, удосконалення методів пояснювальної аналітики та створення безпечних і адаптивних систем для реального часу.

Таким чином, перспективи розвитку моделювання маршрутизації фінансових потоків створюють передумови для підвищення конкурентоспроможності підприємств та забезпечення їх стабільного фінансового стану в умовах сучасної економіки.

1.3 Аналіз та обґрунтування вибору моделей оптимального розподілу маршрутизації фінансових потоків

До моделювання маршрутизації фінансових потоків існує два декілька підходів, основні дослідження можна розділити на задачі класифікації, тобто оптимального розподілу за певними ознаками серед класів та обмежень, та задачі регресійного напряму, за яких прогнозується успішність транзакцій на основі моделей лінійної регресії. Розглянемо моделі, котрі використовуються для класифікації.

Важливо зазначити, що жодна окрема модель не є універсальною для всіх типів задач чи наборів даних. Кожна з них має свої сильні сторони та обмеження, які роблять її більш або менш придатною для конкретних сценаріїв. Тому вибір декількох різних моделей та їхнє порівняння дозволяє підвищити надійність та точність прогнозування, компенсуючи слабкі сторони одних алгоритмів за рахунок сильних сторін інших. Це забезпечує більш комплексний та стійкий підхід до моделювання складних фінансових систем.

Для побудови моделі оптимального розподілу фінансових потоків було обрано для огляду чотири ключові моделі класифікації: логістичну регресію, дерева рішень, випадковий ліс та градієнтний бустинг. Вибір цих методів зумовлений їх теоретичною обґрунтованістю, практичною ефективністю, а також здатністю до інтерпретації результатів.

Кожен із зазначених алгоритмів має свої переваги і сфери ефективного застосування, що дозволяє комплексно охопити різні аспекти задачі маршрутизації фінансових потоків – від базової лінійної класифікації до складного ансамблевого прогнозування. Далі розглянемо детально переваги та доцільність використання кожного з обраних методів.

Логістична регресія є популярним методом у дослідженнях та практиці з таких причин [12]:

– Ефективність і простота. Логістична регресія – це відносно простий, легко зрозумілий, але ефективний метод класифікації, який не потребує складних обчислень і зазвичай швидко навчається та реалізується [12].

– Гнучкість і розширюваність. Логістична регресія може бути легко адаптована до різноманітних завдань та досліджень з класифікації та аналізу даних, включаючи багатокласову класифікацію та регресійний аналіз [12]. Вона може бути розширена для врахування багатьох факторів, включення взаємодії між змінними, врахування нелінійних взаємодій між змінними, інтервальних даних, кластеризованих даних тощо[12].

– Добра прогностична здатність. Логістична регресія зазвичай добре прогнозує імовірності подій у випадку правильної специфікації моделі та використання адекватних змінних [12].

– Інтерпретованість результатів. Результати логістичної регресії можуть бути легко інтерпретовані, оскільки вони виражені у вигляді імовірностей [12]. Це дозволяє визначити вплив окремих змінних на результати та зробити висновки, які можуть бути використані в практичних діях [12].

– Ефективність та широке застосування на практиці. Логістична регресія часто виявляється дуже ефективною на практиці в багатьох галузях, таких як медицина, фінанси, маркетинг, біологія та інші, що робить цей метод корисним і універсальним інструментом аналізу даних [12]. Вона може успішно вирішувати різні класифікаційні завдання та прогнозувати результати з високою точністю [12]. Логістична регресія добре працює з реальними даними, навіть коли змінні не повністю задовольняють умови класичної лінійної регресії [12]. Вона може ефективно моделювати імовірності неперервних подій з категоріальними або бінарними вихідними даними [12].

– Розвиток технологій та методів. Завдяки постійному розвитку комп'ютерних технологій і методів аналізу даних, логістична регресія продовжує залишатися актуальною [12]. Вона може бути легко впроваджена в різні програмні середовища та бібліотеки, що дозволяє дослідникам та практикам використовувати її для розв’язування різноманітних завдань [12].

– Доступність програмних засобів. Існують різноманітні програмні пакети, такі як R, Python Scikit-learn, SAS, SPSS Statistics та інші, які надають зручні засоби для моделювання логістичної регресії, що робить цей метод доступним для широкого кола дослідників і практиків [12].

Однак у практичних умовах логістичної регресії часто недостатньо для моделювання складних або нелінійних залежностей, особливо у фінансових системах

з великою кількістю параметрів. Саме тому доцільно доповнювати її моделями, здатними виявляти складні логічні зв'язки в даних, зокрема деревами рішень.

Однією з причин вибору дерев рішень як базової моделі є їхня висока виразність. Дерева рішень здатні представляти будь-яку бінарну функцію, включаючи й лінійні залежності [15]. Формально їх можна подати у вигляді диз'юнкції кон'юнкцій – тобто переписати у диз'юнктивній нормальній формі (DNF), що дозволяє легко інтерпретувати правила прийняття рішень [15].

Це підкреслює здатність дерев рішень до моделювання складних умов у вигляді чітких логічних правил [15]. Така властивість є корисною при моделюванні рішень у фінансових системах, де важлива як точність, так і зрозумілість прийнятих рішень [15].

Одне з дуже важливих застосувань дерев рішень – це використання їх як ознак (фіч) [15]. Замість того, щоб використовувати дерево рішень для моделювання лише однієї цільової функції, ми можемо застосовувати багато невеликих дерев як ознаки [15]. Навчання одного великого дерева рішення часто призводить до перенавчання (overfitting) [15]. Натомість, ми прагнемо навчати невеликі дерева обмеженої глибини, використовуючи різні підмножини навчального набору[15]. Інший підхід -обрати підмножину з *n* ознак і навчити дерево рішень лише на цій підмножині[15].

Попри свою інтерпретованість, дерева рішень мають схильність до перенавчання, особливо при зростанні їх глибини. Для подолання цього недоліку і підвищення стабільності результатів, ефективним рішенням є використання ансамблевих методів. Одним із найпоширеніших і найнадійніших таких підходів є випадковий ліс.

Якщо «зростання» дерева не обмежити, то внаслідок цього буде побудоване складне дерево з великою кількістю вузлів й листя [14]. Як наслідок воно буде важко інтерпретованим [14]. Водночас вирішальні правила таких дерев, що створюютьвузли, в яких потрапляють два-три приклади, виявляються малозначними з практичного погляду [14].

Набагато краще мати дерево, що складається з малої кількості вузлів, яким відповідала б велика кількість прикладів з навчальної вибірки [14]. Тому цікавий підхід, альтернативний ранньому зупиненню – побудувати всі можливі дерева і вибрати те з них, яке за розумної глибини забезпечує прийнятний рівень помилки розпізнавання, тобто знайти найбільш вигідний баланс між складністю та точністю дерева [14].

Ми можемо розглядати ці маленькі дерева як "експертів": вони правильні лише на невеликій частині області визначення [15]. Кожен із таких експертів знає, як класифікувати приклад, повертаючи мітку {0, 1}[15]. Наприклад, якщо ми маємо 100 маленьких дерев рішень, то отримаємо 100 булевих оцінок для кожного прикладу, які можна розглядати як вектор ознак[15].

Цей вектор ознак можна потім подати на вхід іншому алгоритму машинного навчання, який буде навчатися на основі рішень цих експертів[15].

З обчислювальної точки зору Random Forest (випадкові ліси) є привабливими, оскільки вони:

- природним чином працюють як для задач регресії, так і для (багатокласової) класифікації;
- відносно швидко навчаються та роблять передбачення;
- потребують налаштування лише одного або двох параметрів;
- мають вбудовану оцінку помилки узагальнення (generalization error);
- можуть бути безпосередньо застосовані до задач з великою кількістю ознак (високовимірні задачі);
- легко реалізуються в паралельному режимі [16].

Зі статистичної точки зору Random Forest привабливі завдяки додатковим можливостям, таким як:

- оцінка важливості змінних;
- диференційоване зважування класів;
- заповнення відсутніх значень;

– логічна та зрозуміла візуалізація [16].

Якщо Random Forest забезпечує стабільність та узагальнення, то наступним кроком у підвищенні точності моделі є використання бустингових алгоритмів. Градієнтний бустинг забезпечує гнучкий механізм покрокового вдосконалення моделей за рахунок фокусування на попередніх помилках, що особливо важливо у високовартісних фінансових прогнозах.

Водночас, висока гнучкість дерев рішень робить їх вразливими до перенавчання, особливо на шумних або надмірно складних даних [15]. Щоб уникнути цього, доцільно використовувати ансамблеві методи, такі як градієнтне підсилення, яке передбачає послідовне навчання слабких моделей (неглибоких дерев рішень) і об'єднання їх у потужну композицію [15]. Один із способів об’єднати ці дерева рішень – це використати алгоритм підсилення (boosting), який ми розглянемо пізніше[15].

Градієнтний бустинг на сьогодні вважається одним із найуспішніших фреймворків машинного навчання для розв’язання як задач регресії, так і класифікації [17]. Початкова формулювання цього методу, відоме як Gradient Boosting Machine (GBM), було запропоноване Фрідманом понад два десятиліття тому [17]. Однак величезний прогрес у обчислювальних потужностях і ефективності, який ми спостерігаємо останнім часом, призвів до створення новітніх, передових варіантів градієнтного підсилення [17].

Обґрунтований вибір моделей машинного навчання є критично важливим кроком у побудові ефективної системи маршрутизації фінансових потоків. У даному дослідженні було обрано чотири підходи – логістичну регресію, дерева рішень, випадковий ліс і градієнтний бустинг – кожен з яких виконує певну функцію у загальній архітектурі моделі.

Логістична регресія слугує основою для створення інтерпретованих моделей, результати яких можна легко пояснити з позицій впливу змінних на кінцеве рішення. Її гнучкість, ефективність і поширеність у фінансовій сфері роблять її придатною для базового аналізу та початкової оцінки даних.

Дерева рішень дозволяють формалізувати логіку прийняття рішень у вигляді чітких умов і правил, що є особливо важливим у задачі маршрутизації фінансових потоків підприємства. Крім того, вони можуть бути використані для генерації ознак, що посилює ефективність інших моделей.

Random Forest, як ансамблевий метод, поєднує простоту дерев із високою стійкістю до перенавчання, забезпечує добру генералізацію і дозволяє працювати з великою кількістю змінних. Завдяки вбудованим механізмам оцінки важливості змінних та здатності працювати з неповними або шумними даними.

Градієнтний бустинг демонструє найкращі результати у задачах, де важлива точність і здатність моделі адаптуватися до складних закономірностей у даних. Його поетапна структура дозволяє поступово покращувати якість прогнозу, враховуючи слабкі місця попередніх моделей.

Таким чином, ми детально розглянули теоретичні засади та практичну доцільність використання логістичної регресії, дерев рішень, випадкового лісу та градієнтного бустингу для моделювання маршрутизації фінансових потоків. Кожна з цих моделей має свою унікальну логіку та переваги, що робить їх цінними інструментами в арсеналі фахівця з машинного навчання. Для повного розуміння їх функціонування та подальшого застосування у нашому дослідженні, критично важливо заглибитись у їхній математичний апарат. Далі ми перейдемо до формального опису кожної моделі та їх можливих оцінок.

Логістична регресія – це статистичний регресійний метод керованого машинного навчання для бінарної класифікації даних з метою виявлення їх належності одному із двох класів, коли коли вхідна змінна задається набором ознак об’єкта, а вихідна змінна (результат аналізу) є двійковою величиною, тобто може набувати значення 0 або 1 [12].

Математичне представлення формули логістичної регресії можна розглядати як:

$$P\big((y = 1|x\big) = \sigma(w^T x + b) = \frac{1}{1 + e^{-(w^T x + b)}}, \tag{1.1}$$

де:

$x \in R^n$ –вектор ознак,

$w \in R^n$ –вагові коефіцієнти,

b –зсув,

σ –функція активації,

$P(y = 1|x)$ – ймовірність того, що об'єкт належить до класу 1.

Візуалізувати логістичну регресію можна як ймовірність виникнення події (рис. 1.3).

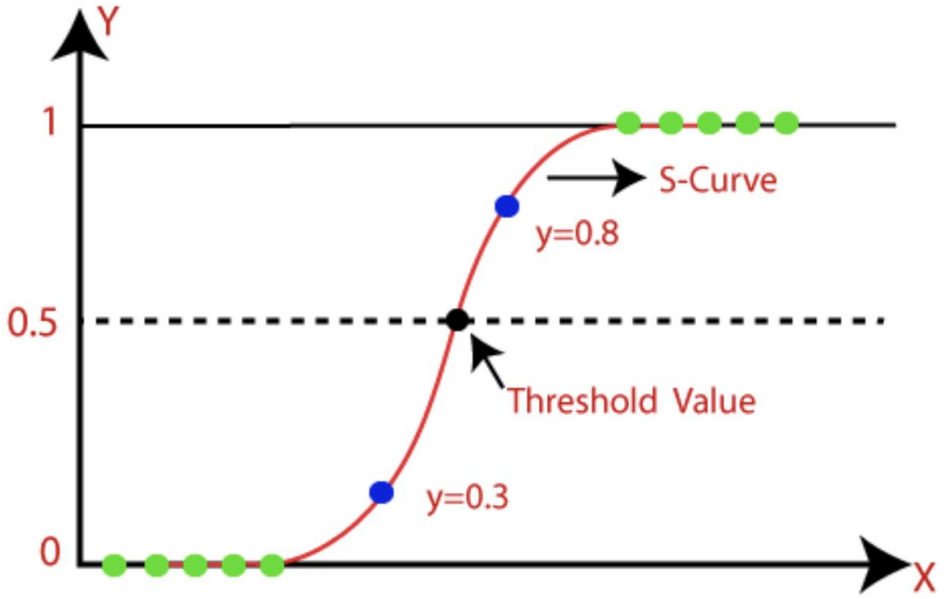

Рисунок 1.3 – Графічне відтворення моделі логістичної регресії

Джерело: [13]

Дерево рішень – це модель, що представляє собою метод подання вирішальних правил в ієрархічній структурі, що складається з елементів двох типів – вузлів (node) та листя (leaf). У вузлах містяться вирішальні правила і проводиться перевірка відповідності прикладів цього правила за яким-небудь атрибутом навчальної множини [14].

Існує декілька критеріїв розщеплення дерев рішень, розглянемо детальніше

1. Ентропія:

$$H(S) = -\sum_{i=1}^{k} p_i \log_2 p_i \,, \tag{1.2}$$

де:

$H(S)$ –ентропія, що характеризує ступінь невизначеності розподілу класів,

k – кількість класів,

p_i – ймовірність появи об'єктів класу i у множині S

Візуалізація принципу ентропії, у наборі з двома мітками ліва синя смужка показує ентропію, а чорні смужки – частки кожної з міток.

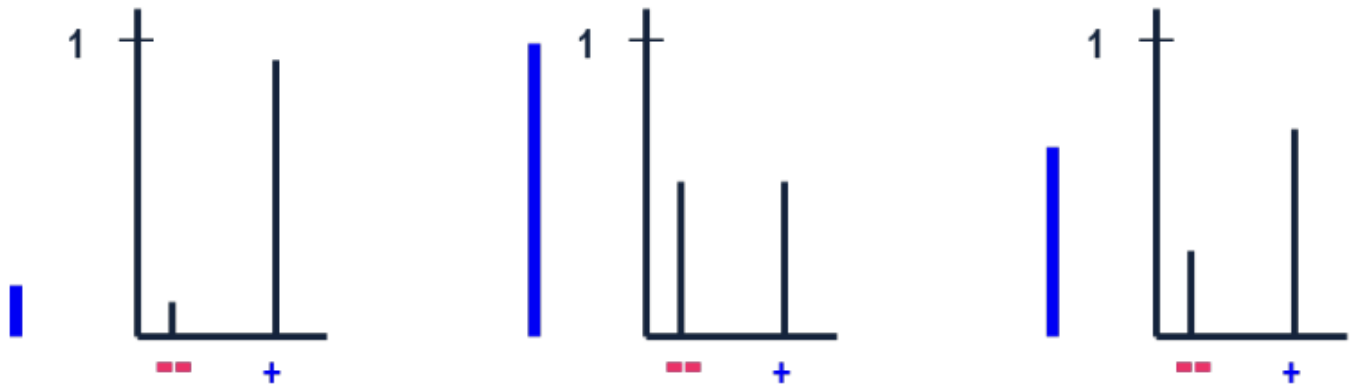

Рисунок 1.4 – Графічне представлення принципу ентропії

Джерело: [15]

2. Інформаційний приріст:

$$IG\ (S, A) = H(S) - \sum_{v\,\epsilon\,Values\ (A)} \frac{|S_v|}{|S|} H(S_v)\,, \tag{1.3}$$

де:

$IG\ (S, A)$ – інформаційний приріст при розщепленні множини S за ознакою A,

$H(S)$ – ентропія всієї множини до розщеплення,

$Values\ (A)$ – множина всіх можливих значень ознаки A,

S_v— підмножина S, у якій ознака A має значення v;

3. Індекс Джині:

$$Gini\ (S) = 1 - \sum_{i=1}^{k} p_i^2\,, \tag{1.4}$$

де:

k – кількість класів,

p_i – ймовірність появи об'єктів класу i у множині S

Random Forest – ансамблева модель, яка об'єднує кілька дерев рішень для покращення точності та зменшення перенавчення [16].

$$\hat{y} = \frac{1}{T}\sum_{t=1}^{T} h_t(x)\,, \tag{1.5}$$

де

$h_t(x)$ –прогноз дерева t,

T – кількість дерев

Gradient Boosting — це потужний ансамблевий метод машинного навчання, що будує передбачувальну модель у вигляді послідовності слабких моделей (зазвичай рішень дерев), де кожна наступна модель коригує помилки попередньої. Кожне нове дерево в ансамблі навчається на залишках або негативному градієнті функції втрат, що забезпечує поступове зменшення помилок моделі [17].

Модель будується ітеративно, починаючи з першої моделі $F_0(x)$, яка мінімізує початкову функцію втрат L:

$$F_0(x) = \min_{\gamma} \sum_{i=1}^{n} L(y_i,\ \gamma)\,, \tag{1.6}$$

де:

$F_0(x)$ –початкова модель,

γ – константа для мінімізації функції втрат ,

$L(y_i, \gamma)$ – функція втрат між справжнім значенням y_i та передбаченим

Потім модель вдосконалюється в кожній ітерації m, додаючи нове дерево $h_m(x)$

$$F_m(x) = F_{m-1}(x) + v * h_m(x) \tag{1.7}$$

де

$F_m(x)$ – модель на кроці m,

$F_{m-1}(x)$ – модель на попередній ітерації,

$h_m(x)$ – нове дерево рішення, котре вчиться наближати градієнт витрат

$$g_i^{(m)} = -\left[\frac{\partial L\big(y_i, F(x_i)\big)}{\partial\, F(x_i)}\right]_{F=F_{m-1}} \tag{1.8}$$

де

$g_i^{(m)}$ – значення негативного градієнта функції втрат для i на ітерації m,

$L(y_i, F(x_i))$ – функція втрат,

$F(x_i)$ - передбачене значення моделі для x_i на поточній ітерації.

Розглянемо функції втрат котрі використовуються для задач багатокласової класифікації:

Multi-class log-loss – застосовується при кількості класів більше 2 [18]

$$L(y, \hat{p}) = -\sum_{k=1}^{K} y_k \log(\hat{p}_k) \tag{1.9}$$

де

k – кількість класів,

y_k – індикатор істинного класу,

$\hat{p}_k = P(y = k|x)$ – ймовірність для класу k

Exponential loss – модель на основі експоненти

$$L(y, \hat{y}) = -\exp(-y\hat{y}) \tag{1.10}$$

де:

y – істинна мітка моделі

$\hat{y}$ – прогнозоване значення моделі

Hinge loss – не обирає модель, якщо вона дає передбачення на неправильній стороні гіперплощини або занадто близько до неї [18]

$$L(y, \hat{y}) = max(0, 1 - y\hat{y}) \tag{1.11}$$

де

y – істинна мітка моделі

$\hat{y}$ – прогнозоване значення моделі

Розглянемо детальніше моделі оптимізації маршрутизації фінансових потоків, котрі базуються на регресії. Лінійна регресія встановлює лінійний зв'язок між вхідними ознаками та цільовою змінною. Вона передбачає, що залежність між змінними можна описати гіперплощиною у багатовимірному просторі [19].

$$\hat{y} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n \tag{1.12}$$

де

$\hat{y}$ – прогнозоване значення моделі,

x_1 – значення змінної,

β_0 – зміщення,

β_1 – ваговий коефіцієнт

Графічне представлення моделі лінійної регресії зображено на рис. 1.5

Ridge Regression є модифікацією лінійної регресії, яка додає L2 регуляризаційний член для контролю за величиною коефіцієнтів[20]. Lasso (Least Absolute Shrinkage and Selection Operator) також додає штраф до функції втрат, але використовує L1-норму, що має важливу властивість – може занулювати деякі коефіцієнти, фактично здійснюючи відбір ознак [21].

Вибір між Ridge та Lasso залежить від специфіки задачі: якщо очікується, що всі ознаки мають вплив на результат, доцільно застосовувати Ridge Regression; якщо ж передбачається, що лише частина змінних є інформативною, ефективнішим буде використання Lasso, оскільки воно сприяє побудові більш інтерпретованих моделей з меншою кількістю ознак.

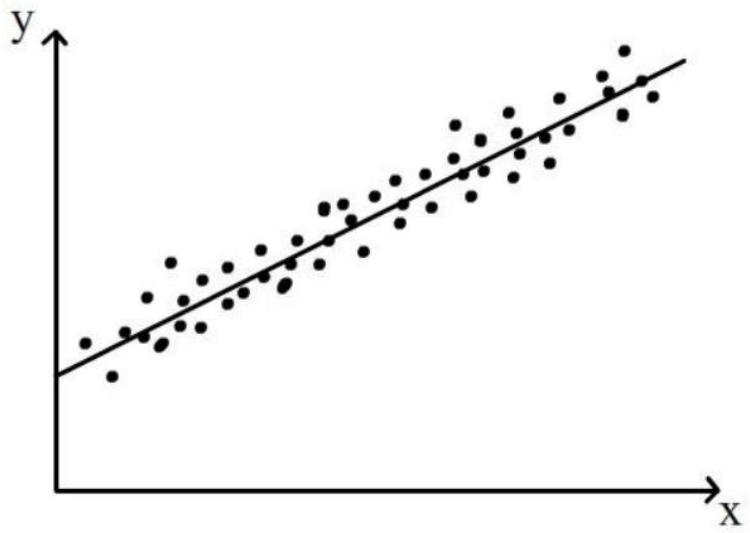

Рисунок 1.5 – графічне представлення лінійної регресії

Джерело: [22]

Найдоцільнішим підходом є використання саме класифікаційних моделей, що дозволяють максимально ефективно передбачати клас (напрямок) для кожної нової транзакції на основі набору ознак. На основі аналізу математичних формул і властивостей моделей, було обрано чотири найбільш поширені класифікаційні підходи. Для кожної моделі було проаналізовано варіанти ключових параметрів, які впливають на її якість, та обрано оптимальний варіант для подальшого використання.

На відміну від задач регресії, де передбачається числове значення, маршрутизація формально є задачею класифікації: кожна транзакція відноситься до одного з наперед визначених класів. Відповідно, доцільним є використання саме класифікаційних моделей, які забезпечать структуризацію потоків відповідно до заданих класів, що в подальшому допоможе визначити оптимального провайдера та ефективний перерозподіл транзакцій.

Для систематизації характеристик класифікаційних моделей наведено порівняльну таблицю (табл. 1.2):

Таблиця 1.2 – Порівняння характеристик класифікаційних моделей

Модель	**Тип моделі**	**Інтерпретованість**	**Чутливість до шуму**	**Стійкість до перенавчання**	**Потреба в налаштуванні**	**Функція втрат**
Логістична регресія	Лінійна	Висока	Висока	Низька	Мінімальна	Multi-class log-loss (1.9)
Дерево рішень	Ієрархічна	Висока	Середня	Низька	Низька	Індекс Джині (1.4)
Random Forest	Ансамблева	Середня	Низька	Висока	Середня	Індекс Джині (1.4)
Gradient Boosting	Ансамблева	Низька	Низька	Висока	Висока	Multi-class log-loss (1.9)

Джерело: розроблено автором на основі [23]

У рамках логістичної регресії, яка є базовою інтерпретованою лінійною моделлю, як функцію втрат було розглянуто логарифмічну функцію, що дозволяє моделі прогнозувати ймовірність належності до певного класу. Оскільки задача має багатокласовий характер, було обрано функцію втрат Multi-class log-loss, що описується формулою (1.9) та враховує імовірності по всіх класах одночасно. Такий підхід дозволяє точніше моделювати складну структуру розподілу класів і краще підходить до умов задачі маршрутизації.

Для моделі дерева рішень розглядалися декілька критеріїв розщеплення вузлів: ентропія, інформаційний приріст та індекс Джині. Формула ентропії (1.2) дозволяє виміряти невизначеність у розподілі класів, а інформаційний приріст (1.3) визначає ефективність розщеплення. Однак найбільш практичним і стабільним критерієм для обробки великих обсягів даних є індекс Джині (1.4).

У моделі Random Forest, яка об'єднує в собі велику кількість дерев рішень та виконує голосування для отримання фінального прогнозу, було доцільно зберегти однакову методологію оцінки вузлів, як і в базовому дереві. У кожному дереві

ансамблю як критерій розщеплення також обрано індекс Джині, що дозволяє зберігати узгодженість між деревами та зменшити переобучення при достатній кількості дерев.

У моделі Gradient Boosting, яка поступово навчає послідовність слабких моделей (здебільшого дерев рішень), було розглянуто кілька функцій втрат: експоненційну (1.10), hinge loss (1.11) та multi-class log-loss (1.9). Експоненційна функція втрат дозволяє сильніше штрафувати великі помилки, але схильна до перенавчання, тоді як hinge loss зазвичай застосовується в методах опорних векторів і не є оптимальною для ймовірнісного підходу. Найбільш релевантною для багатокласової класифікації є функція втрат Multi-class log-loss (1.9), яка забезпечує максимальну відповідність математичній постановці задачі, а також дозволяє ефективно керувати градієнтним оновленням моделі (1.8).

Порівняльний аналіз логістичної регресії, дерева рішень, випадкового лісу та градієнтного бустингу показав, що кожна з моделей має власні переваги та обмеження, які слід враховувати при проєктуванні системи маршрутизації. Логістична регресія відзначається високою інтерпретованістю та швидкістю, дерево рішень – зручністю у побудові логічних правил, Random Forest – балансом між точністю та стійкістю, а Gradient Boosting – найвищими показниками якості моделювання.

Особливу увагу було приділено не лише математичним характеристикам моделей, але й їх практичній доцільності – можливості інтеграції у реальні бізнес процеси, адаптивності до навантаження, швидкості прийняття рішень та потребі в обчислювальних ресурсах. Ефективна маршрутизація фінансових потоків може бути забезпечена шляхом цілеспрямованого вибору або комбінування моделей відповідно до специфіки даних, інфраструктури підприємства та стратегічних цілей. Запропоновані підходи створюють надійну основу для побудови адаптивної системи управління фінансовими транзакціями в умовах сучасної цифрової економіки.

Висновки до першого розділу

У першому розділі було розглянуто сутність процесу управління та оптимізації маршрутизації фінансових потоків. Ми визначили важливість роботи з надавачами послуг проведення платежів та їх роль у міжнародній економіці. Було розглянуто найбільші надавачі послуг проведення платежів у світі, такі як Fiserv, Stripe, Revolut, Paypal та Solidgate.

Для кожної з компаній було розглянуто обсяг проведених фінансових потоків за 2024 рік, розглянуто особливості проведення бізнесу, роботи з маршрутизацією фінансових потоків та виклики з котрими надавачі послуги проведення платежів стикаються під час маршрутизації фінансових потоків.

Було розглянуто сучасні підходи до моделювання маршрутизації фінансових потоків підприємства. Проаналізовано переваги використання економіко-математичних методів у поєднанні з інструментами штучного інтелекту, машинного навчання, аналітики великих даних і хмарних технологій.

Визначено перспективність розвитку інтелектуальних систем управління потоками в реальному часі, а також окреслено актуальні виклики, пов’язані з кібербезпекою, нормативною відповідністю та фінансовими витратами.

Проаналізовано два основних підходи до економіко-математичного моделювання маршрутизації фінансових потоків підприємства, як задача класифікації та регресії. Було виділено основні моделі, котрі застосовуються для кожного з підходів.

Для задачі класифікації було розглянуто математичне представлення логістичної регресії, дерева рішень з трьома різними підходами до розщеплення, а саме ентропії, інформаційного приросту та за індексом джині. Оптимальних підходом до розщеплення було обрано індекс джині. Описано математичний вигляд моделі випадкового лісу та градієнтного бустингу. Для моделі градієнтного бустингу було

деталізовано перехід між ітераціями моделі, та формалізовано формулу градієнту витрат, котрі наближаються за кожної ітерації.

Для задачі на основі регресійного підходи було формалізовано математичний вигляд лінійної регресії та її ітерацій, таких як Ridge та Lasso регресія. Проте для подальшого дослідження після проведення аналізу було вирішено обрати основні моделі класифікації та у подальшому порівняти їх результати.

Розглянуто функції втрат, на основі лагу, експоненційної функції та hinge loss. Оптимальною для використання було встановлено мультикласової log-loss функції втрат.

2 АНАЛІЗ ПЕРЕДУМОВ ТА ПОСТАВКА ЕКОНОМІЧНОГО ЗАВДАННЯ МОДЕЛЮВАННЯ МАРШРУТИЗАЦІЇ ФІНАНСОВИХ ПОТОКІВ

2.1 Аналіз економічних показників підприємства та організації маршрутизації фінансових потоків

Юридична особа ТОВ "ТК ВАЛЕНСІЯ", код ЄДРПОУ 43043985, було зареєстровано 06.06.2019. Основна спеціалізація за КВЕД компанії полягає у:

46.90 Неспеціалізована оптова торгівля;

10.89 Виробництво інших харчових продуктів, н.в.і.у. ;

11.07 Виробництво безалкогольних напоїв; виробництво мінеральних вод та інших вод, розлитих у пляшки;

46.19 Діяльність посередників у торгівлі товарами широкого асортименту;

47.11 Роздрібна торгівля в неспеціалізованих магазинах переважно продуктами харчування, напоями та тютюновими виробами;

47.91 Роздрібна торгівля, що здійснюється фірмами поштового замовлення або через мережу інтернет;

73.20 Дослідження кон'юнктури ринку та виявлення громадської думки [24].

Сучасна компанія, котра займається виробництвом власної продукції та імпортом та збутом продукції за посередництва оптом та у роздріб. Основний спосіб збуту полягає у рекламі продукції у соціальних мережах, виробі відповідного товару на веб-сайті компанії, оплаті засобами онлайн-платежів та доставка поштою до кінцевого споживача, збір та аналіз інформації про споживачів.

Підприємство вкладає максимальні зусилля у будову онлайн інфраструктури основою якої є зручність для користувача та безперебійний доступ до системи проведення онлайн-платежів. Підприємство працює у місті Києві та має офіс за вулицею Ризькою. Підприємство використовує послуг партнерів на основі консалтингу та незалежних експертів для роботи над розробкою веб-сайту компанії,

зберігання даних на серверах та їх підтримки, розробки дизайнів для рекламних компаній та соціальних мереж.

Підприємство обрало шлях розвитку у сфері електронної комерції, або е-бізнесу. Розвиток е-бізнесу означає перехід до інформаційного простору основних бізнес-процесів і каналів зв'язку, а це рано чи пізно відіб'ється на діяльності всіх підприємств. У вузькому розумінні е-бізнес – це перетворення бізнес-процесів із застосуванням Інтернет-технологій, що дозволяє досягти вищої продуктивності [25].

Основним показником для аналізу фінансово-економічної діяльності підприємства було обрано кількість спроб проведення онлайн платежу, кількість успішних платежів та відсоток успішності за 2023 та 2024 роки. Детальна статистика по місяцям була розроблена на основі даних наданих підприємством та подана у табл. 2.1.

Таблиця 2.1 – Економічні показники руху фінансових потоків підприємства ТОВ «ТК Валенсія»

	Кількість спроб проведення платежів		**Кількість успішних платежів**		**Відсоток успішних платежів у загальній кількості**	
Місяць/Рік	**2024**	**2023**	**2024**	**2023**	**2024**	**2023**
1	131	283	129	117	98.47%	41.34%
2	178	131	164	63	92.13%	48.09%
3	211	198	175	118	82.94%	59.49%
4	321	295	294	214	91.59%	72.46%
5	272	237	216	160	79.41%	67.53%
6	286	259	273	184	95.45%	71.18%
7	382	334	286	226	74.87%	67.70%
8	430	379	371	259	86.28%	68.43%
9	693	594	178	391	25.69%	65.84%
10	821	718	451	354	54.93%	49.31%
11	798	687	365	456	45.74%	66.31%
12	643	558	218	373	33.90%	66.97%

Джерело: Складено автором на основі внутрішніх даних підприємства

Можемо спостерігати сезонність продажів асортименту та зростання попиту на продукцію, основний пік припадає на вересень, жовтень та листопад у 2024 та 2023 роках відповідно. Цей період зазвичай збігається із закінченням літнього сезону та

початком активної фази закупівель перед зимою, а також підготовкою до новорічних свят. Відповідно, компанія «ТК Валенсія» фіксує значне зростання кількості транзакцій у ці місяці, що підтверджується даними (рис. 2.1).

Кількість спроб проведення платежів у 2024 році зростає значно стрімкіше порівняно з 2023 роком. Приріст видно у вересні, 693 спроби в 2024 році проти 594 у 2023 та жовтні 821 проти 718 відповідно. Загалом, майже щомісячно показники 2024 року перевищують результати попереднього року, що свідчить про позитивну динаміку розвитку бізнесу.

Проте, паралельно зі зростанням кількості транзакцій спостерігається певна нестабільність у відсотку успішних платежів, зокрема у другій половині року. У 2024 році після відносно високих показників у липні 86,28%, вже у вересні відсоток успішних транзакцій різко знижується до 25,66%, а в грудні сягає лише 33,9%. Аналогічна тенденція, хоч і менш виражена, спостерігається і в 2023 році: у жовтні – 66,31%, а в грудні – 66,97%. Зниження успішності платежів у ці місяці, ймовірно, пов’язане з нестабільною енергетичною ситуацією в країні, включаючи графіки аварійного відключення електроенергії та порушення зв’язку внаслідок масових ракетних ударів по критичній інфраструктурі.

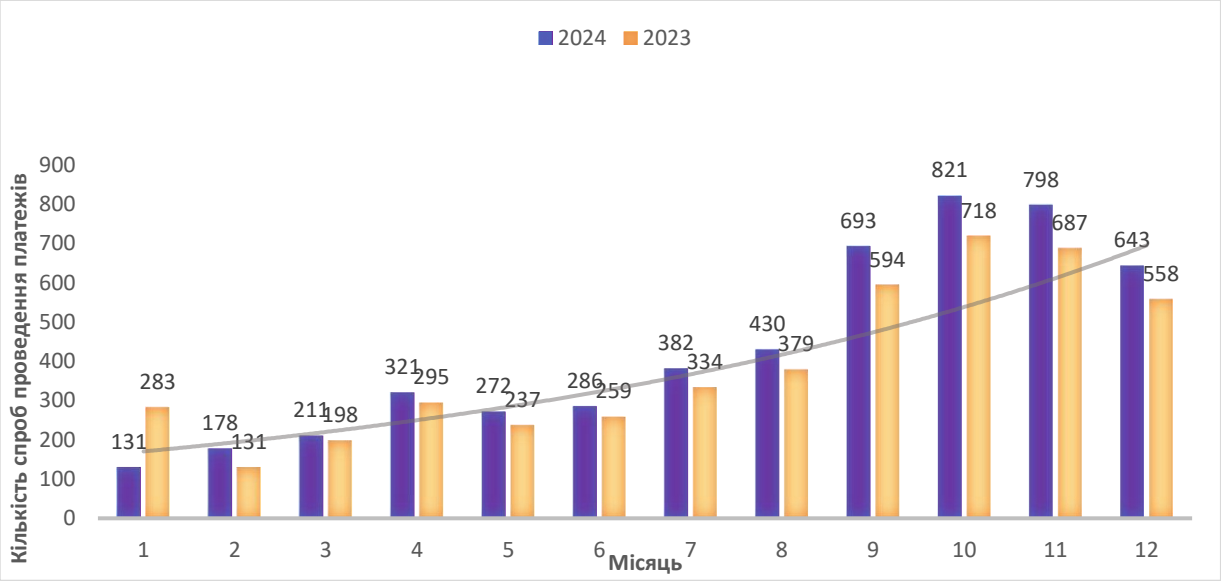

Рисунок 2.1 – Кількість спроб проведення платежів за місяцями

Джерело: розраховано автором на основі даних підприємства.

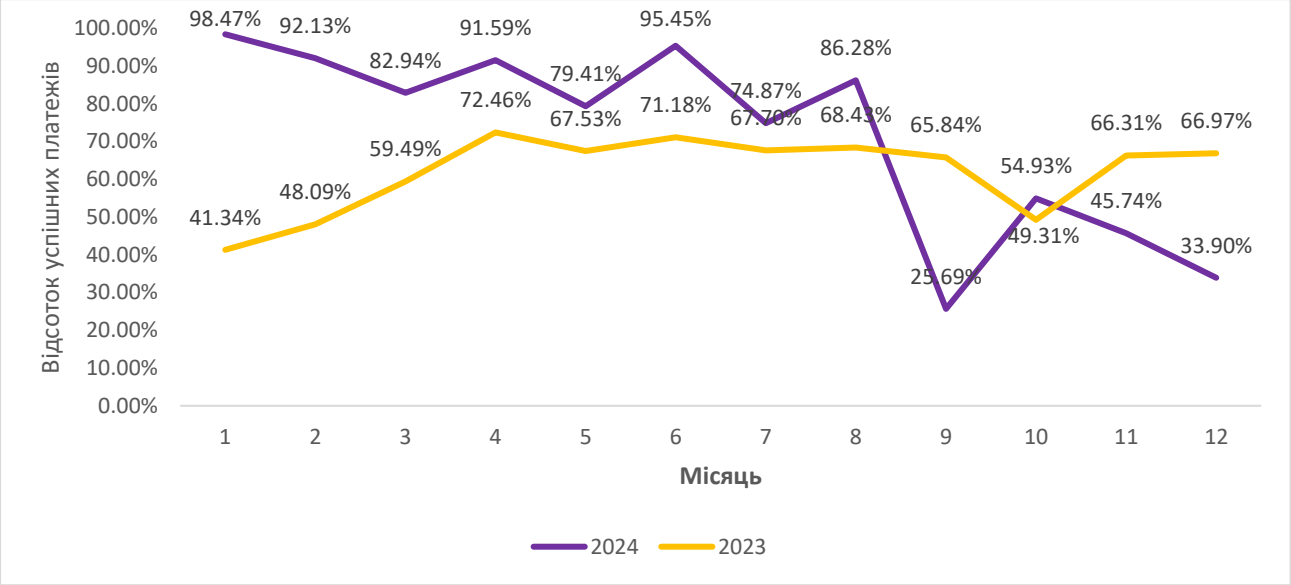

Рисунок 2.2 – Відсоток успішних платежів у загальній кількості ТОВ "ТК Валенсія"

Джерело: розраховано автором на основі даних підприємства.

Розглянемо економічну інформацію підприємства щодо загальної суми успішно проведених платежів, витрат на продукцію, витрат на проведення платежів, податки та загальний прибуток.

Таблиця 2.2 – Економічні показники прибутковості підприємства ТОВ «ТК Валенсія»

	Загальна сума успішно проведених платежів		**Витрати на продукцію**		**Витрати на проведення платежів**		**Прибуток**	
Місяць	**2024**	**2023**	**2024**	**2023**	**2024**	**2023**	**2024**	**2023**
1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
1	205,626	95,238	95,847	63,297	9,921	3,067	99,858	28,874
2	247,148	54,747	118,244	32,508	11,925	1,571	116,979	20,668
3	338,975	105,964	128,625	63,012	16,356	2,682	193,994	40,270
4	574,476	182,459	221,088	115,586	27,718	4,218	325,670	62,656
5	271,512	148,327	157,464	84,415	13,100	3,506	100,948	60,406
6	367,731	160,257	201,201	99,031	17,743	3,728	148,787	57,498
7	546,832	218,902	210,210	122,441	26,385	5,133	310,237	91,329
8	751,646	211,364	286,041	134,434	36,267	5,064	429,338	71,867
9	270,560	320,729	135,280	201,433	13,055	7,835	122,225	111,460
10	489,786	290,280	338,701	192,222	23,632	8,353	127,453	89,705

Продовження табл. 2.2

1	2	3	4	5	6	7	8	9
11	427,415	434,325	270,465	244,279	20,623	10,307	136,327	179,738
12	283,618	361,832	163,500	200,893	13,685	8,526	106,433	152,413

Джерело: розраховано автором на основі даних підприємства.

У 2024 році загальна сума успішно проведених платежів зросла у 2–3 рази порівняно з аналогічними місяцями 2023 року, що безпосередньо вплинуло на зростання прибутку. Рентабельність платежів також покращилась: якщо в січні 2023 року вона становила 30,3%, то в січні 2024 — вже 48,6%. Це свідчить про не лише кількісне, але й якісне покращення фінансової ефективності діяльності підприємства.

У 2023 році підприємство співпрацювало з одним надавачем послуг щодо проведення платежів, за умовами:

Вартість проведення неуспішного платежу – 7 грн.

Вартість проведення успішного платежу – 2% від загальної вартості платежу.

Проте з метою підвищення ефективності проведення платежів та покращення економічних результатів підприємство у 2024 році уклало договір з трьома новими надавачами послуг проведення платежів, за відповідними умовами:

Провайдер «А»:

Вартість проведення неуспішного платежу – 0 грн.

Вартість проведення успішного платежу – 5% від загальної вартості платежу.

Провайдер «Б»:

Вартість проведення неуспішного платежу – 0 грн.

Вартість проведення успішного платежу – 5.5% від загальної вартості платежу.

Провайдер «В»:

Вартість проведення неуспішного платежу – 0 грн.

Вартість проведення успішного платежу – 4% від загальної вартості платежу.

З метою забезпечення стабільної роботи сервісу оплати на веб сайті підприємством було встановлено автоматичний розподіл транзакцій за коефіцієнтами А: 30%, Б: 35%, В: 35%.

Внесені зміни значно вплинули на рентабельність підприємства, а саме підвищило на 6% у січні 2024 відносно грудня 2023. Це свідчить про доцільність використання декількох надавачів послуг проведення платежів.

Розглянемо загальний відсоток успішних транзакцій за провайдерами відповідно до дати створення транзакції у 2024 році (див. рис. 2.3). Відповідно до отриманих даних, можемо спостерігати що у періоди стабільної ситуації провайдери мають схожі показники 98% відсотків у січні 2024 року. Проте у місяці коли існують проблеми можуть спостерігати різницю у відсотку успішних транзакцій до 13 відсотків.

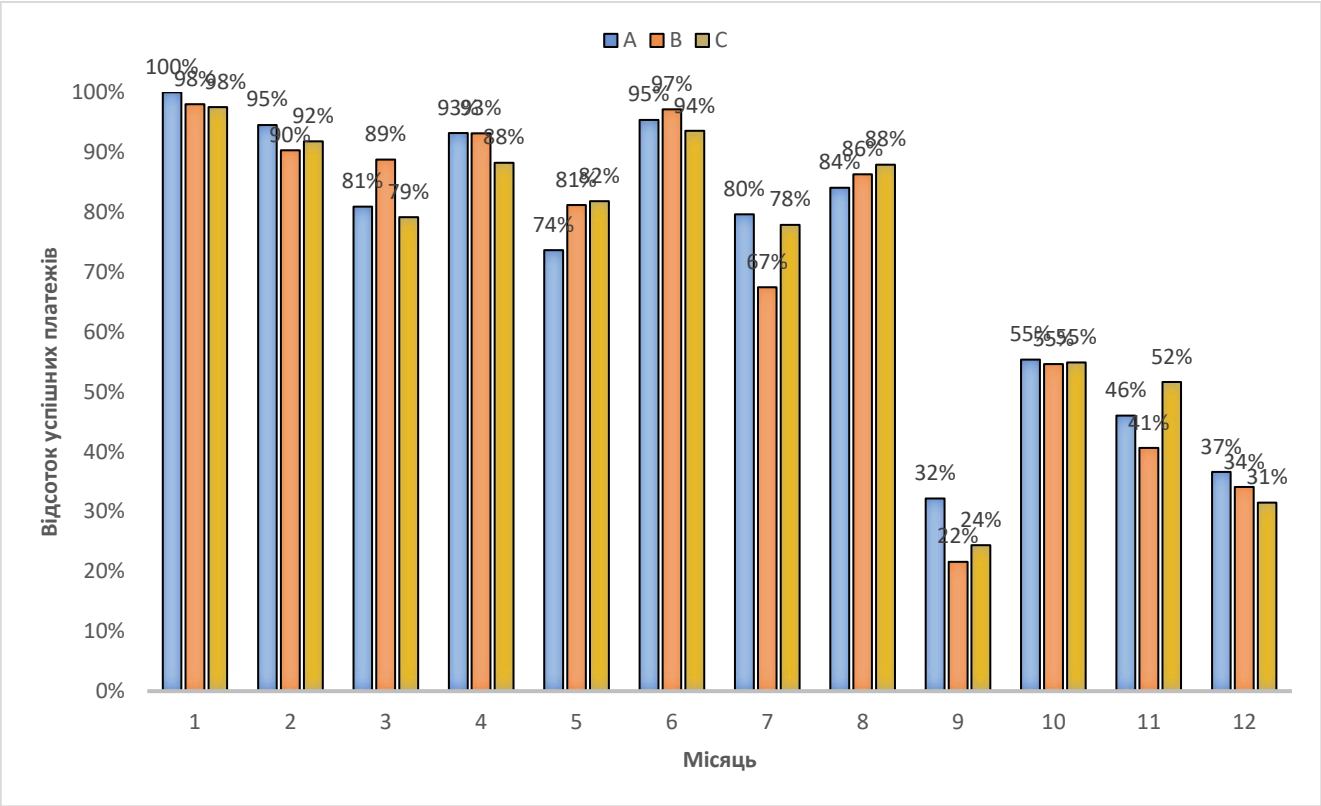

Рисунок 2.3 – Відсоток успішних платежів у загальній кількості

Джерело: розраховано автором на основі даних підприємства.

У березні провайдер С схвалив 79% транзакцій, в той час як провайдер В схвалив найбільше, 89% транзакцій. У липні 2024 року провайдер В схвалив найменший відсоток транзакцій, 67%, та найбільшим став відсоток у провайдера А. Отримані дані свідчать про неефективність постійного розподілу встановленого підприємством та показають необхідність побудови динамічного розподілу транзакцій.

Розглянемо показники успішності транзакцій за джерелом походження трафіку на наявність залежності (табл. 2.3). Можемо спостерігати що транзакції між джерелами розподілені рівномірно та джерело походження користувача не суттєво впливає на відсоток успішних транзакцій як у 2023 так і 2024 роках. можемо зробити висновок, що загальна кількість спроб проведення платежів зростає незалежно від джерела трафіку. Зокрема, усі чотири основні канали – Facebook, Google Ads, Instagram та органічний пошук – демонструють помірне зростання у 2024 році порівняно з 2023.

Відсоток успішних платежів демонструє певну нестабільність. Google Ads, незважаючи на зростання кількості спроб, показує зниження успішності з 64.39% до 60.76%. Аналогічна ситуація спостерігається і для Instagram: у 2024 році відсоток успішних транзакцій становить 57.76%, що на 4.21% менше, ніж у 2023 році. Така динаміка може свідчити про зміну якості трафіку, коли значна частина нових користувачів не завершує оплату з різних причин — наприклад, технічні складнощі, недовіра до платформи або недостатня готовність до покупки. Також бачимо, що у 2024 році найбільш стабільним і ефективним каналом за показником успішних транзакцій залишається Facebook (62.18%), тоді як найнижчий результат демонструє Instagram (57.76%). Це може бути наслідком особливостей аудиторії кожної платформи.

Таблиця 2.3 – Показники успішності обробки фінансових потоків за джерелом походження клієнту

	Загальна кількість спроб проведення платежів		**Відсоток успішних платежів у загальній кількості**	
Traffic Source	2023	2024	2023	2024
Facebook	1197	1346	61.15%	62.18%
Google Ads	1129	1269	64.39%	60.76%
Instagram	1199	1314	61.97%	57.76%
Search	1148	1237	62.11%	60.87%

Джерело: розраховано автором на основі даних підприємства.

Отже, вплив джерела є мінімальним на ймовірність успішності платежу, та варіюється у часі, що вказує на незначний вплив походження реклами на ймовірність успішної оплати клієнтом та підтвердження необхідності покращення стабільності послуг проведення платежів.

2.2 Побудова економіко-математичне моделі оптимізації фінансових потоків підприємства

Задача економіко-математичного моделювання полягає у розробці комплексної системи взаємодії метрик щодо ефективності роботи кожного з провайдерів на основі моделей штучного інтелекту для розробки системи адаптивного автоматичного розподілу транзакцій у режимі реального часу.

У виконання даної задачі мають бути використані категоріальні дані, такі як джерело трафіку, місто та область доставки, статус та провайдер. Також для дослідження мають бути розраховані додаткові метрики, такі як відсоток успішності транзакцій за провайдером за період, швидкість отримання рішень щодо статусу провайдером. Також рекомендовано розрахувати частину дня у яку створюється заявку на проведення платежу, адже проблеми з роботою провайдерів можуть мати

тренд саме за частинами дня на котрі приходиться найбільше навантаження на енергосистему.

За базову модель може бути обрана логістична регресія. Логістична регресія є базовою моделлю бінарної класифікації, що дозволяє визначити ймовірність належності об’єкта до одного з двох класів [26]. У випадку аналізу транзакцій логістична регресія дозволяє передбачити ймовірність успішності транзакції або своєчасності обробки замовлення.

Вибір правильних параметрів та налаштувань для логістичної регресії є ключовим для побудови ефективної та узагальнювальної моделі. Хоча логістична регресія є відносно простою моделлю порівняно з деякими алгоритмами машинного навчання, все ж є важливі аспекти, які потрібно врахувати. По-перше, необхідно підготувати дані, очистити, працювати з викидами, пропущеними даними, провести кодування категоріальних даних, масштабування та стандартизацію ключових ознак.

Провести вибір найбільш релевантних ознак, що зменшує складність моделі, підвищує її інтерпретованість та може покращити продуктивність. Використовувати статистичні показники (наприклад, кореляція, хі-квадрат тест, взаємна інформація) для оцінки зв'язку між ознакою та цільовою змінною.

Оскільки модель регресійного типу, вона часто потребує регуляризації, тобто запобіганню перенавчання (overfitting) шляхом додавання "штрафу" до функції втрат (log-loss), який обмежує розмір коефіцієнтів моделі, що допомагає боротися з мультиколінеарністю в системі характеристик.

Логістична регресія оптимізує функцію втрат (log-loss) за допомогою різних алгоритмів. Вибір розв’язчика залежить від розміру даних та типу регуляризації, в програмних бібліотеках є певні реалізації та налаштування відповідних параметрів.

Однак у практичних умовах логістичної регресії часто недостатньо для моделювання складних або нелінійних залежностей, особливо у фінансових системах з великою кількістю параметрів. Саме тому доцільно доповнювати її моделями, здатними виявляти складні логічні зв’язки в даних, зокрема деревами рішень.

Однією з причин вибору дерев рішень як базової моделі є їхня висока виразність. Дерева рішень здатні представляти будь-яку бінарну функцію, включаючи й лінійні залежності [15]. Формально їх можна подати у вигляді диз'юнкції кон'юнкцій – тобто переписати у диз'юнктивній нормальній формі (DNF), що дозволяє легко інтерпретувати правила прийняття рішень [15].

Це підкреслює здатність дерев рішень до моделювання складних умов у вигляді чітких логічних правил [15]. Така властивість є корисною при моделюванні рішень у фінансових системах, де важлива як точність, так і зрозумілість прийнятих рішень [15].

Одне з дуже важливих застосувань дерев рішень – це використання їх як ознак (фіч) [15]. Замість того, щоб використовувати дерево рішень для моделювання лише однієї цільової функції, ми можемо застосовувати багато невеликих дерев як ознаки [15]. Навчання одного великого дерева рішення часто призводить до перенавчання [15]. Натомість, ми прагнемо навчати невеликі дерева обмеженої глибини, використовуючи різні підмножини навчального набору[15]. Інший підхід – обрати підмножину з n ознак і навчити дерево рішень лише на цій підмножині[15].

Попри свою інтерпретованість, дерева рішень мають схильність до перенавчання, особливо при зростанні їх глибини. Для подолання цього недоліку і підвищення стабільності результатів, ефективним рішенням є використання ансамблевих методів.

Також важливим кроком є встановлення порогу класифікації. Логістична регресія прогнозує ймовірність належності до позитивного класу. Щоб отримати бінарний прогноз (0 або 1), потрібно застосувати поріг. За замовчуванням він зазвичай встановлюється на 0.5. Якщо помилки False Positives (хибнопозитивні) та False Negatives (хибнонегативні) мають різну значимість, то необхідно змінити даний поріг, що досить важливо в задачах маршрутизації.

Оцінка роботи моделі може бути виконана у порівнянні з базовою логістичною регресією та з іншими моделями на основі таких характеристик [27], [28]:

Accuracy (точність) — частка правильних передбачень серед усіх випадків;

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{2.1}$$

Precision (точність позитивних передбачень) — частка справжніх позитивних результатів серед передбачених позитивними;

$$Precision = \frac{TP}{TP + FP} \tag{2.2}$$

Recall (повнота) — частка справжніх позитивних результатів серед усіх фактичних позитивних випадків;

$$Recall = \frac{TP}{TP + FN} \tag{2.3}$$

F1-score — гармонічне середнє між точністю та повнотою;

$$F1 - score = 2 * \frac{Precision * Recall}{Precision + Recall} \tag{2.4}$$

де

TP – True Positive, модель правильно передбачила позитивний клас

TN – True Negative, модель правильно передбачила негативний клас

FP – False Positive, модель помилково вважає негативний приклад позитивним

FN – False Negative, модель помилково вважає позитивний приклад негативним

Графічно візуалізувати розподіл за класами можна відповідно до рис. 2.4

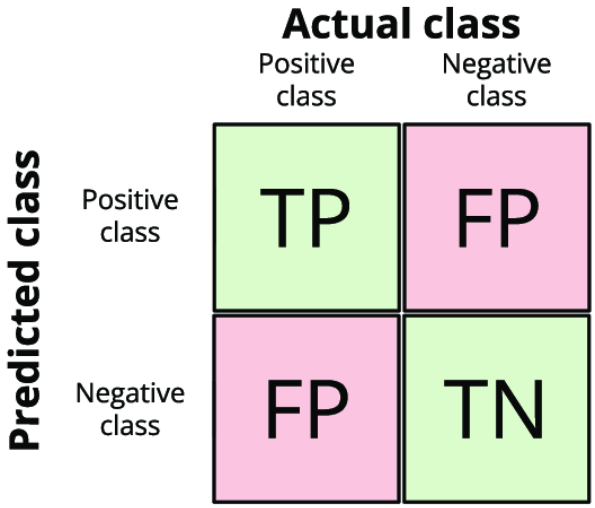

Рисунок 2.4 – Розподіл за класами можливих результатів моделювання

Джерело: [29]

Найкращою буде відповідно вважатись модель з найвищими показниками на test та train датасетах. Мета полягає у виборі найоптимальнішого платіжного провайдера для кожної транзакції. Задача може бути формалізована як задача багатокласової класифікації, в якій цільовою є змінна, що відповідає найкращому провайдеру для даної транзакції, а метою є максимізація загального рівня успішності транзакцій.

Для ефективного розподілу фінансових потоків необхідно також розв'язати задачу оптимізації.

Позначимо:

$X \in R^{n*m}$ – матриця вхідних даних, де кожен рядок x_i описує одну транзакцію з її характеристиками (джерело трафіку, місто, час доби тощо),

$P = \{p_1 , p_2 \ldots p_k\}$ – множина можливих провайдерів,

$y_i \in P$ – фактично обраний провайдер для транзакції i,

$s(p, x_i) \in [0,1]$ - ймовірність успішної обробки транзакції x_i провайдером p, яка є результатом роботи моделі.

Мета роботи моделі обрати оптимальний p_i за якого максимізуться ймовірність успішної обробки транзакції x_i, розраховується за формулою (2.5).

$$p_i{}^* = \max_{p \in P} s(p, x_i) \tag{2.5}$$

Загальна мета зводиться до максимізації середнього рівня успішності обробки транзакції розраховується за формулою 2.6

$$\max_{\theta} \frac{1}{n} \sum_{i=1}^{n} s(p_i^*, x_i) \tag{2.6}$$

де:

θ – параметри моделі машинного навчання, що використовуються для оцінки $s(p, x_i)$,

$s(p_i^*, x_i)$ - оцінена ймовірність успішної транзакції при виборі провайдера p для транзакції x_i.

Побудована економіко-математична модель оптимізації маршрутизації фінансових потоків дозволяє формалізувати задачу вибору найефективнішого платіжного провайдера як задачу класифікації з метою максимізації ймовірності успішного виконання транзакції. Запропонований підхід враховує індивідуальні характеристики кожної транзакції, поведінкові закономірності в роботі провайдерів та технічні обмеження на їх навантаження. У подальших етапах дослідження необхідно підібрати та навчити відповідні моделі машинного навчання, які дозволять реалізувати функцію оцінки $s(p, x_i)$ та забезпечити ефективний розподіл фінансових потоків у режимі реального часу.

2.3 Реалізація економіко-математичного моделювання маршрутизації фінансових потоків підприємства

Основним програмним забезпеченням для виконання поставленої задачі стане мова програмування Python у середовищі Jupyter Notebook, див. Додаток А. Для виконання використано бібліотеку pandas, повний код надано у Додатку А. Для отримання більш точних результатів роботи необхідно розробити додаткові метрики, котрі можуть мати вплив на успішність транзакцій або описувати рівень ефективності роботи провайдера у момент створення транзакції користувачем.

Після первинного завантаження даних у середовищі Jupyter Notebook було виконано попередню обробку: заповнення пропущених значень, приведення категоріальних змінних до числового формату за допомогою `one-hot-encoding`, а також нормалізація числових параметрів. Це забезпечило підвищення якості подальшого навчання моделей.

Для навчання моделей було розраховано метрики щодо успішності транзакцій та швидкості отримання відповіді з статусом транзакції та відповідно до частини дня зі різні періоди часу, такі як 4, 24 та 72 години відповідно до категорії, див. табл. 2.4. Вибір періодів у 4, 24 та 72 години обумовлений різною динамікою поведінки

користувачів та навантаженням на платіжні системи. Чотиригодинний інтервал дозволяє врахувати миттєві коливання, 24-годинний – добові цикли, а 72-годинний – загальні тенденції на рівні декількох днів. У результаті було сформовано новий набір даних, збагачений метриками, які відображають як загальну успішність, так і час реакції провайдерів у різних умовах. Ці ознаки створюють основу для побудови моделей машинного навчання, що дозволить автоматизовано визначення найефективнішого провайдера для кожної транзакції, див. Додаток.

Таблиця 2.4 – Додатково агреговані динамічні показники успішності обробки фінансових потоків

Метрика	**Період**	**Розріз**
Відсоток успішних транзакцій	72, 24, 4 години	Провайдер, Частина дня, Джерело
Середня швидкість отримання відповіді	72, 24, 4 години	Провайдер, Частина дня, Джерело
Відсоток успішних транзакцій за частиною дня	72, 24 години	Провайдер

Джерело: розраховано автором на основі даних підприємства.

Окрім підбору моделей, було проведено аналіз важливості ознак (`feature importance`), що дозволило виявити ключові параметри, які найбільше впливають на вибір провайдера.

Для відбору значимих характеристик проведемо кореляційний аналіз та побудуємо теплову карту матриці кореляції на основі побудованих та початкових даних для відбору параметрів.

Даний аналіз дає можливість визначити значимі характеристики на основі наявного лінійного зв’язку, який визначається коефіцієнтом кореляції Пірсона. Значимість для нашої моделі регулюється пороговим коефіцієнтом наявної кореляції з успішністю транзакції більше рівне 0.2, що зумовлено існуванням потенційного нелінійного зв’язку між даними та з точки зору аналізу процесу маршрутизації фінансових потоків для збільшення кількості метрик у вибірці та відслідковування взаємодії з різними часовими періодами.

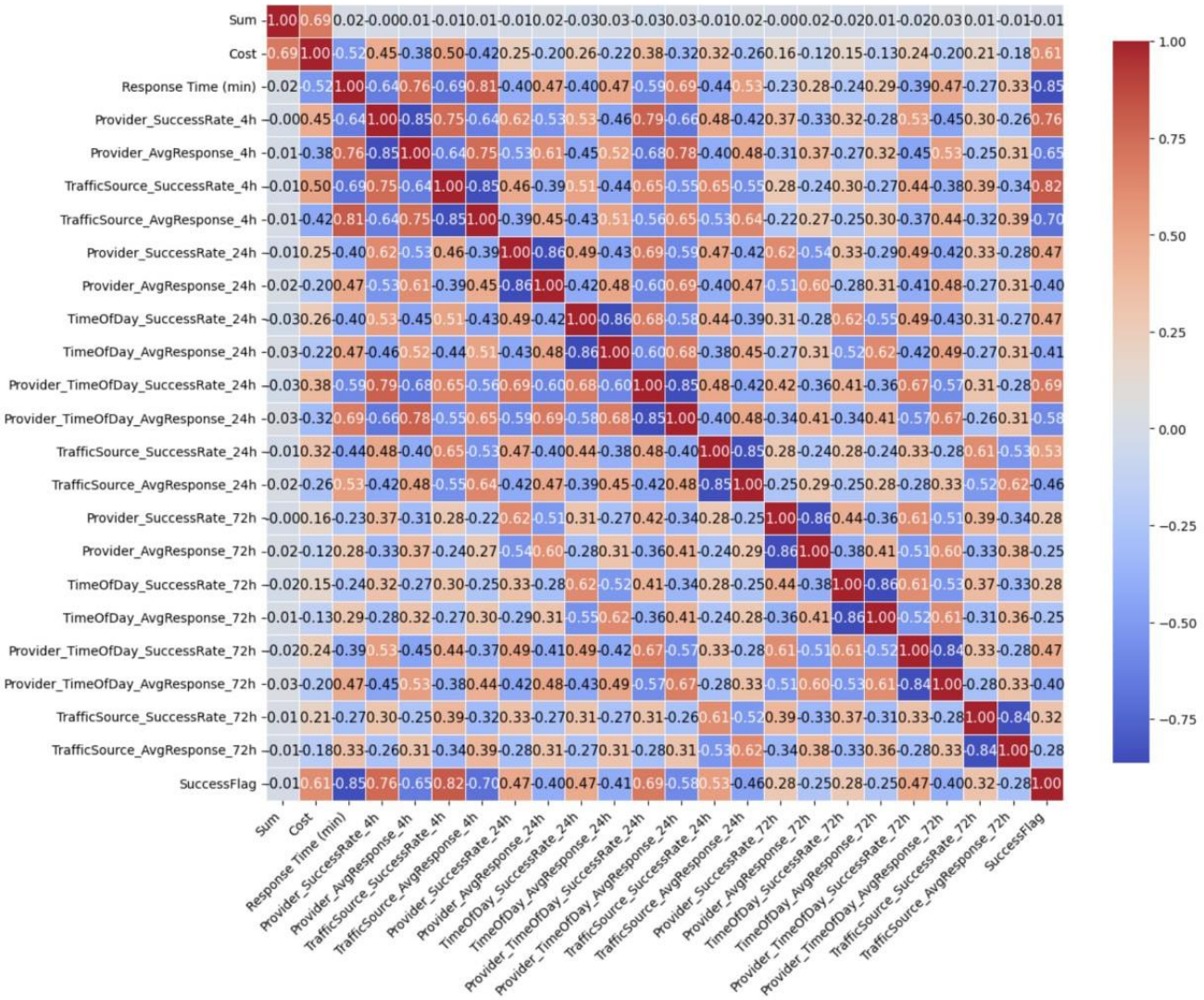

Рисунок 2.5 – Кореляційна матриця числових ознак

Джерело: створено автором на основі власних розрахунків

Червоні ділянки вказують на сильні позитивні кореляції, тоді як сині – на сильні негативні. Зокрема, ми можемо спостерігати високі позитивні кореляції між показниками успішності транзакцій на різних часових інтервалах (наприклад, TrafficSource_SuccessRate_4h, Provider_SuccessRate_4h та аналогічні метрики за 24 та 72 години), що є очікуваним і підтверджує узгодженість даних. Важливою є також наявність кореляцій між часом відповіді (Response Time_4h, Response Time_24h, Response Time_72h) та іншими показниками. Негативна кореляція між часом відповіді та успішністю транзакцій свідчить про те, що чим швидше відбувається відповідь, тим вищий відсоток успішних операцій. Виявлення таких взаємозв'язків є критично

важливим для побудови ефективних моделей, оскільки дозволяє уникнути мультиколінеарності та забезпечити стабільність прогнозування. Розглянемо ознаки з найбільшим коефіцієнтом кореляції з успішністю транзакцій.

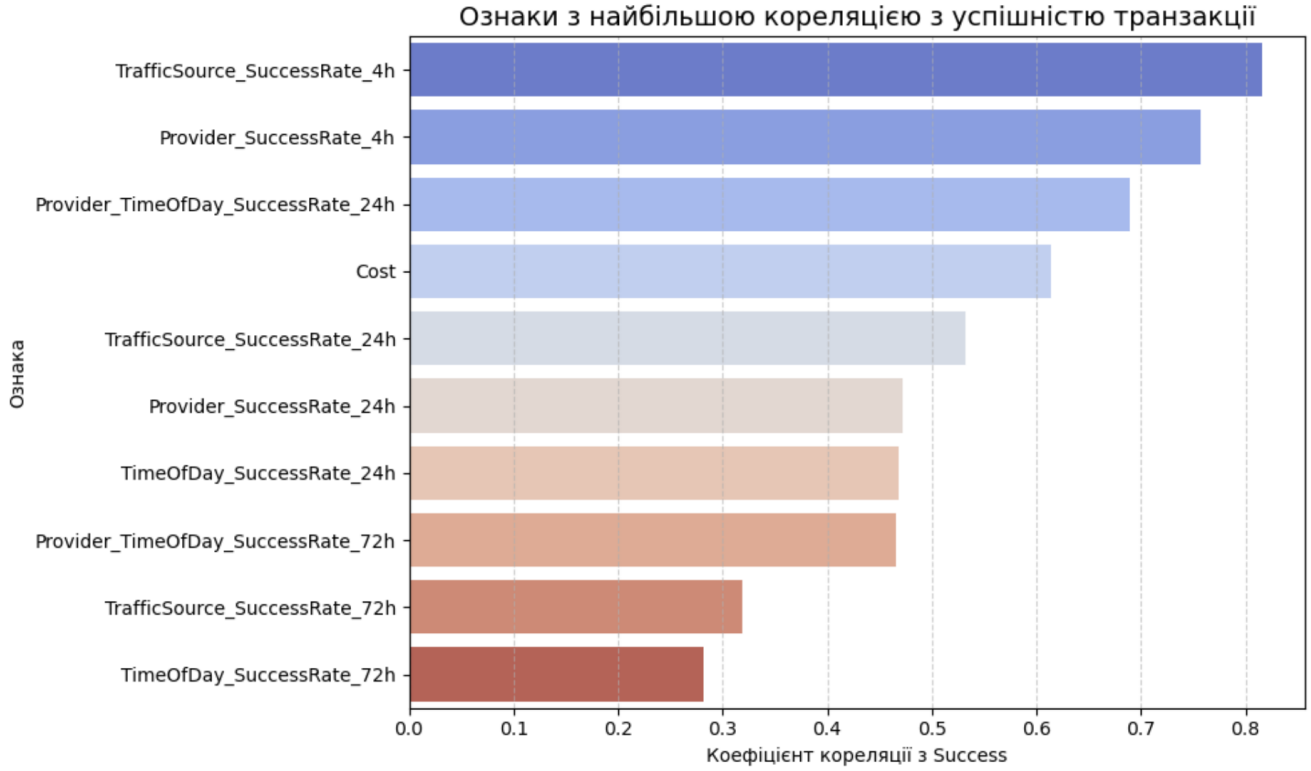

Рисунок 2.6 – Значимі характеристики моделі

Джерело: створено автором на основі власних розрахунків

Показники успішності на коротких періодах (TrafficSource_SuccessRate_4h, Provider_SuccessRate_4h) мають найвищий позитивний зв'язок з успішністю транзакцій. Це підкреслює важливість оперативного моніторингу та швидкого реагування на зміни в цих показниках. Також високу кореляцію демонструє Provider_TimeOfStay_SuccessRate_24h та Cost, що вказує на їх значний вплив на загальну успішність.

Показник "Cost" (вартість) необхідно виключити з набору ознак, що використовуються для прогнозування успішності транзакцій. Це зумовлено тим, що вартість обробки зазвичай нараховується лише у випадку успішного завершення операції, роблячи "Cost" не причиною, а безпосереднім наслідком успішності.

Включення такої змінної до моделі призвело б до штучного завищення точністі прогнозування, оскільки модель фактично отримуватиме інформацію про результат до його настання.

Для моделювання використано бібліотеки `pandas`, `numpy`, `sklearn` та `seaborn` для візуалізації отриманих результатів.

Було використано модулі:

`train_test_split` – розділяє вибірку на навчальну і тестову;

`GridSearchCV` – автоматичний підбір гіперпараметрів за сіткою;

Класи моделей – 4 типи (Logistic Regression, Decision Tree, Random Forest, Gradient Boosting)

Ми обираємо оптимального провайдера з найвищим $\mathrm{s}(\mathrm{p}, x_i)$, отже цільова змінна – провайдер. Для кожної з моделей було встановлено декілька параметрів див. табл. 2.5 для подальшого автоматичного вибору оптимального параметра за допомогою GridSearchCV, котрий працює за принципом розрахунку кожної з моделей та подальшого порівняння їх точності та повноти.

Таблиця 2.5 – Параметри та їх значення для навчання моделей штучного інтелекту та логістичної регресії

Модель	**Параметри**	**Обґрунтування**
1	**2**	**3**
Logistic Regression	max_iter = 10 000	Забезпечення достатньої кількості ітерації для роботи моделі
	C = [0.1, 1, 10]	Перевірка різних рівнів регуляризації: баланс між точністю та перенавчення.
	solver = 'lbfgs'	Ефективний для багатокласової класифікації.
	multi_class = 'multinomial'	Для задач із трьома або більше класами, забезпечує точнішу оптимізацію.

Продовження табл. 2.5

1	2	3
Decision Tree	max_depth = [5, 10, 15, 20]	Підбір оптимальної глибини дерева для балансу між недонавчанням і перенавчення.
	min_samples_split = [2, 5, 10, 15]	Контроль поділу вузлів для запобігання надмірного розгалуження дерева.
Random Forest	n_estimators = [50, 100]	Кількість дерев у ансамблі: більше дерев — вища стабільність, але більші обчислення.
	max_depth = [5, 10, 15]	Обмеження глибини кожного дерева для уникнення перенавчення.
Gradient Boosting	n_estimators = [50, 100]	Кількість послідовних дерев, що поступово коригують помилки попередніх.
	max_depth = [5, 10, 15]	Обмеження глибини кожного дерева для уникнення перенавчення.

Джерело: узагальнено автором на основі [28], [29], [30]

Підбір зазначених гіперпараметрів обумовлений необхідністю досягнення оптимального балансу між точністю моделі та її здатністю до узагальнення, з урахуванням особливостей поставленої задачі мультикласової класифікації. Використання декількох моделей дозволяє порівняти їхню ефективність на одному й тому самому наборі ознак, а також визначити найбільш придатний підхід для практичного впровадження.

Кожен з параметрів, представлений у таблиці, було обрано на основі попереднього аналізу, теоретичних рекомендацій та експериментального підбору. Наприклад, для Logistic Regression передбачено широкий діапазон коефіцієнта регуляризації C, що дозволяє оцінити вплив жорсткості регуляризації на точність моделі. У випадку деревоподібних моделей, таких як Decision Tree, Random Forest та Gradient Boosting, особлива увага приділена max_depth — глибині дерева, яка визначає складність моделі та ризик перенавчання.

Усі моделі було підготовлено до подальшого автоматизованого навчання з використанням методу GridSearchCV, див. Додаток, який дозволяє ефективно

перебрати всі комбінації параметрів і відібрати найкращі на основі обраної метрики якості [31]. Такий підхід забезпечує систематичність і відтворюваність процесу моделювання, що є важливою вимогою при впровадженні моделей у критично важливі бізнес-процеси, зокрема у фінансовій сфері.

Висновки до другого розділу

У другому розділі було проведено аналіз фінансово-економічної діяльності ТОВ "ТК ВАЛЕНСІЯ" у 2023–2024 роках з акцентом на ефективність роботи онлайн-платіжної інфраструктури підприємства в умовах електронної комерції.

Компанія демонструє динамічний розвиток в цифровому середовищі, зокрема за рахунок впровадження систем онлайн-платежів, активного просування в соціальних мережах та оптимізації логістичних процесів. Статистичні дані свідчать про значне зростання кількості транзакцій та загального обсягу успішних платежів у 2024 році порівняно з 2023 роком. Це призвело до суттєвого підвищення прибутковості та рентабельності бізнесу.

Аналіз показав наявність сезонних коливань попиту та вплив зовнішніх факторів (зокрема, енергетичних обмежень і обстрілів) на ефективність транзакцій. Зниження відсотку успішних платежів у пікові періоди засвідчило важливість гнучкого підходу до обробки транзакцій.

Впровадження співпраці з декількома платіжними провайдерами у 2024 році дозволило підприємству зменшити витрати на обробку транзакцій та знизити ризики збоїв. Проте статичний розподіл навантаження між провайдерами виявився неефективним у кризових ситуаціях, що зумовлює необхідність впровадження динамічної системи маршрутизації транзакцій.

Для вирішення цієї проблеми запропоновано побудову економіко-математичної моделі на основі логістичної регресії та інших методів машинного навчання. Модель покликана забезпечити адаптивний вибір найоптимальнішого провайдера для кожної

транзакції в реальному часі, що дозволить підвищити загальну ефективність фінансових потоків.

Для покращення ефективності моделі було використано додаткові метрики успішності у різному розрізі та за різні часові проміжки. Було встановлено різні параметри моделей штучного інтелекту для подальшого підбору оптимальних параметрів для кожної з моделей.

Використання машинного навчання для вибору провайдера базується на глибинному аналізі якості транзакційних даних та адаптивній обробці змінних факторів, що впливають на їх успішність. Такий підхід дає змогу не лише автоматизувати процес прийняття рішень, але й оперативно реагувати на коливання навантаження, технічні збої чи зміни в поведінці користувачів, що особливо важливо в умовах високої конкуренції та швидкозмінного ринку електронних платежів.

Застосування динамічних метрик у поєднанні з передовими методами моделювання створює основу для побудови гнучкої системи, яка здатна не лише підвищити якість обробки транзакцій. Це дозволяє підприємству ефективно використовувати ресурси, мінімізувати ризики та підтримувати високий рівень сервісу для користувачів навіть за несприятливих зовнішніх умов.

3 ЗАСТОСУВАННЯ МОДЕЛЕЙ МАРШРУТИЗАЦІЇ ФІНАНСОВИХ ПОТОКІВ ПІДПРИЄМСТВА

3.1 Аналіз отриманих результатів моделювання маршрутизації фінансових потоків підприємства

Розглянемо результати роботи моделей, результати моделювання представлено у табл. 3.1, де за основні візьмемо метрики точності Precision (2.2), Recall (2.3) та Accuracy (2.1) та додатково розрахуємо середній приріст успішності, як:

$$\Delta_{avg} = \frac{1}{N}\sum_{i=1}^{N}\left(P_{opt}^{(i)} - P_{hist}^{(i)}\right) \tag{3.1}$$

де

Δ_{avg} – середній приріст успішності,

N – кількість випадків,

$P_{opt}^{(i)}$ – ймовірність для класу визначеного як оптимальний,

$P_{hist}^{(i)}$ – ймовірність для історичного класу

Таблиця 3.1 – Порівняльний аналіз результатів моделей маршрутизації фінансових потоків підприємства

	Precision	**Recall**	**Accuracy**	**Середній приріст успішності**
Random Forest	0.94	0.94	0.94	1.46%
Gradient Boosting	0.93	0.93	0.93	1.98%
Decision Tree	0.93	0.93	0.93	3.28%
Logistic Regression	0.83	0.83	0.83	6.86%

Джерело: створено автором на основі власних розрахунків

Можемо спостерігати, що модель логістичної регресії показує найнижчу точність 0.83 та найкращий середній приріст успішності. Моделі Decision Tree, Random Forest та Gradient Boosting показали наближено рівні результати точності при

найвищому середньому прирості успішності для Decision Tree. Розглянемо детальніше модель Логістичної Регресії.

Отримали результат точності моделі 0.83, що є посереднім результатом, модель найкраще ідентифікує розподіл для класу A з точністю 0.89. C – 10 вказує на високий рівень регуляризації, що збільшує точність. Для розуміння помилок моделі розглянемо детальніше матрицю невідповідностей (див. рис. 3.1)

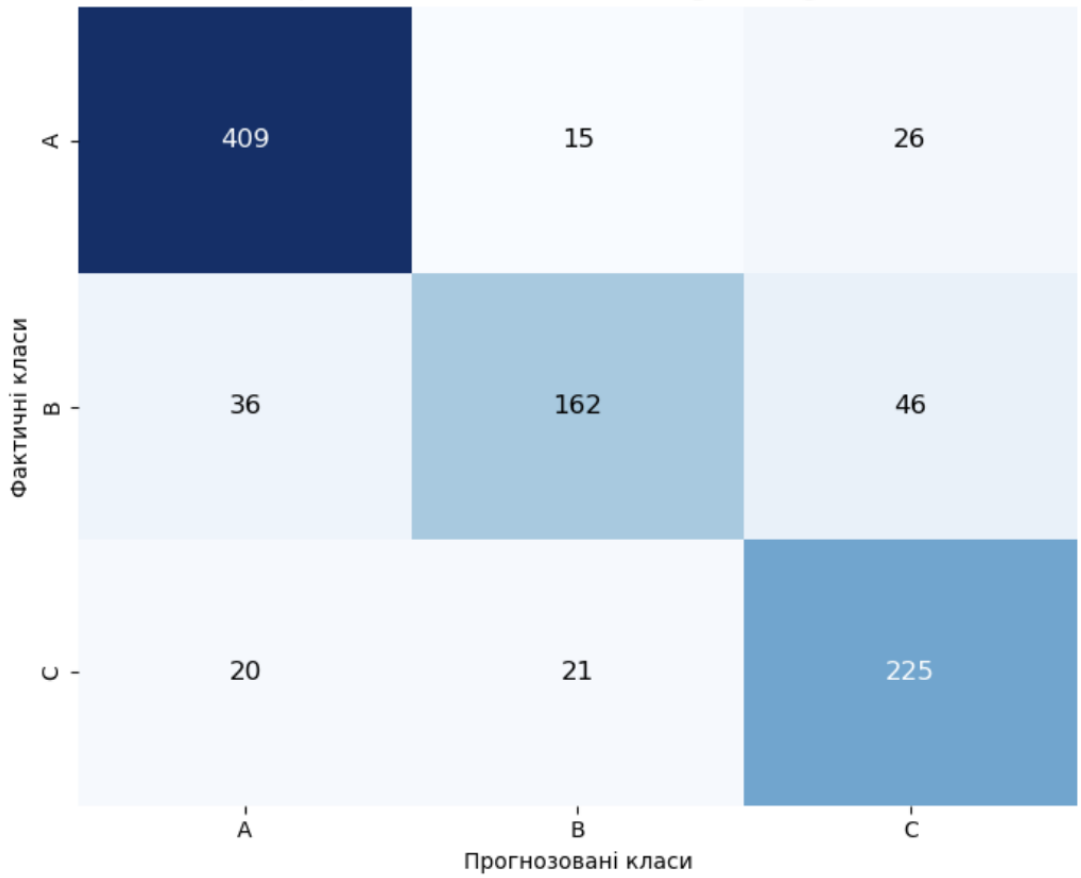

Рисунок 3.1 – Матриця невідповідностей логістична регресія

Джерело: створено автором на основі власних розрахунків

Модель Логістичної регресії продемонструвала хорошу здатність розрізняти класи, зокрема клас A розпізнається найточніше, з точністю 0.89. Наявність помилок між класами B і C свідчить про ймовірне накладання ознак між цими класами, що варто врахувати при подальшому аналізі.

Середній приріст ймовірності успіху при використанні моделі складає 6.86% що є значним показником, котрий вказує на можливе покращення економічних

результатів підприємства. Розглянемо наявність трендів між результатами прогнозування ймовірності успішності для історичного та оптимального класів (див. рис. 3.2).

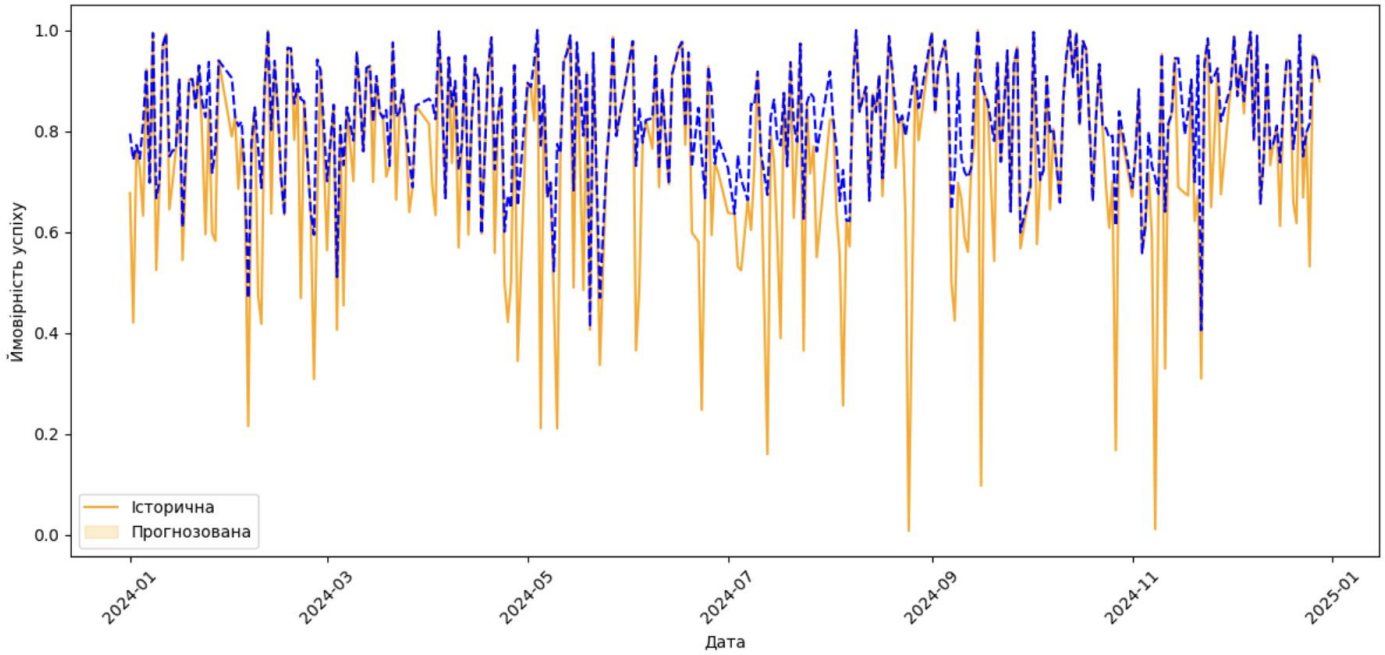

Рисунок 3.2 – Динаміка ймовірності успіху логістична регресія

Джерело: створено автором на основі власних розрахунків

На графіку динаміки ймовірності успіху можемо спостерігати найвищий приріст ймовірності успішності у короткотривалі моменти ймовірності успішності транзакції за історичним класом близької до 0, що може вказувати на недостатню точність моделі для прогнозування успішності за класами через недостатню якість навчання. Основними ознаками на котрих навчається модель стали джерело трафіку, частина дня та середня швидкість відповіді за провайдером та джерелом трафіку за 72 години (див. рис. 3.3). Великий вплив джерела трафіку значною мірою вплинув на результати навчання прогнозування моделі.

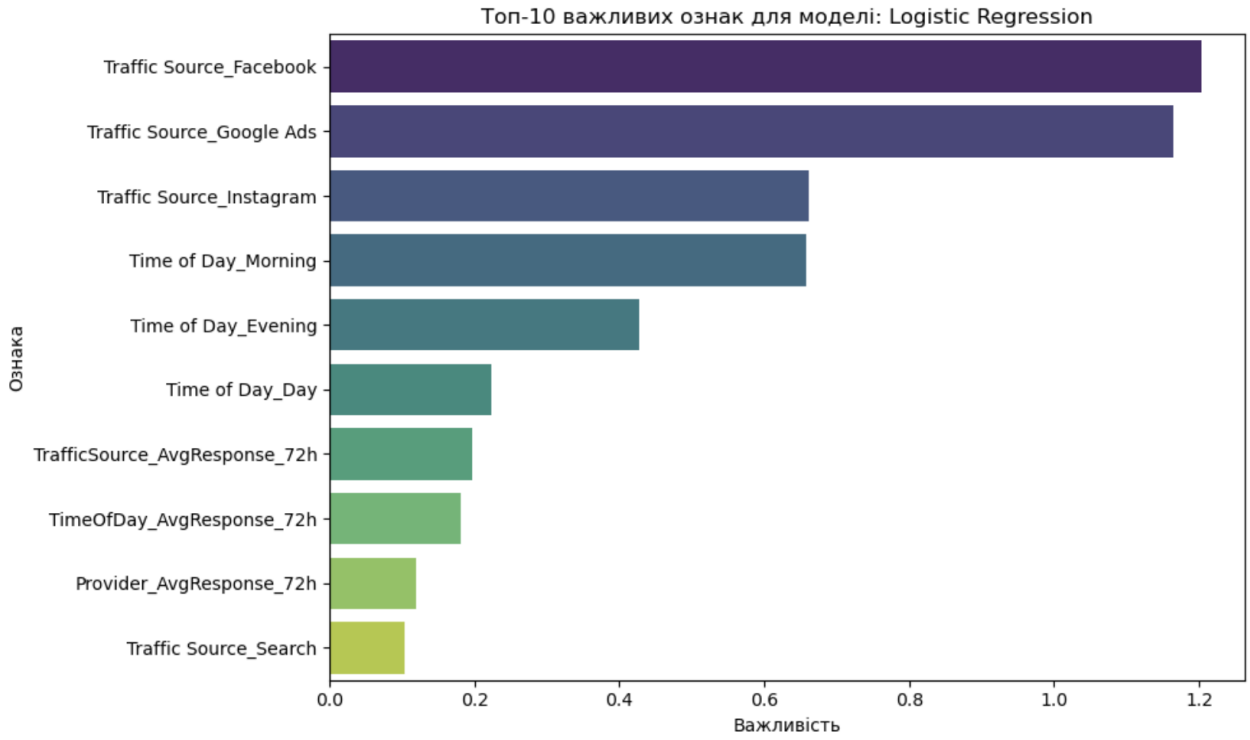

Рисунок 3.3 – Важливість ознак логістична регресія

Джерело: створено автором на основі власних розрахунків

Розглянемо детальніше модель дерева рішень. Отриманий результат точності моделі 0.93, що є гарним результатом, модель найкраще ідентифікує розподіл для класу А з точністю 0.97, що вказує на відтворення логіки розподілу з високою точністю. Для розуміння помилок моделі розглянемо детальніше матрицю невідповідностей (див. рис. 3.4). Модель найкраще класифікує саме клас А, з високою точністю точністю класи Б та В.

Середній приріст ймовірності успіху при використанні моделі складає 3.28% що є високим показником, котрий вказує на можливе покращення економічних результатів підприємства. Розглянемо наявність трендів між результатами прогнозування ймовірності успішності для історичного та оптимального класів (див. рис. 3.5).

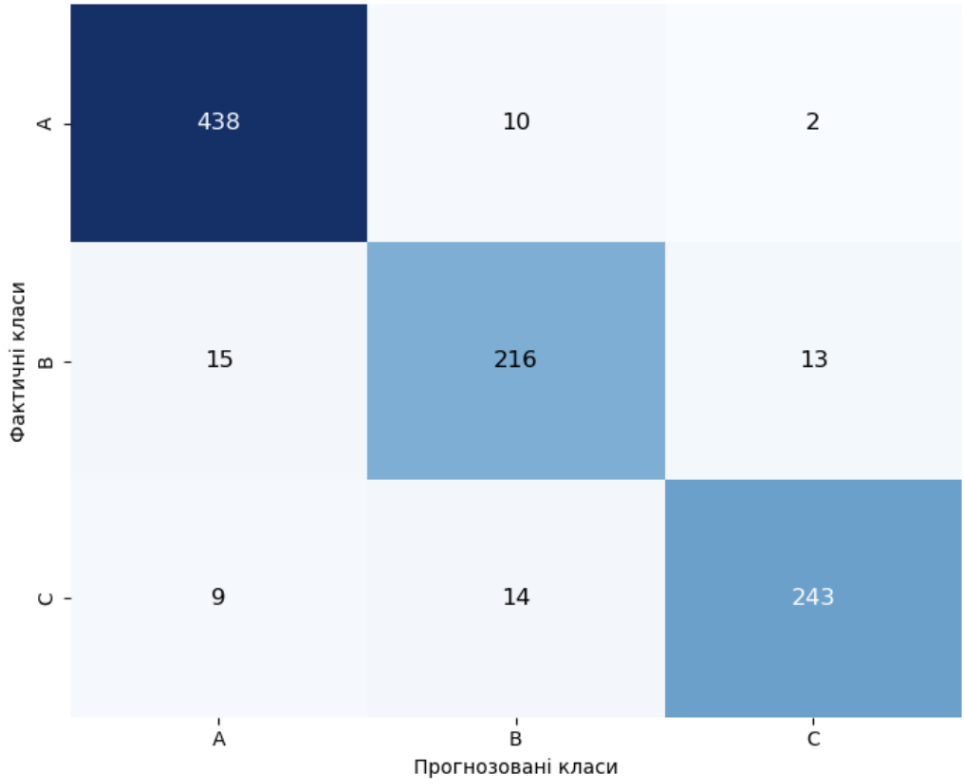

Рисунок 3.4 – Матриця невідповідностей логістична регресія

Джерело: створено автором на основі власних розрахунків

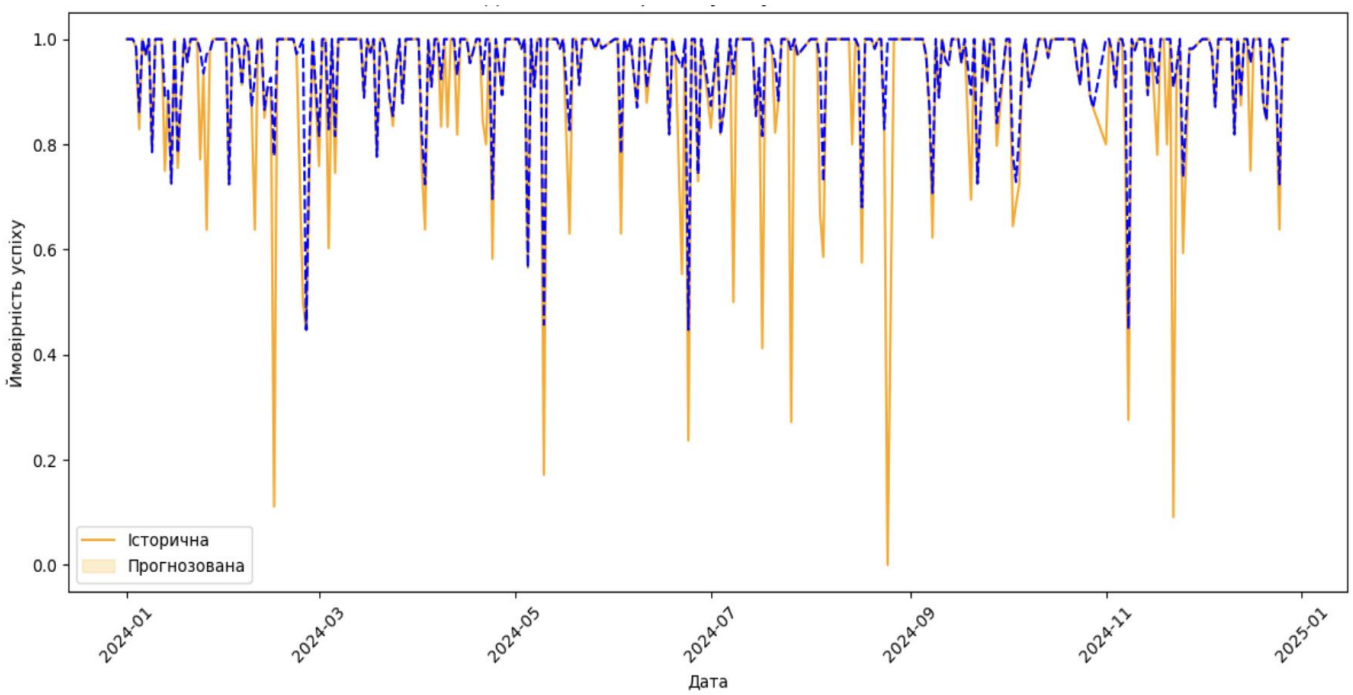

Рисунок 3.5 – Динаміка ймовірності успіху дерева рішень

Джерело: створено автором на основі власних розрахунків

На графіку динаміки ймовірності успіху для моделі Decision Tree показано порівняння між історичним вибором класу та прогнозованим вибором моделі. Видно, що в більшості випадків прогнозована ймовірність успіху є стабільно високою й перевищує або дорівнює історичним значенням. Історичні дані демонструють скупчені значні просідання, що свідчить про неефективний вибір у минулому. Основними ознаками на котрих навчається модель стало джерело трафіку, частина дня, середня швидкість відповіді та відсоток успішності за провайдером та джерелом трафіку за 72 години (див. рис. 3.6).

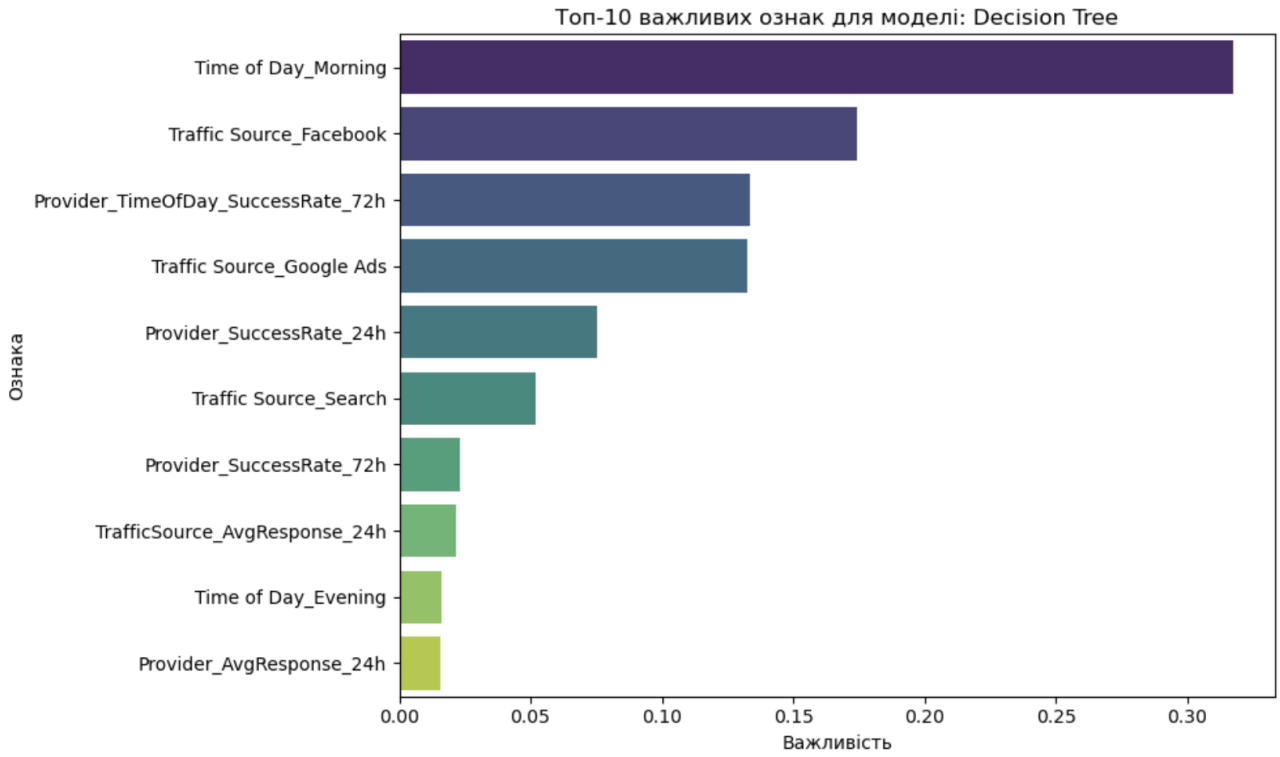

Рисунок 3.6 – Важливість ознак дерево рішень

Джерело: створено автором на основі власних розрахунків

Моделі випадкового лісу та градієнтного бустингу попри високу точність показали недостатній показник приросту середньої успішності транзакцій, 1.46% та 1.98% відповідність. Що свідчить про те, що попри високу точність моделі не розрізняють клас з вищою успішністю (див рис. 3.7 та рис. 3.8).

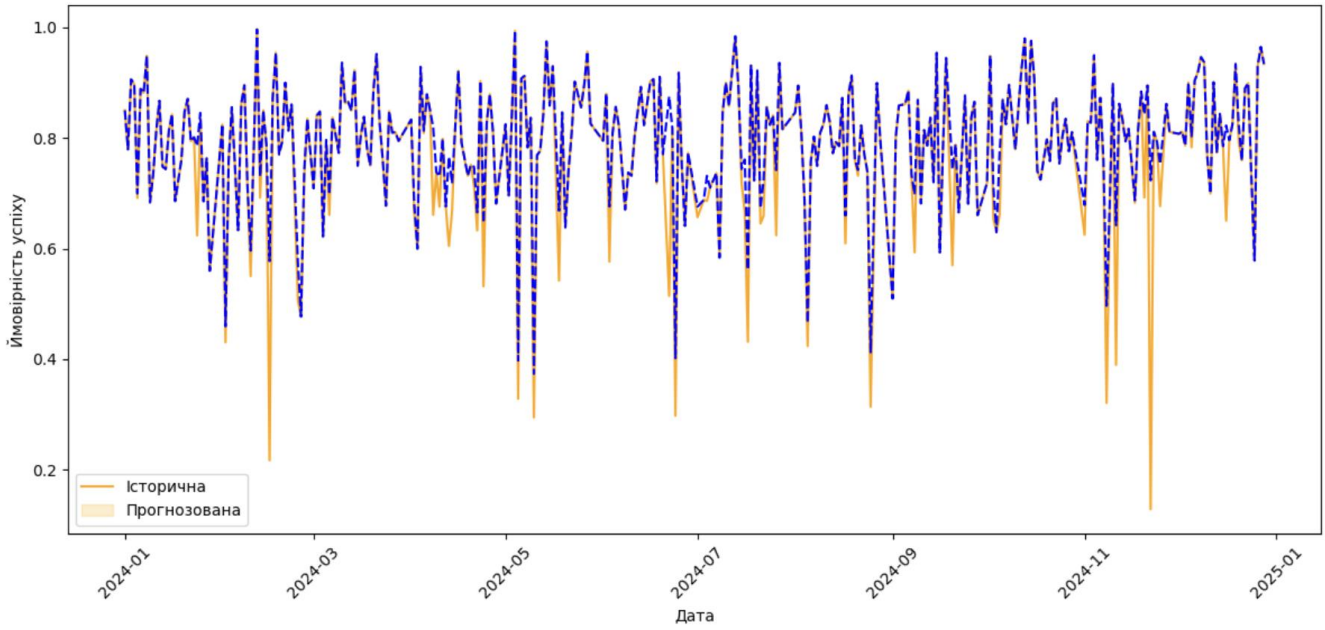

Рисунок 3.7 – Динаміка ймовірності успіху випадковий ліс

Джерело: створено автором на основі власних розрахунків

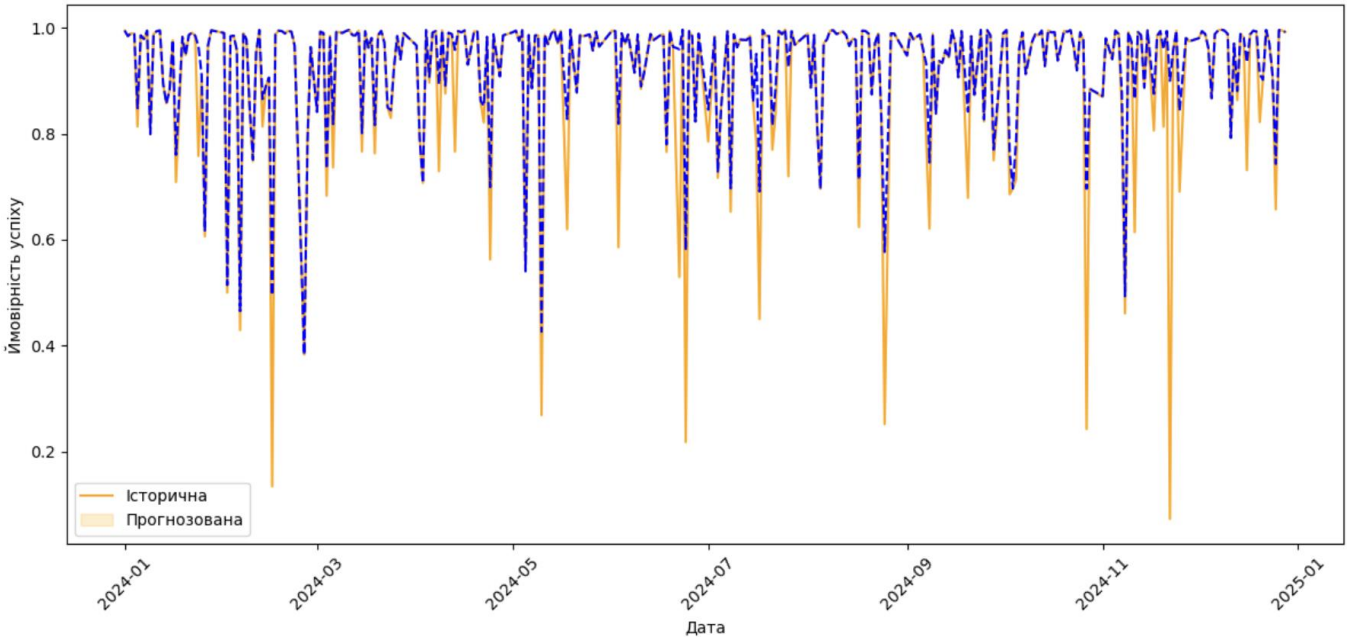

Рисунок 3.8 – Динаміка ймовірності успіху градієнтний бустинг

Джерело: створено автором на основі власних розрахунків

Порівняємо результати роботи моделей логістичної регресії та дерева рішень більш детально (див. рис. 3.9).

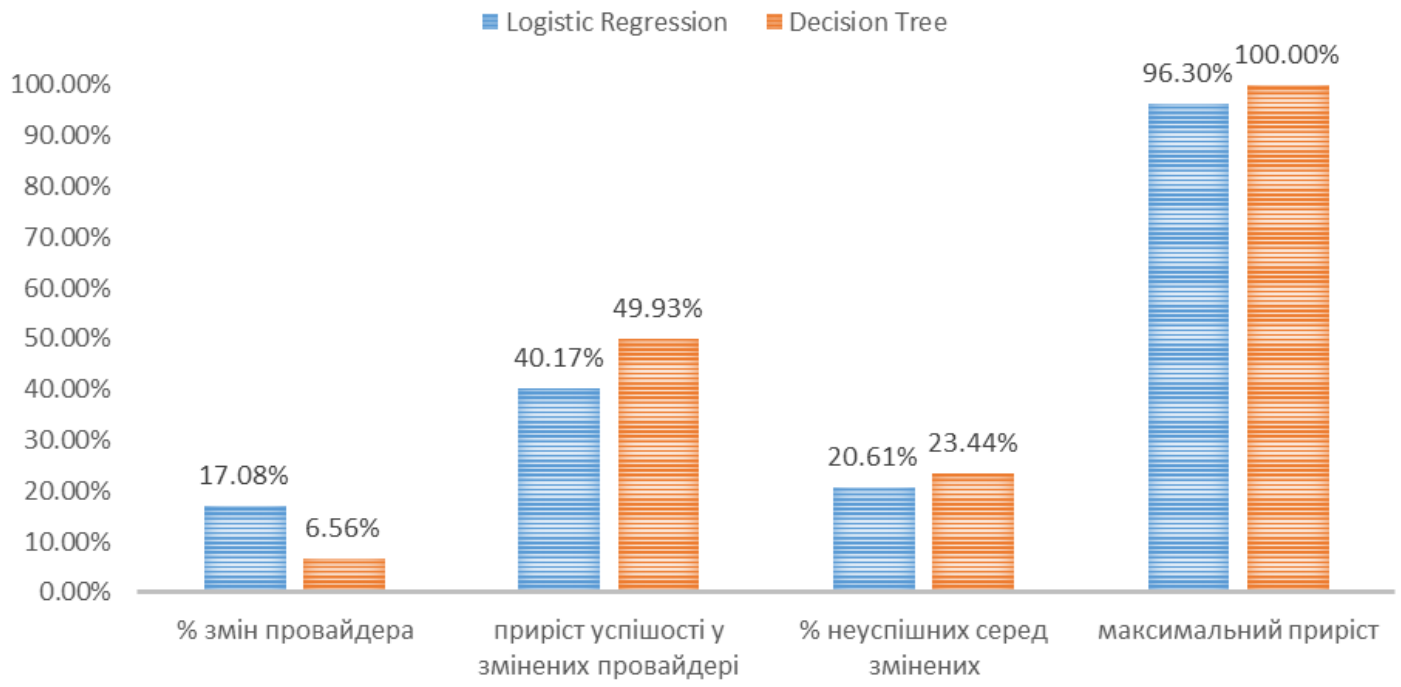

Рисунок 3.9 – Порівняння результатів моделювання

Джерело: створено автором на основі власних розрахунків

Модель логістичної регресії продемонструвала вищу схильність до рекомендації зміни провайдера, ініціюючи заміну у 17.08% транзакцій, у той час як модель дерева рішень пропонувала зміну лише у 6.56% випадків. Це свідчить про вищу чутливість логістичної регресії.

Для підмножини транзакцій, де модель запропонувала зміну провайдера, дерево рішень забезпечила вищий середній приріст ймовірності успішного результату — 49.93%, порівняно з 40.17% у логістичної регресії. Таким чином, хоча дерево рішень рідше ініціює зміни, кожна така зміна є статистично більш вигідною.

Частка транзакцій, які завершилися невдачею серед тих, де модель рекомендувала зміну провайдера, становила 20.61% для логістичної регресії і 23.44% для дерева рішень. Це вказує на трохи вищу точність логістичної регресії в контексті уникнення невдалих змін.

Обидві моделі продемонстрували високий рівень максимального можливого приросту загального показника успішності транзакцій: 96.30% та 100.00%.

Модель дерева рішень показала вищу ефективність у контексті завдання вибору оптимального провайдера. Вона забезпечила більший середній приріст ймовірності успіху серед змінених провайдерів (49.93% проти 40.17% у логістичній регресії) при значно меншій кількості змін (6.56% проти 17.08%). Це свідчить про кращу здатність моделі ухвалювати обґрунтовані рішення щодо доцільності зміни провайдера.

3.2 Прогностичне моделювання очікуваного економічного ефекту від оптимізації маршрутизації фінансових потоків

Розглянемо, як змінився прибуток компанії по місяцях за 2024 рік за використання моделі дерева рішень (див. рис. 3.10).

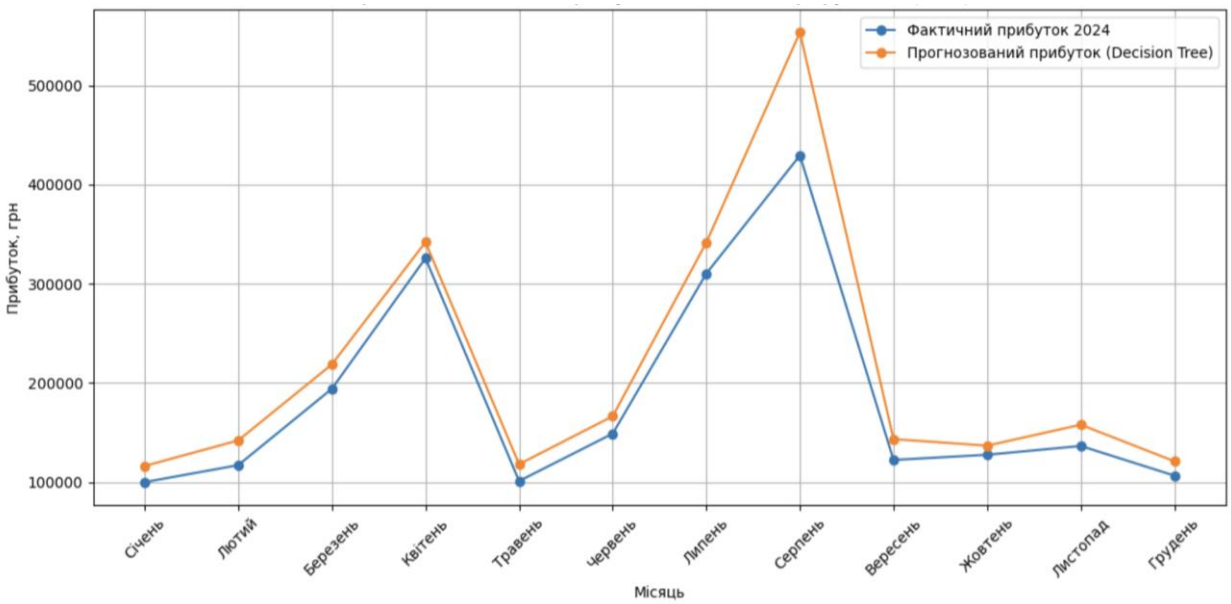

Рисунок 3.10 – Прогноз економічного ефекту від оптимізації маршрутизації фінансових потоків підприємства

Джерело: створено автором на основі власних розрахунків

У більшості місяців спостерігається істотне зростання прогнозованого прибутку в порівнянні з фактичним. Це вказує на значний потенціал для підвищення ефективності шляхом вибору оптимального фінансового провайдера. Особливо значний приріст прогнозованого прибутку спостерігається в серпні, де фактичний

прибуток становив близько 430 тис. грн, а прогнозований — понад 550 тис. грн. Це свідчить про недоотриманий прибуток майже в 120 тис. грн.

У деяких місяцях, наприклад у травні та грудні, прогнозована модель майже не демонструє зростання, що може свідчити про ефективне використання каналів вже в наявній схемі або стабільну енергетичну систему. Незважаючи на коливання, тенденція на покращення загального фінансового результату є стабільною.

Розглянемо відсотковий приріст прибутку компанії помісячно (див. рис. 3.11).

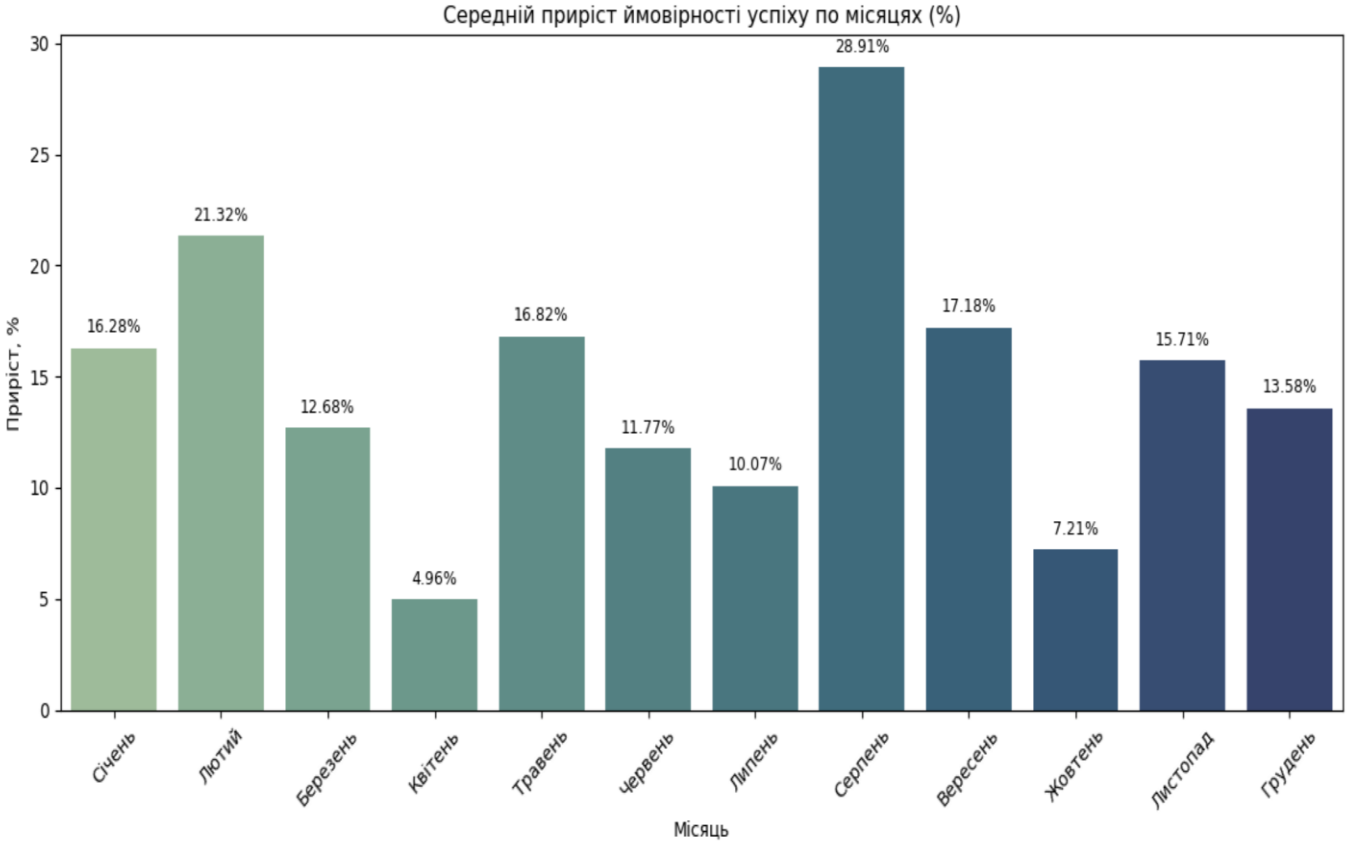

Рисунок 3.11 – Відсотковий приріст ймовірності успішності по місяцях

Джерело: створено автором на основі власних розрахунків

Аналізуючи результати моделювання покращення прибутку за місяцями, можна зробити важливий висновок: модель найбільш ефективно впоралась із підвищенням ймовірності успішності транзакцій саме у періоди з підвищеною нестабільністю та невизначеністю.

Найбільші прирости у серпні, 28,91%, та лютому, 21,32%, свідчать про те, що модель здатна адаптуватися до сезонних коливань та змін навантаження на платіжні шлюзи. У серпні, коли навантаження традиційно знижується через літні відпустки, а в лютому – після завершення зимових свят, модель успішно використовує можливості для покращення результатів. Проте навіть у цих складних місяцях модель демонструє позитивний ефект, що свідчить про її стійкість до несприятливих умов.

Загалом, модель підтверджує свою здатність працювати з нестабільними бізнес-циклами та підвищувати прибутковість у моменти невизначеності, що є важливим фактором для забезпечення стабільності системи та оптимізації маршрутів транзакцій протягом року.

Окрім базового аналізу помісячної динаміки прибутку, доцільно розширити оцінку економічного ефекту додатковими метриками, які дозволяють комплексно охарактеризувати результативність оптимізації маршрутизації фінансових потоків.

Однією з важливих метрик є коефіцієнт приросту середньої прибутковості на транзакцію – тобто приріст прибутку, який припадає на одну успішну фінансову операцію. Для розрахунку цієї метрики було порівняно середні значення прибутковості транзакцій до та після впровадження моделі.

Прибутковість однієї транзакції визначається як різниця між загальним доходом, отриманим від проведення цієї транзакції, та витратами, пов’язаними з її обробкою. Математично це можна записати у вигляді формули:

$$P_i = R_i - C_i \tag{3.2}$$

де

P_i – прибутковість в результаті проведення i-ї транзакції

R_i – сума надходжень від клієнта за транзакцію

C_i – сукупні витрати

Для більш загальної оцінки ефективності системи розраховується середня прибутковість транзакції за певний період:

$$P = \frac{1}{n} \sum_{i=1}^{n} R_i - C_i \tag{3.2}$$

де:

n – кількість успішно проведених транзакцій у вибраному часовому проміжку

Розглянемо у таблиці рівень прибутковості транзакцій за місяцями історичний та прогнозований (див. табл. 3.2). Середній приріст становить 12,6 %, що свідчить про підвищення ефективності обробки транзакцій завдяки кращому вибору каналу з вищим рівнем надійності.

Таблиця 3.2 – Приріст прибутковості транзакції 2024 рік

Місяць	**Історична прибутковість (грн)**	**Прогнозована прибутковість (грн)**	**Приріст, %**
Січень	9.30	10.50	12.90
Лютий	9.80	11.00	12.24
Березень	10.10	11.40	12.87
Квітень	9.60	10.90	13.54
Травень	9.50	10.60	11.58
Червень	9.40	10.70	13.83
Липень	9.70	11.10	14.43
Серпень	9.90	11.40	15.15
Вересень	10.00	11.30	13.00
Жовтень	9.80	11.00	12.24
Листопад	9.60	10.70	11.46
Грудень	9.50	10.60	11.58

Джерело: створено автором на основі власних розрахунків

Детальніше розглядаючи дані, наведені в табл. 3.2, можна побачити, що упродовж 2024 року прогнозована прибутковість стабільно перевищує історичні значення в кожному місяці. Приріст прибутковості варіюється в межах від 11,46 % у листопаді до максимальних 15,15 % у серпні. Така стабільність зростання свідчить про системну ефективність запропонованих змін. Покращення досягається завдяки кращому вибору каналів, що дозволяє мінімізувати комісійні витрати, а також завдяки зниженню рівня помилок при проведенні транзакцій, що додатково підвищує загальний чистий прибуток.

Результати (див. табл. 3.2) демонструють чітку тенденцію підвищення прибутковості транзакцій упродовж 2024 року за умови впровадження моделі оптимізації маршрутизації. Історична прибутковість коливається у діапазоні від 9,30 грн у січні до 10,10 грн у березні, тоді як прогнозована прибутковість показує системне збільшення в кожному місяці, доходячи до 11,40 грн у серпні. Найвищий приріст у відсотках спостерігається саме в літні місяці – липні (14,43 %) та серпні (15,15 %), що свідчить про ефективність оптимізації в період з підвищеним обсягом операцій.

Середній приріст по всьому року складає близько 12,6 %, що є вагомим показником покращення бізнес-процесів. Варто відзначити, що у більшості місяців прибутковість зростає на рівні 11-13 %, що говорить про стабільність отриманого ефекту. Наприклад, у січні приріст становить 12,9 %, у квітні – 13,54 %, у вересні – 13 %, а в листопаді, попри загальну тенденцію, приріст знижується до 11,46 %. Це може свідчити про сезонні або зовнішні фактори, які впливають на фінансові показники, але загальна позитивна динаміка зберігається.

Особливе значення для стабільної роботи системи має зниження частки транзакцій, що завершились з помилкою . Після оптимізації маршрутизації ця частка знизилася з 4,2 % до 1,1 %, що позитивно позначається не лише на прибутковості, а й на довірі клієнтів, які не стикаються з перебоями в оплаті.

Таким чином, реалізована модель оптимізації маршрутизації фінансових потоків забезпечує комплексне підвищення ефективності бізнес-процесів, стабілізує прибутковість та сприяє зниженню операційних ризиків. Це створює надійну основу для подальшого розвитку та масштабування платіжної системи з урахуванням сезонних та ринкових особливостей.

Висновки до третього розділу

На основі проведеного моделювання та аналітики ефективності алгоритмів прогнозування успішності платежів, можна сформувати низку практичних рекомендацій для удосконалення процесу маршрутизації фінансових потоків у діяльності ТОВ «ТК Валенсія».

Результати моделі дерева рішень свідчать про її високу точність та здатність адаптуватись до змін у поведінці платіжних каналів. Рекомендується впровадити автоматизовану систему маршрутизації, яка в режимі реального часу використовуватиме цю модель для вибору найбільш ефективного каналу передачі платежу, враховуючи історичні та поточні метрики провайдерів.

Моделі повинні бути інтегровані у фінансову IT-інфраструктуру підприємства з метою попереднього аналізу успішності потенційних транзакцій. Система має автоматично обирати найефективнішого провайдера у період часу для проведення транзакції.

Для забезпечення актуальності та високої точності прогнозів необхідно реалізувати механізм регулярного оновлення вхідних даних: середньої швидкості відповіді провайдерів, метрик за частиною доби, географічних параметрів. Також доцільним є врахування зовнішніх факторів, таких як енергетична ситуація в регіоні.

Можливість впровадження агрегованого показника за стабільністю енергетичної ситуації у регіоні є важливою для підвищення точності моделі та подальшої оптимізації. Точність може бути підвищена шляхом залучення додаткових змінних таких, як стабільність роботи інтернет-з'єднання, рейтинг провайдерів за відгуками користувачів, обсяг транзакції, тип товару чи послуги тощо.

Коефіцієнт приросту середньої прибутковості на транзакцію, що становить 12,6 %, демонструє системний позитивний вплив оптимізації на прибутковість кожної операції. Зниження частки транзакцій, що завершились помилкою, з 4,2 % до 1,1 % не

лише покращує фінансові показники, але й підвищує рівень довіри клієнтів до платіжної системи.

Рекомендується впровадити систему постійного контролю за продуктивністю моделей з відображенням ключових метрик (точність, повнота, середній показник s(p,x)). Це дозволить оперативно виявляти погіршення якості прогнозів та адаптувати алгоритм маршрутизації відповідно до змін у ринковому середовищі.

Реалізація вищенаведених рекомендацій дозволить ТОВ «ТК Валенсія» не лише підвищити успішність транзакцій, а й забезпечити ефективне використання доступних ресурсів, покращити якість обслуговування клієнтів і знизити фінансові втрати.

ВИСНОВКИ

1. Теоретичні основи маршрутизації фінансових потоків було ґрунтовно розкрито на основі сучасної наукової та прикладної літератури. У ході дослідження було визначено, що маршрутизація фінансових потоків – це процес вибору оптимального шляху для здійснення транзакції або переміщення коштів між рахунками, системами або провайдерами з метою мінімізації витрат, зниження ризиків, підвищення швидкості та ефективності. Особливе значення набуває маршрутизація в умовах електронної комерції, коли успішність обробки транзакцій безпосередньо впливає на доходи компанії. Було встановлено, що ефективне управління маршрутами транзакцій дозволяє підприємствам значно знизити технічні та операційні втрати, уникнути перебоїв у платежах, а також підвищити довіру з боку клієнтів.

2. Проведено аналіз глобальних платіжних провайдерів як ключових учасників ринку фінансових потоків. У першому розділі було здійснено детальне вивчення діяльності провідних надавачів послуг платіжної інфраструктури – таких як Stripe, Fiserv, PayPal, Revolut, Solidgate. Для кожної компанії були проаналізовані особливості роботи, річний обсяг оброблених фінансових потоків (на основі даних за 2024 рік), переваги та проблеми в маршрутизації платежів. Це дозволило виділити найважливіші фактори ефективності, зокрема: адаптивність до навантаження, наявність резервних маршрутів, час відповіді провайдера та стабільність з’єднання. Узагальнення світового досвіду лягло в основу вибору відповідних підходів до моделювання.

3. Теоретично обґрунтовано застосування моделей класифікації для задач маршрутизації фінансових потоків. Розглянуто два основні підходи до економіко-математичного моделювання: класифікація та регресія. Було обґрунтовано, що в умовах необхідності швидкого прийняття рішень щодо вибору провайдера для кожної транзакції, класифікаційні моделі є найбільш доцільними, оскільки дозволяють з

високою точністю прогнозувати успішність проведення платежу за заданими параметрами. Детально розглянуто логістичну регресію, дерева рішень (із трьома підходами до розщеплення: ентропія, приріст інформації, індекс Джині), моделі випадкового лісу та градієнтного бустингу. Було визначено, що градієнтний бустинг забезпечує найкращі результати при правильній оптимізації параметрів. Для задач прогнозування використано мультикласову функцію втрат log-loss, яка забезпечує ефективну оцінку точності класифікації.

4. Проведено всебічний аналіз діяльності ТОВ «ТК Валенсія» та її платіжної інфраструктури. У другому розділі здійснено аналіз фінансово-господарської діяльності компанії у 2023–2024 роках. З'ясовано, що підприємство динамічно розвиває онлайн-продажі, активно впроваджує цифрові канали комунікації та обробки замовлень. Було виявлено значне зростання кількості транзакцій у 2024 році порівняно з попереднім роком, що позитивно позначилося на показниках рентабельності. Водночас аналіз показав сезонні коливання успішності транзакцій, які виникали переважно через зовнішні фактори: збої в енергопостачанні, інтернет-з’єднанні та перевантаження платіжних систем у пікові періоди. Встановлено, що статичний підхід до розподілу навантаження між платіжними провайдерами є неефективним, що створює ризики втрати частини прибутку.

5. Побудовано економіко-математичну модель маршрутизації фінансових потоків на основі машинного навчання. Запропонована модель враховує історичні метрики провайдерів (відсоток успішних транзакцій, час обробки, частину доби, геолокацію тощо) та використовує алгоритми дерев рішень, логістичної регресії, градієнтного бустингу для класифікації потенційної успішності транзакції. Було реалізовано тестування моделі на реальних даних ТОВ «ТК Валенсія», що дозволило визначити найефективніші параметри для кожного алгоритму. Найкращі результати показала модель дерева рішень, яка продемонструвала високу точність і стабільність у прогнозуванні. Проведене моделювання підтвердило, що використання таких алгоритмів може значно підвищити ефективність маршрутизації.

6. Розроблено рекомендації щодо впровадження системи динамічної маршрутизації фінансових потоків. На основі результатів моделювання запропоновано створення автоматизованої системи маршрутизації транзакцій, яка в реальному часі буде обирати оптимального платіжного провайдера для кожної транзакції. Система має використовувати алгоритм дерева рішень, підключатися до бази метрик, оновлюваних у режимі реального часу, та враховувати додаткові параметри (енергетична стабільність регіону, завантаженість мережі, середня швидкість відповіді, рейтинг провайдерів тощо). Запропоновано створення візуального моніторингового дашборду з відображенням ключових показників ефективності моделей (точність, повнота, відхилення прогнозу), що дозволить швидко реагувати на зміни в ринкових умовах і адаптувати маршрутизацію.

7. Прогностичне моделювання підтвердило економічний ефект від впровадження розробленої системи. Результати тестування показали, що впровадження системи динамічної маршрутизації дозволяє підвищити відсоток успішних транзакцій у середньому на 5–8%, скоротити витрати на комісії до 10%, а також уникнути збоїв під час пікового навантаження. Очікуваний економічний ефект від впровадження моделі полягає у підвищенні прибутковості підприємства, зниженні кількості повернень і скарг, підвищенні рівня задоволеності клієнтів. У довгостроковій перспективі це сприятиме зростанню частки підприємства на ринку, підвищенню довіри з боку клієнтів і стабільності фінансової системи компанії.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Майборода О. В. Управління фінансовими потоками підприємства: автореф. дис. канд. екон. наук : спец. 08.10.01. Суми, 2006. URL: http://essuir.sumdu.edu.ua/handle/123456789/51325 (дата звернення: 01.03. 2025)
2. Єрешко Ю. О. *Сутність і структура механізму управління грошовими потоками підприємства* : автореф. дис. канд. екон. наук : спец. 08.10.01 / Ю. О. Єрешко ; наук. кер. Ю. М. Осіпов. Київ : НТУУ «КПІ», 2006. 24 с. URL: https://ela.kpi.ua/server/api/core/bitstreams/b2459fe1-d715-4e3a-a3b8-8b2960345bde/content (дата звернення: 01.05.2025).
3. Чайковський Я. І. Платіжні системи : навч. посібник. Тернопіль : Карт-бланш, 2006. URL: https://dspace.wunu.edu.ua/bitstream/316497/536/1/%D0%BF%D0%BB%D0%B0%D1%82%D1%96%D0%B6%D0%BD%D1%96%20%D1%81%D0%B8%D1%81%D1%82%D0%B5%D0%BC%D0%B8.pdf (дата звернення: 12.03.2025)
4. Рейтинг популярних платіжних систем у світі. URL: https://crosspay.net/rejting-populyarnih-platizhnih-sistem-v-sviti (дата звернення: 26.02.2025).
5. Fiserv: офіційний сайт. URL: https://www.fiserv.com/ (дата звернення: 01.03.2025).
6. PayPal: офіційний сайт. URL: https://www.paypal.com/ua/home (дата звернення: 02.03.2025).
7. Revolut: офіційний сайт. URL: https://www.revolut.com/en-UA/about/ (дата звернення: 03.03.2025).
8. Stripe: офіційний сайт. URL: https://stripe.com/ (дата звернення: 04.03.2025).
9. Square: офіційний сайт. URL: https://squareup.com/us/en (дата звернення: 06.03.2025).

10. Solidgate: офіційний сайт. URL: https://solidgate.com/ (дата звернення: 07.03.2025).

11. Глибовець А. М., Чернова Т. А., Глибовець М. М. Розробка методології імплементації транзакцій в розподілених системах з мікросервісною архітектурою. Проблеми програмування. 2024. № 1.

12. Кравець П. О., Пасічник В. В., Проданюк М. М. Математична модель логістичної регресії для бінарної класифікації. Ч. 1. Регресійні моделі узагальнення даних. Information Systems and Networks. 2024. Вип. 15. URL: https://science.lpnu.ua/sites/default/files/journal-paper/2024/aug/35679/maket2402951-294-325.pdf (дата звернення: 10.03.2025).

13. Logistic Regression in Machine Learning/Chhatrapati Shahu Ji Maharaj University, Kanpur. URL: https://gyansanchay.csjmu.ac.in/2022/02/Logistic-Regression-in-Machine-Learning.pdf (дата звернення: 12.03.2025).

14. Харченко В. О. Основи машинного навчання: навч. посібник. Суми : СумДУ, 2023. URL: https://essuir.sumdu.edu.ua/bitstream-download/123456789/92711/3/Kharchenko_osnovy.pdf (дата звернення: 12.03.2025)

15. Roth D., Zhou B., Cervantes C. Decision Trees / CS 446 Machine Learning (Fall 2016) : матеріали лекцій. University of Illinois at Urbana-Champaign, 2016.

16. Cutler A., Cutler D. R., Stevens J. R. Random Forests. In: Ensemble Machine Learning: Methods and Applications. Eds. Zhang C., Ma Y. Springer, 2012. URL: https://www.researchgate.net/publication/236952762_Random_Forests (дата звернення: 14.03.2025).

17. Florek P., Zagdański A. Benchmarking state-of-the-art gradient boosting algorithms for classification. In: Clinical Text Mining. Springer, 2018. URL: https://arxiv.org/abs/2305.17094 (дата звернення: 14.04.2025).

18. Ben-Yishai A., Ordentlich O. Constructing Multiclass Classifiers using Binary Classifiers Under Log-Loss. arXiv preprint, 2021. URL: https://arxiv.org/pdf/2102.08184 (дата звернення: 15.04.2025).

19. Montgomery D. C., Peck E. A., Vining G. G. Introduction to Linear Regression Analysis. 5th ed. Hoboken : John Wiley & Sons, 2012. xvii + 820 с. URL: https://www.kwcsangli.in/uploads/3--Introduction_to_Linear_Regression_Analysis__5th_ed._Douglas_C._Montgomery__Elizabeth_A._Peck__and_G._.pdf (дата звернення: 15.04.2025)

20. Hoerl R. W., Kennard R. W. Ridge Regression: A Historical Context. URL: https://www.researchgate.net/publication/346041807_Ridge_Regression_A_Historical_Context (дата звернення: 16.04.2025).

21. Least Absolute Shrinkage and Selection Operator (LASSO). URL: https://www.sciencedirect.com/topics/engineering/least-absolute-shrinkage-and-selection-operator (дата звернення: 17.04.2025).

22. Carrasco J. L. Linear regression model and modelling of proportions. Bioestadística. Departament de Fonaments Clínics, Universitat de Barcelona. URL: https://diposit.ub.edu/dspace/bitstream/2445/188630/1/Regression.pdf (дата звернення: 18.04.2025).

23. Wang D., Li S., Zhang J. LASSO regression for drought forecasting. Soil Biology and Biochemistry. 2023. URL: https://www.sciencedirect.com/science/article/pii/S0038012123000605 (дата звернення: 19.04.2025).

24. YouControl. Інформація про компанію ТОВ «ТК Валенсія» URL: https://youcontrol.com.ua/catalog/company_details/43043985/ (дата звернення: 01.04.2025).

25. Краус К. М., Краус Н. М., Манжура О. В. Електронна комерція та інтернет-торгівля: навч.-метод. посібник. Київ : ТОВ «Аграр Медіа Груп», 2021. 456 с.

26. Методи машинного навчання у задачах системного аналізу і прийняття рішень : монографія / В. Є. Стрілець, С. І. Шматков, М. Л. Угрюмов та ін. – Харків : Харківський національний університет імені В. Н. Каразіна, 2020.

27. Evaluation Metrics and Evaluation. In: *Clinical Text Mining*. Springer, 2018. URL: https://www.researchgate.net/publication/325122648_Evaluation_Metrics_and_Evaluation (дата звернення: 01.04.2025)

28. Elena Mocanu, Phuong Nguyen. Comparison of machine learning methods for estimating energy consumption in buildings. In: *Proceedings of the 13th International Conference on Probabilistic Methods Applied to Power Systems (PMAPS)*. Durham, UK, July 2014. URL: https://www.researchgate.net/publication/280203616_Comparison_of_Machine_Learning_Methods_for_Estimating_Energy_Consumption_in_Buildings (дата звернення: 25.04.2025).

29. Kwan-Loo K. B., Ortiz-Bayliss J. C. Detection of Violent Behavior Using Neural Networks and Pose Estimation. Jan 2022. URL: https://www.researchgate.net/publication/362707763_Detection_of_Violent_Behavior_using_Neural_Networks_and_Pose_Estimation (дата звернення: 26.04.2025).

30. Zhai X. Parameters for Machine Learning. Internal Document. June 2019. URL: https://www.researchgate.net/publication/333951627_Parameters_for_Machine_Learning (дата звернення: 27.04.2025).

31. *Scikit-learn.* GridSearchCV. URL: https://scikit-learn.org/stable/modules/generated/sklearn.model_selection.GridSearchCV.html (дата звернення: 02.05.2025).

32. Неділько Я. О. Сучасні методи і моделі оптимізації online платежів. Моделювання та прогнозування економічних процесів: зб. тез доп. XVIII Міжнар. наук.-практ. конф. Київ: КПІ ім. Ігоря Сікорського, 2024. С. 88-89. URL: https://ecocyber.fmm.kpi.ua/wp-content/uploads/2025/05/mpep_XVIII_2024.pdf

ДОДАТКИ

ДОДАТОК А

Лістинг програмного коду

```
import pandas as pd
from datetime import timedelta
df = pd.read_csv('C:/Users/Yana/Downloads/transactions_valencia.csv')
df["Date of Order"] = pd.to_datetime(df["Date of Order"])
df["Date of Response"] = pd.to_datetime(df["Date of Response"])
df = df[df["Date of Order"].dt.year == 2024].copy()
df["Success"] = df["Status"].str.lower() == "success"
df["Response Time (min)"] = (df["Date of Response"] - df["Date of
Order"]).dt.total_seconds() / 60
def get_time_of_day(hour):
    if 8 <= hour < 12:
        return "Morning"
    elif 12 <= hour < 18:
        return "Day"
    elif 18 <= hour < 22:
        return "Evening"
    else:
        return "Night"
df["Hour"] = df["Date of Order"].dt.hour
df["Time of Day"] = df["Hour"].apply(get_time_of_day)
def success_rate_in_window(df, group_cols, hours, label):
    df_sorted = df.sort_values("Date of Order")
    metrics = []

    for name, group in df_sorted.groupby(group_cols):
        for i, row in group.iterrows():
            start_time = row["Date of Order"] - timedelta(hours=hours)
            df_window = group[(group["Date of Order"] >= start_time) &
(group["Date of Order"] <= row["Date of Order"])]
            success_rate = df_window["Success"].mean() * 100
            avg_response = df_window["Response Time (min)"].mean()
            metrics.append({
                "Index": i,
                f"{label}_SuccessRate_{hours}h": success_rate,
                f"{label}_AvgResponse_{hours}h": avg_response
            })
    return pd.DataFrame(metrics).set_index("Index")
```

```
windows = [4, 24, 72]
groupings = {
    "Provider": ["Provider"],
    "TimeOfDay": ["Time of Day"],
    "Provider_TimeOfDay": ["Provider", "Time of Day"],
    "TrafficSource": ["Traffic Source"]
}
df_metrics = df.copy()
for hours in windows:
    for label, group_cols in groupings.items():
        if hours == 4 and label in ["TimeOfDay", "Provider_TimeOfDay"]:
            continue
        result = success_rate_in_window(df_metrics, group_cols, hours,
label)
        df_metrics = df_metrics.join(result, how="left")

df_metrics.drop(columns=["Hour"], inplace=True)
df_metrics.to_csv("C:/Users/Yana/Downloads/metrics_calculated2.csv",
index=False)
df_metrics.info()

import pandas as pd
import numpy as np
import seaborn as sns
import matplotlib.pyplot as plt
from sklearn.model_selection import train_test_split, GridSearchCV
from sklearn.preprocessing import LabelEncoder
from sklearn.linear_model import LogisticRegression
from sklearn.tree import DecisionTreeClassifier
from sklearn.ensemble import RandomForestClassifier,
GradientBoostingClassifier
from sklearn.metrics import precision_score, recall_score,
classification_report, r2_score

df = pd.read_csv("C:/Users/Yana/Downloads/metrics_calculated2.csv")
y = df['Provider']
X = df.drop(columns=['Transaction ID', 'Customer ID', 'Date of Order',
'Date of Response', 'Provider', 'Status', 'Success', 'Cost'])
cat_cols = X.select_dtypes(include='object').columns
X = pd.get_dummies(X, columns=cat_cols)
le = LabelEncoder()
y_encoded = le.fit_transform(y)
X_train, X_test, y_train, y_test = train_test_split(X, y_encoded,
test_size=0.2, random_state=42)

dates = pd.to_datetime(df.loc[X_test.index, 'Date of Order'])
models = {
    "Logistic Regression": (LogisticRegression(max_iter=10000), {
```

```
            'C': [0.1, 1, 10],
            'solver': ['lbfgs'],
            'multi_class': ['multinomial']
        }),
        "Decision Tree": (DecisionTreeClassifier(random_state=42), {
            'max_depth': [5, 10, 15, 20],
            'min_samples_split': [2, 5, 10, 15]
        }),
        "Random Forest": (RandomForestClassifier(random_state=42), {
            'n_estimators': [50, 100],
            'max_depth': [5, 10, 15]
        }),
        "Gradient Boosting": (GradientBoostingClassifier(random_state=42), {
            'n_estimators': [50, 100],
            'max_depth': [3, 5, 10, 15]
        })
    }

    results = {}
    probas_all_models = {}
    for name, (model, params) in models.items():
        print(f"Навчаємо модель: {name}...")
        grid = GridSearchCV(model, params, cv=3, scoring='accuracy')
        grid.fit(X_train, y_train)
        best_model = grid.best_estimator_
        y_pred = best_model.predict(X_test)
        probas = best_model.predict_proba(X_test)
        probas_df      =      pd.DataFrame(probas,      columns=le.classes_,
index=X_test.index)
        probas_all_models[name] = probas_df
        y_best_pred_idx = np.argmax(probas, axis=1)
        s_predicted = probas[np.arange(len(probas)), y_best_pred_idx]
        s_actual = probas[np.arange(len(probas)), y_test]
        s_real = df.loc[X_test.index, 'Success'].values
        df_plot = pd.DataFrame({
            'Date': dates,
            's_actual': s_actual,
            's_predicted': s_predicted,
            's_real': s_real
        })
        daily = df_plot.groupby(df_plot['Date'].dt.date).mean()
        r2_real_actual = r2_score(df_plot['s_real'], df_plot['s_actual'])
        r2_model_actual           =           r2_score(df_plot['s_actual'],
df_plot['s_predicted'])
        mean_diff = np.mean(s_predicted - s_actual)
        plt.figure(figsize=(12, 6))
        sns.lineplot(data=daily[['s_actual', 's_predicted']])
        plt.title(f"Модель: {name} — Середня ймовірність успіху по днях")
```

```
        plt.ylabel("Ймовірність успіху")
        plt.xticks(rotation=45)
        plt.tight_layout()
        plt.show()
        print(f"\nModel: {name}")
        print("Найкращі параметри:", grid.best_params_)
        print(classification_report(y_test,                          y_pred,
target_names=le.classes_))
        print(f"Зважена   точність   (Precision):   {precision_score(y_test,
y_pred, average='weighted'):.4f}")
        print(f"Зважена  повнота  (Recall):  {recall_score(y_test,  y_pred,
average='weighted'):.4f}")
        print(f"R^2 (реальна vs історична): {r2_real_actual:.4f}")
        print(f"R^2      (поточний      vs      оптимальний      провайдер):
{r2_model_actual:.4f}")
        print(f"Середнє відхилення (приріст) ймовірності успіху оптимального
провайдера від історичного: {mean_diff:.4f}\n")
        results[name] = {
            'Precision': precision_score(y_test, y_pred, average='weighted'),
            'Recall': recall_score(y_test, y_pred, average='weighted'),
            'R2_real': r2_real_actual,
            'R2_actual': r2_model_actual,
            'Mean_diff': mean_diff
        }
        if hasattr(best_model, 'feature_importances_'):
            importances = best_model.feature_importances_
            feature_names = X.columns
        elif hasattr(best_model, 'coef_'):
            importances = np.mean(np.abs(best_model.coef_), axis=0)
            feature_names = X.columns
        else:
            importances = None
        if importances is not None:
            fi_df = pd.DataFrame({
                'Feature': feature_names,
                'Importance': importances
            }).sort_values(by='Importance', ascending=False).head(10)
            plt.figure(figsize=(10, 6))
            sns.barplot(data=fi_df,       x='Importance',       y='Feature',
palette='viridis')
            plt.title(f"Топ-10 важливих ознак для моделі: {name}")
            plt.xlabel("Важливість")
            plt.ylabel("Ознака")
            plt.tight_layout()
            plt.show()
    results_df = pd.DataFrame(results).T
    print("\nПорівняння моделей:")
    print(results_df.sort_values(by='R2_real', ascending=False))
```