

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
Національний університет «Острозька академія»
Економічний факультет
Кафедра економіко-математичного моделювання та інформаційних технологій

Кваліфікаційна робота

на здобуття освітнього ступеня магістра

на тему: «Методи та моделі оцінки ризиків цифрового маркетингу»

Виконала:

студентка 2 курсу, групи МЕК-21
спеціальності 051 «Економіка»
освітньо-професійної програми «Економічна кібернетика»
другого (магістерського) рівня вищої освіти
Грабовська Анна Андріївна

Керівник:

кандидат економічних наук, доцент
Новоселецький Олександр Миколайович

Рецензент:

кандидат економічних наук, доцент кафедри
теоретичної та прикладної системотехніки
Харківського національного університету імені
Н.В.Каразіна
Чуб Ольга Ігорівна

"РОБОТА ДОПУЩЕНА ДО ЗАХИСТУ"

**Завідувач кафедри економіко-математичного моделювання
та інформаційних технологій _____ (проф. Ольга КРИВИЦЬКА)**

Протокол № 5 від 05 грудня 2024 р.

Острог, 2024

МІНІСТЕРСТВО ОСВІТИ І НАУКИ
НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ «ОСТРОЗЬКА АКАДЕМІЯ»

Економічний факультет
Кафедра економіко-математичного моделювання та інформаційних технологій
Спеціальність 051 «Економіка»
Освітньо-професійна програма «Економічна кібернетика»

ЗАТВЕРДЖУЮ
Завідувач кафедри економіко-
математичного моделювання та
інформаційних технологій

проф., д.е.н. Ольга КРИВИЦЬКА
“ ____ ” _____ 202__ року

**З А В Д А Н Н Я
НА КВАЛІФІКАЦІЙНУ РОБОТУ СТУДЕНТА**

Грабовської Анни Андріївни

1. Тема роботи: **Методи та моделі оцінки ризиків цифрового маркетингу.**

керівник роботи: *Новоселецький Олександр Миколайович, кандидат економічних наук, доцент.*

затверджено наказом ректора Національного Університету «Острозька академія» від 03 листопада 2023 р. №98.

2. Термін здачі студентом закінченої роботи на кафедру: *05 грудня 2024 року*

3. Вихідні дані до роботи: *емпіричні дані з email-кампаній (характеристики розсилок, типи взаємодій, географічні фактори, сегментація клієнтів, тематика кампаній), науково-публіцистичні праці з дослідження та оцінки ризиків у маркетингових кампаніях, наукові праці з побудови ансамблевих моделей.*

4. Зміст розрахунково-пояснювальної записки (перелік питань, які потрібно розробити): *1) теоретичні засади оцінки ризиків у маркетингових кампаніях; 2) методичні засади застосування математичних моделей для оцінки ризиків; 3) оцінювання ефективності маркетингових кампаній; 4) розробка рекомендацій для мінімізації ризиків у маркетингових кампаніях.*

5. Перелік графічного матеріалу (з точним зазначенням обов'язкових креслень): *схематично представити ризики в цифровому маркетингу, візуалізувати*

результати моделювання за допомогою матриці невідповідності та таблиці з порівняльними показниками точності, повноти та F1-міри для всіх моделей, відобразити аналіз порівняння моделей та їх ефективності.

6. Консультанти розділів роботи

Розділ	Прізвище, ініціали та посада консультанта	Підпис, дата	
		завдання видав	завдання прийняв
Розділ 1	Новоселецький О.М, кандидат економічних наук, доцент кафедри економіко-математичного моделювання та інформаційних технологій	26..02.2024	26..02.2024
Розділ 2	Новоселецький О.М, кандидат економічних наук, доцент кафедри економіко-математичного моделювання та інформаційних технологій	01.05.2024	01.05.2024
Розділ 3	Новоселецький О.М, кандидат економічних наук, доцент кафедри економіко-математичного моделювання та інформаційних технологій	10.06.2024	10..06.2024

7. Дата видачі завдання: 01 грудня 2023 р.

КАЛЕНДАРНИЙ ПЛАН

№ з/п	Назва етапів кваліфікаційної роботи	Строк виконання етапів роботи	Примітка
1.	Вивчення літератури	до 01.02.2024	
2.	Розробка змісту (плану)	до 20.02.2024	
3.	Ознайомлення керівника з текстом кваліфікаційної роботи (чорновий варіант):		
3.1	Розділ 1	до 11.09.2024	
3.2	Розділ 2	до 16.10.2024	
3.3	Розділ 3	до 11.11.2024	
4.	Ознайомлення керівника з текстом кваліфікаційної роботи із врахуванням зауважень	до 20.11.2024	
5.	Попередній захист кваліфікаційної роботи	до 29.11.2024	
6.	Реєстрація на Moodle, перевірка на рівень унікальності	до 01.12.2024	
7.	Рецензування кваліфікаційної роботи	до 05.12.2024	
8.	Здача роботи на кафедрі	до 05.12.2024	

Студентка  Анна ГРАБОВСЬКА

Керівник роботи _____ Олександр НОВОСЕЛЕЦЬКИЙ

АНОТАЦІЯ
кваліфікаційної роботи
на здобуття освітнього ступеня магістра

Тема: **«Методи та моделі оцінки ризиків цифрового маркетингу»**

Автор: Грабовська Анна Андріївна

Науковий керівник: кандидат економічних наук, доцент Новоселецький
Олександр Миколайович

Захищена «.....».....2024 року.

Короткий зміст праці: Використання методів машинного навчання, таких як Random Forest, штучні нейронні мережі (ANN) та метод Монте-Карло, є актуальним та новаторським, оскільки дозволяє точно оцінювати ризики цифрового маркетингу в умовах обробки великих обсягів неоднорідних даних.

У першому розділі здійснено аналіз теоретичних та методологічних аспектів оцінки ризиків у цифровому маркетингу. Розглянуто класифікацію ризиків, визначено їх основні характеристики та обґрунтовано вибір сучасних методів моделювання ризиків. Розкрито важливість переходу від традиційних до інноваційних підходів.

У другому розділі викладено основи застосування методів машинного навчання та симуляційного моделювання для аналізу ризиків. Охарактеризовано математичні основи моделей Random Forest, ANN та методу Монте-Карло. Проведено моделювання для оцінки впливу різних факторів.

У третьому розділі виконано порівняльний аналіз побудованих моделей. Розроблено рекомендації для мінімізації ризиків, включно з оптимізацією часу відправки листів та сегментації клієнтів. Побудовано алгоритм практичного застосування моделей для підвищення ефективності маркетингових кампаній.

Ключові слова: оцінка ризиків, цифровий маркетинг, моделі прогнозування, email-маркетинг, оптимізація кампаній, управління ризиками, аналіз даних.

 Анна ГРАБОВСЬКА

ANNOTATION
qualification work
to obtain a master's degree

Topic: «**Methods and models for risk assessment in digital marketing**»

Author: Hrabovska Anna

Academic supervisor: PhD in Economics, Associate Professor Oleksandr Novoseletskyi

Protected by ".....".....2024.

Summary of the work: The use of machine learning methods such as Random Forest, Artificial Neural Networks (ANN), and the Monte Carlo method is both relevant and innovative, as it enables accurate risk assessment in digital marketing under conditions of processing large volumes of heterogeneous data.

The first chapter analyzes the theoretical and methodological aspects of risk assessment in digital marketing. The classification of risks is considered, their main characteristics are identified, and the choice of modern risk modeling methods is justified. The importance of transitioning from traditional to innovative approaches is highlighted.

The second chapter outlines the fundamentals of applying machine learning methods and simulation modeling for risk analysis. The mathematical foundations of Random Forest, ANN, and the Monte Carlo method are described. Modeling is conducted to evaluate the impact of various factors.

The third chapter performs a comparative analysis of the constructed models. Recommendations for risk minimization are developed, including optimizing email dispatch timing and client segmentation. An algorithm for the practical application of the models is constructed to enhance the effectiveness of marketing campaigns.

Keywords: risk assessment, digital marketing, predictive models, email marketing, campaign optimization, risk management, data analysis.



Anna HRABOVSKA

Зміст

ВСТУП.....	5
РОЗДІЛ 1.	8
ТЕОРЕТИЧНІ ТА МЕТОДОЛОГІЧНІ АСПЕКТИ ОЦІНКИ РИЗИКІВ У ЦИФРОВОМУ МАРКЕТИНГУ.....	8
1.1. Класифікація та характеристика ризиків у цифровому маркетингу.....	8
1.2. Огляд та обґрунтування вибору математичних моделей і методів для оцінки ризиків у цифровому маркетингу	24
1.3. Формулювання гіпотез дослідження.....	33
РОЗДІЛ 2.	40
МОДЕЛЮВАННЯ ТА ОЦІНКА РИЗИКІВ МАРКЕТИНГУ ЗА ДОПОМОГОЮ МАТЕМАТИЧНИХ МОДЕЛЕЙ ТА МЕТОДІВ.....	40
2.1. Побудова моделі «Random Forest» для оцінки ризиків у цифровому маркетингу.....	40
2.2. Моделювання ризиків за допомогою штучних нейронних мереж (ANN)....	50
2.3. Використання методу Монте-Карло для симуляції ризикових подій.	59
2.4. Комбінування моделей та аналіз результатів.....	68
РОЗДІЛ 3.	78
ОЦІНКА РИЗИКІВ ТА РОЗРОБКА РЕКОМЕНДАЦІЙ ДЛЯ ПІДПРИЄМСТВА	78
3.1. Аналіз результатів та їх відповідність фактичним даним.	78
3.2. Розробка рекомендацій використання запропонованих моделей для мінімізації ризиків у маркетингових кампаніях.	90
3.3. Оцінка економічної ефективності запропонованих заходів.....	98
ВИСНОВКИ	108
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ.....	111
ДОДАТКИ	114

ВСТУП

Враховуючи сучасні умови економіки, цифровий маркетинг є найбільш популярним та ефективним інструментом комунікації з клієнтами. Компанію мають можливість достатньо швидко та з невеликими фінансовими затратами доносити інформацію до споживачів, що збільшує їх залученість та підвищує конверсію. Незважаючи на це, великий обсяг інформації, який оточує споживачів, зростає. Це призводить до збільшення конкуренції за увагу клієнта, підвищує вимоги до якості контенту.

Здійснення маркетингової діяльності в цифровому середовищі є досить ризиковим, що обумовлює необхідність постійного моніторингу, ідентифікації та управління ризиками, що й обумовлює актуальність дослідження. Зазвичай традиційні способи оцінки ризиків не здатні бути достатньо точними та гнучкими для опрацювання складних, неоднорідних даних. Як результат виникає потреба у пошуку нових підходів за використанням машинного навчання, як от Random Forest, штучні нейронні мережі (ANN) та метод Монте-Карло. Залучення таких методів дозволить краще проаналізувати великий обсяг даних, визначити закономірності та зробити точніші прогнози відносно ризиків у цифровому маркетингу. Компанії матимуть можливість краще планувати стратегії маркетингу, що підвищить ефективність та знизить ймовірність негативних кейсів. Вкрай важливо щоб такі підходи були досліджені та розроблені в умовах сучасних викликів. Щоб підвищити конкурентоспроможність компаній на ринках та зробити внесок у дослідження цієї теми потрібно розробити моделі для оцінки та управління ризиками.

Мета дослідження полягає у розробці моделі для оцінки ризиків у цифровому маркетингу завдяки застосуванню сучасних математичних моделей і методів аналізу даних. Результат має надати точніше прогнозування основних ризиків, таких як потрапляння електронних листів до спам-фільтрів, низький рівень відкриттів та високий рівень відписок.

Мета дослідження безпосередньо включає в себе:

- Оцінку моделей Random Forest і штучних нейронних мереж (ANN) шляхом їх розробки та застосування для аналізу впливу різних факторів на показники результатів маркетингових кампаній.

- Симуляцію можливих сценаріїв виникнення ризиків та визначення ймовірності їх настання, що забезпечить більш глибоке розуміння потенційних результатів шляхом інтеграції методу Монте-Карло.

- Розробку методичних рекомендацій на основі отриманих результатів моделювання, які сприятимуть мінімізації ризиків та підвищенню ефективності маркетингових кампаній.

Також важливим є практичне застосування моделей для оцінки ризиків у реальних умовах, що дозволить створити інструменти для прийняття більш обґрунтованих рішень та покращення управління маркетинговими кампаніями.

Для досягнення поставленої мети необхідно виконати наступні завдання:

- Опрацювати та дослідити оцінку ризиків з теоретичної точки зору.
- Визначити ключові ризики цифрового маркетингу.
- Обґрунтувати вибір методів моделювання ризиків і пояснити їх застосування.
- Провести моделювання на основі Random Forest та ANN для аналізу взаємодії користувачів маркетинговими кампаніями.
- Використати метод Монте-Карло для оцінки ймовірності ризиків та їхнього впливу.
- Зробити порівняльний аналіз отриманих результатів та їх відповідність фактичним даним.
- Розробити рекомендації для мінімізації ризиків у маркетингових кампаніях.

Об'єктом дослідження є процеси маркетингових кампаній у сучасному цифровому середовищі.

Предмет дослідження: Предметом дослідження є методи оцінки ризиків, зокрема математичні моделі, які використовуються для аналізу ризиків у маркетингу.

Методи дослідження: У процесі дослідження використовувались методи машинного навчання (Random Forest, штучні нейронні мережі) та метод Монте-Карло. Використання цих методів дозволило змодельовати ймовірність виникнення ризиків, визначати важливість ознак та оцінити точність прогнозів для ухвалення рішень у сфері маркетингу.

РОЗДІЛ 1.

ТЕОРЕТИЧНІ ТА МЕТОДОЛОГІЧНІ АСПЕКТИ ОЦІНКИ РИЗИКІВ У ЦИФРОВОМУ МАРКЕТИНГУ

1.1. Класифікація та характеристика ризиків у цифровому маркетингу

Враховуючи важливість раннього та швидкого розвитку транзакцій, цифровий маркетинг як концепція висвітлює набір профільних процесів, які охоплюють всі цифрові канали, доступні для просування продукту або послуги або для створення цифрового бренду.

З економічної практики, цифровий маркетинг спочатку розвивався навколо вебсайтів, щоб згодом продемонструвати стійкість, перенаправляючи трафік на рекламну індустрію. За словами Роба Стокса, цифровий маркетинг можливий через Інтернет як робоче середовище для демонстрації [20].

Цифровий маркетинг призвів і призведе до ключових змін не лише в бізнесі, але й у поведінці споживачів. Для цього дуже важливо мати унікальну платформу для компаній, вміти визначити та розуміти потреби клієнтів і створювати можливості для них на основі часу і місця.

За останнє десятиліття цифровий маркетинг безперервно та стрімко розвивався. Про це свідчить інтенсивне використання інтернету багатьма компаніями у світі, переважно для реклами або корпоративного просування. Компанії в повній мірі використали і повністю задіяли у своїй функціональній системі всі можливості Інтернету, вважаючи його сучасним засобом, абсолютно необхідним при проектуванні та розвитку бізнесу, характерного для цифрової епохи.

Для того, щоб зрозуміти механізм цифрового маркетингу, ми коротко представимо його сфери [17].

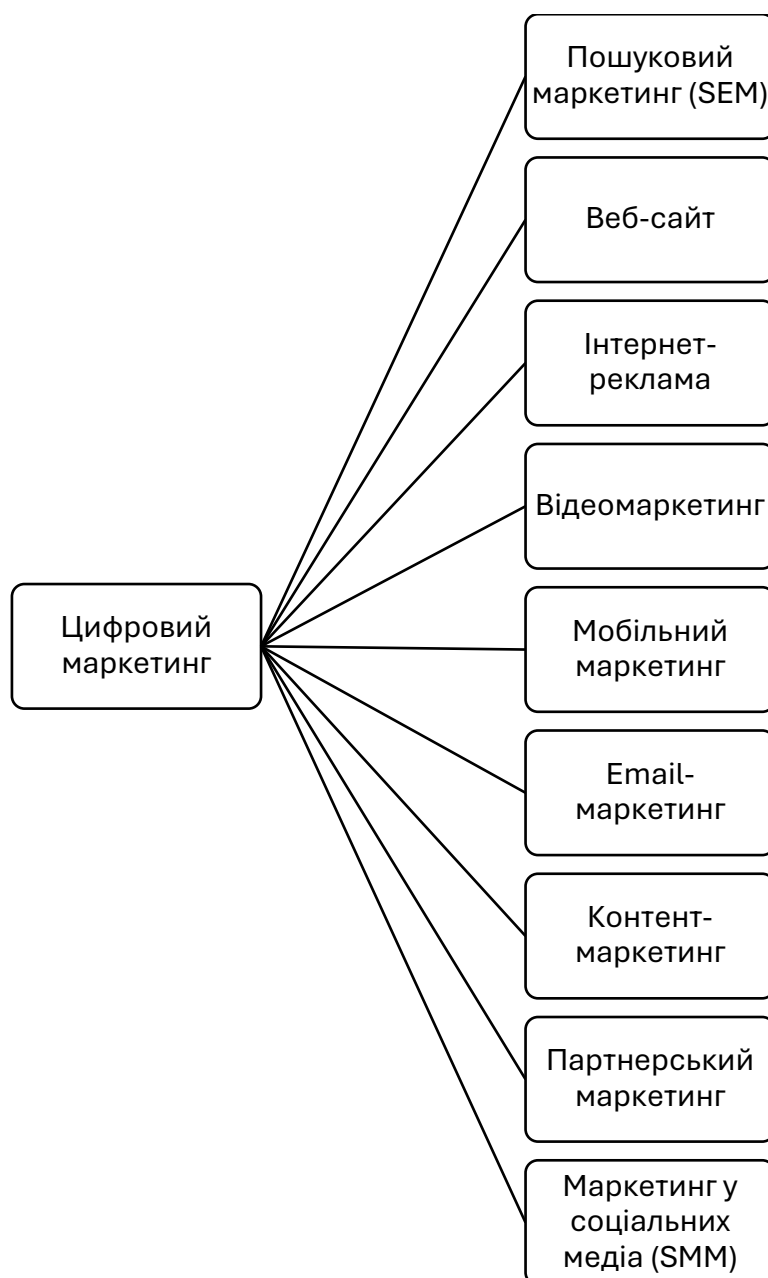


Рис. 1. 1 Смарт-схема сфер маркетингу.

Веб-сайт - це або особистий сайт, корпоративний веб-сайт або навіть особистий блог розміщений на автоматизованій маркетинговій платформі. Для компанії веб-сайт корисний як відправна точка, так і як пункт призначення. Наприклад, маркетологи проводять цифрову маркетингову кампанію з просування веб-сайту (відправна точка) для того, щоб отримати більше відвідувачів (пункт призначення).

Існує кілька доступних варіантів, які складають комунікаційний мікс, що підходить для збільшення кількості відвідувачів на сайті компанії.

Пошуковий маркетинг (SEM) – є основою інтернет-маркетингу завдяки двом важливим функціям:

- «Пошукова оптимізація (SEO)»;
- «Платна пошукова реклама» (PSA).

SEO передбачає оптимізацію методів для отримання більш високих позицій в пошукових системах, відповідно до варіантів продуктів, послуг, пропозицій, постачальників, дистриб'юторів, клієнтів, кінцевих користувачів тощо.

PSA призначене для залучення відвідувачів для розміщення цікавих для них оголошень у пошукових системах. Найпопулярнішим підходом є «Оплата за клік» (PPC), а Google Adwords є найпопулярнішим інструментом, який дозволяє користувачам розміщувати рекламу в Google та партнерських мережах Google.

Маркетинг у соціальних медіа (SMM) – ще одна функціональна частина цифрового-маркетингу, яка стає досить важливою в останні роки. З функціональної точки зору соціальні засоби масової інформації представляють «групу інтернет-базованих додатків, які дозволяють створювати і обмінюватись контентом користувачів». Маркетингові дії в соціальних мережах - це можливості, які включають текст, зображення, відео для всіх бізнес-партнерів, розроблених відповідно до маркетингових методів.

Мільйони користувачів проводять значну частину свого часу (професійного чи особистого) у Facebook, Twitter, Google+, Instagram, Pinterest, YouTube, LinkedIn та інших SSM-платформах. Це спонукає маркетологів компаній приділяти належну увагу SSM і включити його в цифрові маркетингові кампанії.

Партнерський маркетинг. Багато експертів у цій галузі стверджують, що цей вид маркетингу працює у відносних термінах наступним чином:

- трейдер або спонсор, який створив партнерську програму;
- посередник, людина, яка приймає умови партнерської програми;
- власник мережі партнерської програми;
- власник мережі (афіліат).

Посередник слідкує за тим, щоб відповідати вимогам партнерського маркетингу, встановленим бізнесом (торговцем), і бізнесу (торговця), за що отримуватиме комісію за кожного відвідувача сайту компанії (яка застосовує метод).

Контент-маркетинг. У попередньому підході до SEO в основному виділяли можливість створення посилань шляхом застосування методів пошукової оптимізації. Починаючи з 2011 року ця тенденція змінилася, і контент-маркетинг став новим SEO.

Цей аспект розкриває важливість добре розробленого контенту як основи SEO, що призводить до збільшення присутності в Інтернеті, а не через способи створення дешевих посилань.

Email-маркетинг є одним з найбільш важливих інструментів будь-якої цифрової маркетингової кампанії. Незважаючи на зростаючий вплив соціальних мереж, електронна пошта все ще залишається найбільш ефективним способом перетворення відвідувачів або читачів на клієнтів.

Розглядаючи маркетинг електронною поштою як концепцію поняття, то це процес збору електронних адрес від поточних та потенційних покупців (фізичних та/або юридичних осіб), зацікавлених у продуктах чи звертати більше уваги на послуги компанії, з метою надсилання їм інформаційних розсилок, пропозицій тощо.

Мобільний маркетинг. Все більше користувачів мобільних пристроїв відвідують веб-сайти для інтеграції в соціальні мережі або для ідентифікування продуктів (послуг), щоб купити їх негайно або пізніше. Насправді, за оцінками, 30% трафіку на веб-сайтах припадає на відвідувачів з мобільних телефонів [24].

На думку експертів [24] веб-сайти повинні мати сумісні з мобільними пристроями версії. Власник мобільного телефону може стати потенційним покупцем, а за короткий час - клієнтом компанії, що продає продукт або послугу.

Основною перевагою цього методу є зниження витрат і ширше охоплення, але цей метод має свої обмеження, оскільки онлайн-клієнт може легко проігнорувати отриманий рекламний контент або помістити до спаму.

Відеомаркетинг. Відео можна додати на сайт, щоб збільшити трафік. Повідомлення, що надсилаються цільовій аудиторії повинні бути релевантними для задоволення її потреби, а отже, відео має бути розроблене таким чином, щоб відповідати сприйняттю та здатністю розуміння як вже існуючих клієнтів, так і потенційних, інакше реалізація цієї ініціативи буде марною.

Корпоративне відео. Цей метод дозволяє розробляти спільні бізнес-дії в інтерактивному режимі онлайн відео, а тому є корисним у донесенні бажаного повідомлення до цільової аудиторії, маючи можливість легкого та економічно ефективного спільного використання з точки зору створення, наприклад можливість глобального доступу.

Інтернет-реклама. Це найпоширеніша техніка, яка використовується в цифровому маркетингу, де маркетологи використовують віртуальний простір для передачі та просування маркетингового контенту на веб-сторінках, створених виключно для того, щоб привернути увагу інтернет-користувачів; підвищити впізнаваність бренду та стимулювати продажі. Цей метод передбачає розміщення рекламних повідомлень на екрані комп'ютера.

Якщо звичайний маркетинг створює попит, то цифровий маркетинг веде до створення попиту, використовуючи можливості Інтернету. Інтернет-бізнес може отримати цінність у вигляді часу, уваги та підтримки з боку споживача. Для користувача цінність може бути додана у вигляді, зокрема, розваг.

Інтернет пропонує нову парадигму того, як споживачі зв'язуються з брендами та один з одним. З цією метою бренди будують свою лояльність через користувачів, які люблять їхні продукти чи послуги. Користувачі закохуються в продукти та послуги, коли вони пристосовуються до їхніх потреб, а не для того, щоб служити бренду.

Більше, ніж будь-який інший вид маркетингу, цифровий маркетинг можна виміряти. Це дає бренду можливість створювати оптимізований профільний досвід, розроблений спеціально для споживачів.

Очевидно, що Інтернет кардинально змінив звичайний процес продажу і те, як покупець та продавець взаємодіють один з одним. Тепер клієнти мають

багато можливостей, вплив і реальну владу в інтернеті, адже бренди просуваються засобами, придатними для електронних транзакцій на потенційних ринках. З цією метою традиційним маркетинговим агенціям довелося адаптувати свої нові практики та розширити свої цифрові маркетингові операції.

Таким чином, найважливішою місією для маркетологів у цифровому середовищі є просування та торгівля брендами, а також формування лояльності клієнтів. Відомо, що люди занадто насичені величезною хвилею щоденної реклами в онлайн-середовищі та все більш турбулентному світі. Це означає нові виклики для цифрового маркетингу, який зараз більше, ніж будь-коли, повинен бути здатним задовольнити онлайн-споживачів.

Глобалізація додала цифровому маркетингу нових рівнів складності, вимагаючи глибокого розуміння різноманітних ринків, крос-культурних нюансів та регуляторних середовищ.

Нові технології змінили методи збору та розповсюдження новин. Ці технології змінили потік інформації; замінили централізовану модель «зверху-вниз» і створили спеціальні мережі розповсюдження інформації[14]. Непослідовність виникає тоді, коли подія розвивається, але інформація не оновлюється регулярно.

Соціальні медіа є новим джерелом інформації для покращення ситуаційної обізнаності, але це не все. Багато людей використовують соціальні мережі як єдине джерело інформації, тому вони є легкою і вразливою мішенню для зловмисних атак або дезінформації. Навіть якщо першочерговою метою платформ не було просування брендів, сьогодні майже всі компанії використовують їх як комунікаційну платформу, оскільки вбудовані алгоритми, які використовують організації, дозволяють краще і простіше таргетувати потенційних клієнтів, причому з меншими витратами.

Ми хочемо визначити чотири основні ризики, спричинені новими технологіями, з якими мають боротися маркетологи, щоб зберегти безперервність бізнесу та забезпечити організаційну стійкість:

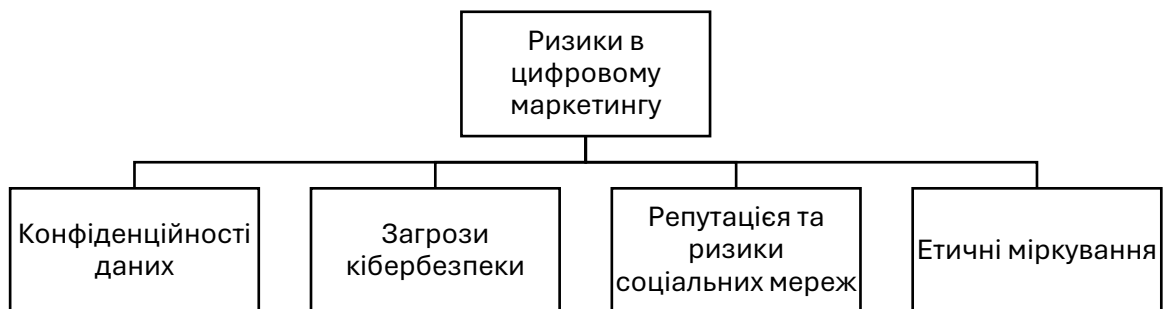


Рис.1.2 Схема ризиків в цифровому маркетингу.

Щодо конфіденційності даних та дотримання законодавства, бізнес більше не обмежується локальними сферами; він є невід'ємною складовою глобальної мережі постачальників, партнерів та клієнтів. Подолання викликів глобалізації вимагає стратегічного підходу, який визнає взаємопов'язаність ринків. Оскільки всі компанії прагнуть проводити персоналізовані кампанії, збір та використання даних про споживачів є важливим для покращення цифрових маркетингових кампаній. Однак неналежне поводження з цими даними створює значні ризики.

Каліфорнійський закон про конфіденційність споживачів (CCPA) [5] та Загальний регламент про захист даних (GDPR) Європейського Союзу [6] - два приклади нормативно-правових актів, що стосуються конфіденційності споживачів, які були прийняті останнім часом. Обидва закони зосереджені на правах користувачів на недоторканність приватного життя та поверненні контролю над персональними даними в руки користувачів. У той час як GDPR застосовується до компаній, які володіють персональними даними резидентів ЄС/ЄЕЗ, CCPA застосовується до комерційних компаній, які володіють персональними даними резидентів Каліфорнії.

Порушення законодавства про захист даних може мати серйозні наслідки. Це можуть бути великі штрафи, судові позови та репутаційні збитки. Наприклад,

штрафи у випадку CCPA можуть сягати 2500 доларів США за кожне порушення і 7500 доларів США за кожне умисне порушення. У випадку GDPR штрафи можуть сягати 2% від річного обороту за менш серйозні порушення і до 20 мільйонів євро або 4% від річного обороту, залежно від того, яка сума є більшою, за серйозні порушення [6]. Конфіденційність даних клієнтів навіть стала геополітичним питанням, оскільки законодавчі проекти ЄС, такі як Закон про цифрові ринки [6] та Закон про цифрові послуги [6], що запроваджують додатковий захист даних фізичних та юридичних осіб ЄС (під заголовком «суверенітет даних»), на які скаржилися американські компанії, такі як Google та Meta, є цілеспрямованим обмеженням доступу до інформації, на якій вони будують свій рекламний бізнес сьогодні та бізнес у сфері ІІІ.

Безперервність бізнесу опиняється під загрозою, коли довіра клієнтів підривається через ймовірне неправомірне використання конфіденційної інформації. Фактична втрата цих даних може призвести до прямих фінансових збитків, реакції ринку (втрата ринкової вартості) та невизначеності на економічні перспективи.

Стратегії пом'якшення наслідків включають впровадження надійних заходів захисту даних, регулярне оновлення політики конфіденційності та проведення комплексних аудитів для забезпечення її дотримання. Компанії повинні інформувати співробітників про практику конфіденційності та використовувати технологію шифрування для захисту інформації про клієнтів.

Створення чітких і прозорих каналів комунікації з клієнтами щодо використання даних і політики конфіденційності може допомогти побудувати довіру та зменшити потенційний вплив недотримання вимог.

Крім того, для кращого захисту даних ми вважаємо, що компаніям варто використовувати додатки на основі технології блокчейн. Блокчейн - це нова технологія, яка застосовується в багатьох сферах економіки, адміністрації та управління. Вона пропонує можливість здійснювати транзакції або зміни в базі даних без посередників і з підвищеним рівнем безпеки, революціонізуючи моделі організації та надання масових послуг. На найпростішому рівні блокчейн - це

розподілена база даних між багатьма учасниками відповідної мережі, яка використовує специфічні алгоритми, що постійно розвиваються, для перевірки та розповсюдження змін в базі даних автоматично, без центрального органу[27]. Таким чином, формується нова промислова революція, заснована на вирішенні важливої проблеми в організації людської діяльності - проблеми довіри і контролю. Застосування блокчейну набагато ширше, ніж можна було б припустити, якби не зацикленість ЗМІ на криптовалютах, таких як біткойн та інших інструментах фінансових спекуляцій. Ми є свідками стрімкого розвитку нового бізнес-сектору, що розробляє додатки для «смарт-контрактів», управління ланцюгами поставок, спрощення фінансових транзакцій тощо.

Смарт-контракти можуть бути найкращим рішенням для захисту конфіденційних даних клієнтів або замовників. Смарт-контракт - це програмне забезпечення, яке зберігає політику і правила для узгодження умов і дій між сторонами. Концепція смарт-контракту була введена Ніком Сабо в 1994 році, де комп'ютеризований протокол транзакцій виконує умови контракту, маючи на меті зменшити витрати і затримки традиційних контрактів і задовольнити загальні договірні умови [21]. Смарт-контракт включає, по-перше, функції, які отримують вхідні параметри контракту і викликаються при здійсненні транзакцій. Друга - це змінна стану, яка залежить від логіки, розробленої у функціях. Смарт-контракти в блокчейні дозволяють створити самоврядне партнерство з обов'язковими до виконання правилами взаємодії без потреби в центральному органі влади.

Зростаюча діджиталізація маркетингової діяльності також наражає компанії на загрози кібербезпеки. Інформаційна безпека передбачає забезпечення конфіденційності, цілісності та доступності даних та інформації під час зберігання, обробки або передачі за допомогою процедур, впроваджених в організації, та технологій, що використовуються.

Нижче наведено основні категорії ІТ-загроз, з якими може зіткнутися організація:

- Шкідливе програмне забезпечення, яке вимагає встановлення програми, здатної відстежувати активність на певному пристрої. Цей тип атаки

здійснюється за допомогою електронних листів, в яких вимагається надання даних, або навіть тих, які заохочують до переходу на певні сайти за посиланнями, надісланими в електронному листі. Ці посилання можуть також поширюватися через соціальні мережі (Facebook і WhatsApp).

- Програми-вимагачі - категорія шкідливих програм, які блокують доступ до комп'ютера до сплати викупу або шифрують окремі файли, а потім вимагають сплатити певну суму грошей протягом декількох днів, щоб відновити дані або доступ до відповідних даних. Як і у випадку зі шкідливими програмами, атаки з вимогами викупу найчастіше здійснюються через електронну пошту або соціальні мережі. Таким чином, реклама, яка обіцяє певні безкоштовні товари або послуги, наприклад, безкоштовне завантаження книг або фільмів після попереднього встановлення певного програмного забезпечення, швидше за все, атакує комп'ютерну систему. Або сайти, які просять перейти за іншим посиланням, щоб отримати доступ до інформації, яку ви шукаєте за допомогою пошукових систем, також становлять кібернетичну небезпеку.

- Фішинг - термін походить з англійської мови, від дієслова fishing (ловити рибу); точніше, зловмисники використовують приманку, за допомогою якої переконують користувачів певних соціальних мереж або клієнтів певних інтернет-провайдерів розкрити особисту інформацію, наприклад, пароль від банківських додатків, паролі до певних рахунків тощо.

Різновидом фішингу є смішинг, тобто атака ініціюється за допомогою SMS-повідомлення, а потім продовжується аналогічно до інших видів атак - жертву спонукають перейти за посиланням і розголосити певну інформацію.

Щоб протистояти цим ризикам, компанії повинні інвестувати в передові заходи кібербезпеки. Існує широка обізнаність про брандмауери, системи виявлення вторгнень та протоколи шифрування. Регулярне тестування на проникнення допомагає виявити вразливі місця, що дозволяє компаніям усунути їх до того, як ними скористаються зловмисники. На жаль, внутрішня загроза не менш актуальна, якщо не більш небезпечна. Людська помилка часто є ключовим фактором порушення безпеки, тому дуже важливо навчати співробітників

найкращим практикам кібербезпеки, оскільки запобігти їй дешевше, ніж лікувати. Крім того, компанії можуть передати цей тип ризику третій стороні - компанії, що спеціалізується на кібербезпеці. Надійний план реагування на інциденти забезпечує швидке та ефективне реагування на кіберзагрози, мінімізуючи вплив на бізнес-операції. Однак, на нашу думку, одних лише цих стратегій недостатньо. Тому ми пропонуємо розглянути інновації в системах кібербезпеки компаній.

Прикладом інновацій в кібербезпеці є використання штучного інтелекту і машинного навчання в процесах кібербезпеки. Обіцянок щодо застосування штучного інтелекту в кібербезпеці багато, і в основному вони пов'язані з системами ідентифікації ризиків. Хоча автоматизація забезпечує виявлення будь-яких правопорушень, вона також може захистити цілі атаки. Глибоке навчання використовується для відстеження журналів, транзакцій і даних у реальному часі, щоб виявити загрози в мережі. Здатність неконтрольованого машинного навчання полягає в тому, щоб знаходити всі види невідомих закономірностей і виявляти аномалії. Він може «навчитися» виявляти патерни і сигналізувати про потенційну спробу атаки. Такі інноваційні технології вдосконалюються з кожним днем, надаючи цінну інформацію про кроки, яких можна вжити, щоб уникнути проблем, спричинених витонченими методами зловмисників. Переваги технологій штучного інтелекту в механізмах захисту кібербезпеки виявилися дуже корисними для забезпечення точності та швидкості виявлення загроз кібербезпеки, а також пом'якшення їх наслідків, тим самим покращуючи організацію безпеки, автоматизацію та систему реагування операційного центру безпеки.

Організації, які зазнали невеликих кібератак, отримали ранній і менш витратний сигнал про необхідність інвестувати в кібербезпеку і перетворити свою здатність протистояти цим загрозам на порівняльну перевагу на переповненому ринку. Державне регулювання також може створити рівні умови для всіх суб'єктів, накладаючи на них мінімальні зобов'язання щодо готовності до кібербезпеки, тим самим усуваючи стимули для недоінвестування з метою

збільшення прибутків або зменшення операційних витрат за рахунок майбутньої вразливості.

Управління репутацією та соціальні мережі. Існує кілька прикладів ризиків, які можуть виникнути в цій категорії. Майже всі вони вимагають певного рівня довіри між компаніями та стейкхолдерами - у цьому випадку стейкхолдерами є клієнти, а не акціонери.

Шахрайство в рекламі - поширений ризик у цифровому маркетингу, з яким сьогодні все частіше стикаються, оскільки генеративний ШІ може видавати себе за відомих людей, яким довіряють

Негативні відгуки, негативна реакція в соціальних мережах або вірусні суперечки можуть швидко зашкодити репутації бренду, вплинути на довіру та лояльність клієнтів

Протидія рекламному шахрайству передбачає використання передових інструментів аналітики та верифікації оголошень для виявлення та запобігання шахрайським діям. Компанії повинні співпрацювати з авторитетними рекламними платформами та постійно відстежувати й аналізувати ефективність рекламних кампаній. Інформованість про останні тенденції рекламного шахрайства та найкращі практики галузі дозволяє маркетологам коригувати свої стратегії та вживати превентивних заходів.

Що стосується ризику нанесення шкоди репутації бренду, то реалізована стратегія вимагає проактивного підходу. Компаніям слід відстежувати онлайн-канали на предмет згадок про бренд, відгуків клієнтів та нових тенденцій. Добре розроблений план дій на випадок непередбачуваних ситуацій має важливе значення для швидкого вирішення проблем і мінімізації репутаційних збитків, а також для інвестування в позитивну присутність в Інтернеті завдяки справжньому залученню. Кампанії з корпоративної соціальної відповідальності також є потенційними учасниками. Залучення до соціально відповідального маркетингу, підтримка етичних причин та уникнення кампаній, які експлуатують чутливі питання, є важливими для підтримання позитивного іміджу бренду.

Тим не менш, нові технології, такі як штучний інтелект, що використовується як система раннього попередження, можуть допомогти у виявленні підозрілих дій/реклами на ранніх стадіях. По-перше, штучний інтелект підвищить здатність обробляти дані для отримання дієвої інформації для осіб, які приймають рішення. Наприклад, можна буде ідентифікувати банківські рахунки з підозрілою активністю для подальшого розслідування або виявити підозрілі патерни активності в складних системах, щоб запобігти довгостроковому проникненню злочинців. Це дуже важливо, враховуючи, що за даними Coburn et al., «середній час від проникнення до виявлення злочину корпорацією збільшився в усьому світі до 101 дня»[13]. Таким чином, ШІ буде широко використовуватися в системах безпеки і підтримки прийняття рішень саме тому, що він являє собою крок вперед в аналізі нестандартних ситуацій, з високим ступенем невизначеності і з різноманітними параметрами. Все частіше програмне забезпечення буде становити значну частину доданої вартості і диференціації продуктів завтрашнього дня, і ця реальність буде відчутна і в сфері кібербезпеки. ШІ стане активною частиною боротьби з кібервтручаннями в режимі реального часу, зі швидкістю, яку людина не може відстежити, а може лише проаналізувати подію постфактум.

Крім того, компаніям варто диверсифікувати свої маркетингові канали та стратегії. Це передбачає вивчення різних платформ і технологій, щоб охопити ширшу аудиторію. Інформованість про галузеві тенденції та прийняття гнучкого підходу до маркетингу дозволяють бізнесу швидко адаптуватися до змін у платформах чи технологіях. Побудова надійної омніканальної маркетингової стратегії гарантує, що бізнес не буде надмірно залежним від будь-якого одного каналу, зменшуючи вплив перебоїв у його роботі. Регулярний аналіз ключових показників ефективності (KPI) дозволяє маркетологам визначити сфери для вдосконалення та відповідно скоригувати свою контент-стратегію.

Зменшення ризику розриву в навичках передбачає інвестиції в постійне навчання та розвиток маркетингових команд. Компанії повинні надавати пріоритет найму людей з необхідним досвідом і бути в курсі останніх досягнень

галузі. Співпраця із зовнішніми агентствами або консультантами може доповнити внутрішні навички та принести команді спеціалізовані знання. Створення культури безперервного навчання та адаптивності гарантує, що команда маркетологів буде добре підготовленою, щоб орієнтуватися в постійно мінливому ландшафті цифрового маркетингу.

У 2023 році Forbes заявив: «Етика в цифровому маркетингу стосується моральних принципів, якими керується компанія в Інтернеті». Етичні проблеми в цифровому маркетингу стають все більш помітними, оскільки технології продовжують розвиватися, а маркетологи використовують складні інструменти та стратегії для досягнення своєї цільової аудиторії. Етичні міркування мають вирішальне значення для збереження довіри зі споживачами, захисту приватності користувачів та забезпечення відповідальних і прозорих практик. Forbes вважає, що основними принципами, які супроводжують компанію, що заслуговує на довіру, є [4]:

- Чесна реклама.
- Прозорість у спілкуванні.
- Відповідальне управління даними.
- Повага до приватності клієнтів.
- Чесна конкуренція.
- Відмова від неправдивих обіцянок.

Ми всі можемо погодитися з тим, що цифровий маркетинг часто передбачає збір та використання даних користувачів для таргетованої реклами, проте етичні проблеми виникають, коли компанії не отримують належної згоди на збір даних або коли вони зловживають особистою інформацією. Щоб вирішити цю проблему, компаніям слід надавати пріоритет прозорим практикам роботи з даними, отримувати явну згоду від користувачів і дотримуватися відповідних правил захисту даних.

Водночас, хоча персоналізований маркетинг може покращити користувацький досвід, ця практика викликає етичні питання, коли вона стає надмірно інвазивною або спирається на оманливу тактику. У цьому випадку має

існувати баланс між персоналізацією та повагою до кордонів користувача. Прозорість використання персональних даних і надання користувачам можливості контролювати свої вподобання може допомогти пом'якшити етичні проблеми, пов'язані з таргетингом.

Крім того, маркетингові практики з використанням так званих інфлюенсерів є популярною стратегією, але етичні проблеми виникають, коли інфлюенсери не розкривають інформацію про оплачувані партнерства або коли вони просувають продукти, не маючи справжнього досвіду чи віри в них. З одного боку, компанії повинні надавати пріоритет прозорості у відносинах з інфлюенсерами, і в той же час інфлюенсери повинні бути прозорими у своїх відносинах з брендами, щоб зберегти автентичність і довіру.

Тут також є технологічний вимір, оскільки регулювання ШІ спрямоване на сприяння розвитку «етичного та надійного ШІ». Комісія високого рівня з питань штучного інтелекту Європейського Союзу розробила низку рекомендацій щодо категорій застосувань штучного інтелекту, які поділяються за ступенем небезпеки, а отже, потребують більшого чи меншого нагляду та регулювання. Ці додатки проходять гаму від мінімального ризику до неприйняттого, і багато маркетингово-орієнтованих додатків ШІ потрапляють до категорій обмеженого або високого ризику через потенційну дискримінацію, упередженість, неналежне використання даних, потенційні маніпуляції, відсутність захисту для вразливих категорій населення тощо. У зв'язку з цим існують як директиви та інші форми законодавства, що регулюють ШІ в загальних рисах, так і добровільні кодекси поведінки та стандарти, які можуть використовувати відділи маркетингу, щоб уникнути етичних, юридичних та фінансових пасток.

У наступному розділі ми аналізуємо дані email-маркетингу банку, зосередженому на торгівлі криптовалютою. Саме тому ми вважаємо необхідним виділити нестабільність даної сфери як потенційний ризик для маркетингових компаній підприємства. Криптовалютний ринок відомий своєю високою волатильністю. Значні коливання цін на Bitcoin та Ethereum можуть бути спричинені як глобальними економічними подіями, так і коментарями відомих

інвесторів або рішеннями урядів щодо регулювання криптовалют. Ці фактори роблять криптовалютний ринок одночасно привабливим і ризикованим для інвесторів. Ба більше, динамічність цього сектору часто приваблює спекулятивний капітал, що ще більше посилює його нестабільність і непередбачуваність. У цьому середовищі дуже важливо підтримувати цікавість та довіру клієнтів, залишатись лояльними та мати план дій у випадку можливих коливань ринку.

Щоб залишатися конкурентоспроможними в економічному середовищі, яке постійно змінюється, організації повинні бути стійкими, в тому числі завдяки стратегіям розуміння поведінки споживачів і того, як вони приймають рішення про покупку, що полегшить їм задоволення потреб і бажань споживачів. Розуміння поведінки споживачів є темою, що цікавить як компанії, так і дослідників. Отримання більш детальної інформації про те, як клієнти відчують, діють та обирають з доступних варіантів, допоможе маркетологам покращити свої кампанії.

Цифрові технології відкрили нові канали для маркетингових комунікацій, інсайтів на основі даних та посилення залучення клієнтів. Онлайн-платформи, аналітика даних та автоматизація відіграють ключову роль у зміні маркетингового ландшафту. Маркетологи повинні адаптуватися до цієї еволюційної парадигми, використовуючи цифрові інструменти для підвищення видимості, оптимізації процесів і взаємодії з бізнес-клієнтами відповідно до сучасних очікувань, але в той же час приділяти увагу таким елементам, як конфіденційність даних і соціальні медіа, щоб зберегти хорошу репутацію і завоювати довіру клієнтів. У цьому цифровому середовищі кібербезпека стала пріоритетом у швидко мінливому і складному середовищі безпеки, яке переповнене гібридними загрозами. Інвестиції в кібербезпеку можуть гарантувати власникам, що вони мають стійку компанію, яка може зберегти безперервність бізнесу в найкращих умовах.

1.2. Огляд та обґрунтування вибору математичних моделей і методів для оцінки ризиків у цифровому маркетингу

У цифровому маркетингу існує велика кількість факторів, які можуть призвести до виникнення ризиків, таких як зміна алгоритмів соціальних мереж, репутаційні кризи, технічні збої або зловживання конфіденційними даними користувачів. Для ефективного управління цими ризиками необхідно використовувати сучасні математичні моделі та методи, які дозволяють не лише прогнозувати можливі ризикові події, але й оцінювати їхній вплив на бізнес у реальному часі.

Ризик у маркетингу пов'язаний, з багатьма його функціями - організацією, плануванням, управлінням, а також контролем. Усі ці функції несуть певний ступінь ризику, що призводить до необхідності створення адаптивної до ризиків подібного роду системи.

Сучасному світі притаманні висока динаміка подій і значний ступінь невизначеності, тож фактор ризику - суттєвий і невід'ємний атрибут існування. При цьому фактично кожна компанія змушена вести роботу з управління ризиками.

Маркетинг же, у свою чергу, є однією з найважливіших концепцій в управлінні підприємством. Він спрямований як на виявлення і задоволення потреб покупців, так і на досягнення основних цілей організації. Найуспішнішою буде та компанія, яка задовольнить потреби покупців найекономнішим і найприйнятнішим для них способом. Успішна маркетингова діяльність дає змогу організації знайти свій унікальний спосіб надати клієнтам вищу споживчу цінність, яку їхні конкуренти забезпечити не зможуть.

У галузі маркетингу, на жаль, нині комплексному вивченню ризиків не приділяється достатньо уваги. Зазвичай проводяться розробки окремих приватних проблем управління ризиками. У рамках деяких маркетингових досліджень[25, 26, 28], наприклад, приділяється особлива увага різного роду помилкам вибірки під час їхнього проведення, а також розробляються дії,

покликані їх мінімізувати, що безпосередньо не є ризик-менеджментом як таким, але, по суті, є його складовою частиною.

У процесі діяльності фірми ризик-менеджмент і маркетинг тісно пов'язані між собою. Цей зв'язок досить легко побачити на прикладі різних маркетингових компаній, які за своєю суттю покликані відігравати істотну роль у зниженні ризиків підприємства. Таким чином, якоюсь мірою маркетинг є інструментом управління ризиками.

З іншого боку, маркетингова компанія сама несе деякі ризики. Маркетингові ризики - це всі ризики, які притаманні маркетинговій сфері діяльності компанії в цілому. Вони характеризуються ймовірністю виникнення певних подій та їхніх наслідків, які можуть значно ускладнити досягнення цілей на різних етапах маркетингової діяльності фірми.

Таким чином, ризик-менеджмент маркетингових компаній є важливим напрямом, що дає змогу підвищити ефективність діяльності підприємства. Кожен етап маркетингової діяльності має свою специфіку, тому необхідно проводити оцінку ризиків маркетингу на кожному етапі. Найчастіше для виявлення ризиків маркетингових стратегій використовують традиційні методики виявлення економічних ризиків.

При створенні маркетингових стратегій організації в умовах неповної інформації або її відсутності можуть використовуватися такі кількісні методи оцінки ризику, як:

- Методи стохастичного програмування.(додати посилання) По суті, це вивчення задач математичного програмування, де дані мають випадковий характер. Можна виділити три шляхи розв'язання задач стохастичного програмування. Перший - заміна випадкових параметрів їхніми середніми значеннями. Це дає змогу отримати розв'язання задач на рівні математичних очікувань. Інший шлях - вивчення двоетапних задач лінійного програмування з випадковими даними [11]. Під цими задачами розуміють таке: на першому етапі розв'язують задачу з фіксації випадкового попиту, який задають математичним очікуванням; на другому етапі шукають оптимальний вектор управління

виробництвом за фіксованого попиту на цю продукцію. Третій шлях розв'язання розглянутих задач має авторську назву «програмування з імовірнісними обмеженнями». Стохастичне програмування дозволяє працювати із завданнями, де невизначеність може бути описана математично, наприклад, через математичне очікування або ймовірнісні обмеження. Моделі дозволяють знайти оптимальні рішення в умовах невизначеності, враховуючи ризики на різних етапах, зокрема двоетапні задачі лінійного програмування.

Однак, використання середніх значень для випадкових параметрів може призвести до спотворення реальності, оскільки це не враховує можливих крайніх випадків або змін у розподілі. Врахування всієї невизначеності в складних задачах може потребувати обчислювальних потужностей і часу для детального аналізу кожного сценарію.

- Метод «дерева рішень». Його застосовують у тому разі, коли потрібно послідовно ухвалити кілька управлінських рішень. Дерево рішень - графічний метод, він допомагає пов'язати точки ухвалення рішення, можливі стратегії, їхні наслідки з імовірними факторами зовнішнього середовища [19].

Метод дерева рішень дозволяє побудувати зрозумілу структуру ухвалення рішень, особливо коли потрібно врахувати кілька послідовних кроків. Метод добре працює, коли є кілька альтернатив і потрібно оцінити ймовірні наслідки кожної з них.

З іншого боку, кількість рішень і можливих варіантів дуже велика, дерево рішень може стати громіздким і важким для обчислення. Метод обмежений кількістю параметрів, які можна легко візуалізувати і моделювати, що ускладнює його використання в складних системах.

- Метод статистичних рішень. Теорія статистичних рішень є теорією проведення статистичних спостережень, опрацювання цих спостережень та їх використання. Будь-які економічні дані являють собою кількісні характеристики будь-яких економічних об'єктів. Вони формуються під дією безлічі факторів, не всі з яких доступні зовнішньому контролю. Неконтрольовані фактори можуть набувати випадкових значень із деякої множини значень і тим самим

зумовлювати випадковість даних, які вони визначають. Стохастична природа економічних даних зумовлює необхідність застосування спеціальних адекватних їм статистичних методів для їхнього аналізу та обробки. Кількісна оцінка ризику незалежно від змісту конкретного завдання можлива, як правило, за допомогою методів математичної статистики. Головні інструменти цього методу оцінки - дисперсія, стандартне відхилення, коефіцієнт варіації [10].

Метод дає можливість використовувати стандартні інструменти математичної статистики для оцінки ризику та невизначеності. Використання статистичних методів на основі реальних даних дозволяє об'єктивно оцінити ризики та прогнозувати подальші зміни.

Не зважаючи на переваги, метод сильно залежить від доступності та якості статистичних даних, і якщо дані містять помилки або неповні, результати можуть бути неточними. Багато статистичних методів обмежуються аналізом відомих або контрольованих факторів, що може не враховувати непередбачувані зовнішні впливи.

Як ми бачимо, на жаль, перелічені вище методи не завжди задовольняють вимогам забезпечення необхідної в маркетингу надійності оцінок ефективності та ступеня ризику стратегій. За заздалегідь заданого розподілу ймовірності випадкових параметрів як зовнішнього середовища, так і виробничої системи може бути отримано більш точну оцінку ефективності та ризику. Такі розподіли отримують на підставі статистичних даних за попередні часові періоди. У тому разі, якщо підприємство не володіє такою статистикою або вона недостатня для надійної оцінки параметрів розподілу досліджуваних показників, то використовується метод Монте-Карло.

На відміну від стохастичного програмування, яке спрощує проблему шляхом використання середніх значень, метод Монте-Карло моделює численні сценарії на основі випадкових даних, що дає можливість врахувати широкий спектр результатів і точніше оцінити ризики. Цей метод може враховувати будь-яку кількість випадкових факторів і дозволяє симулювати навіть складні та

непередбачувані сценарії, що є складним для методу дерева рішень, особливо у випадку великої кількості рішень і параметрів.

Він дозволяє генерувати випадкові дані на основі різних розподілів і проводити багаторазові симуляції, що забезпечує глибший аналіз ризиків порівняно з методами статистичних рішень, які часто обмежені доступними історичними даними.

Метод Монте-Карло враховує можливі крайні значення результатів, що робить його більш точним у випадках, де потрібно оцінити ризики за найгіршим або найкращим сценарієм, чого не завжди досягають інші методи, особливо стохастичне програмування з його орієнтацією на середні значення.

Модель перераховується після побудови під час кожного нового імітаційного експерименту, в якому на основі генерації псевдовипадкових чисел вибираються значення випадкових змінних. Результати всіх проведених експериментів об'єднуються в одну вибірку та аналізуються за допомогою статистичних методів для отримання розподілу ймовірностей вихідного показника й розрахунку основних показників ризику дослідження.

Метод Монте-Карло успішно використовується в управлінні ризиками маркетингових досліджень. При цьому можна визначити мінімальний рівень ризику за раніше заданої ефективності, і навпаки, вибрати максимальну ефективність стратегії при заданому граничному значенні ризику.

Маркетинг включає в себе численні змінні та невизначеності, такі як поведінка клієнтів, ринкові тенденції та конкуренція. Традиційні моделі часто не можуть охопити весь спектр можливих результатів, що призводить до неоптимальних стратегій. Моделювання методом Монте-Карло, з іншого боку, дозволяє маркетологам це зробити, а саме:

- Оцінити ризик і невизначеність: моделюючи різні сценарії можна визначити потенційні ризики і невизначеності у своїх стратегіях, що дозволяє краще управляти ризиками.

- Оптимізувати розподіл бюджету: моделювання методом Монте-Карло допомагає зрозуміти вплив різних варіантів розподілу бюджету на маркетингові результати, що дозволяє більш ефективно розподіляти ресурси.

- Підвищити точність прогнозів: враховуючи мінливість ключових параметрів, цей метод забезпечує більш точні прогнози, що допомагає у стратегічному плануванні та прийнятті рішень.

Припустимо, компанія планує цифрову маркетингову кампанію і хоче оцінити рентабельність інвестицій (ROI). Ключові змінні, що беруть участь у цьому процесі, включають [15]:

- Коефіцієнт конверсії (CR): Відсоток відвідувачів, які здійснили покупку.
- Середня вартість замовлення (AOV): Середня сума, витрачена на одну покупку.
- Ціна за клік (CPC): Вартість, понесена за кожен клік по оголошенню.
- Кількість кліків (NC): Загальна кількість кліків, яку отримує кампанія.

Кожна з цих змінних є невизначеною і може змінюватися залежно від багатьох факторів. Використовуючи історичні дані, компанія може оцінити розподіл ймовірностей для цих змінних.

У відповідь на обмеження традиційних методів, сучасні моделі машинного навчання, такі як Random Forest і штучні нейронні мережі (Artificial Neural Networks, ANN), пропонують більш інноваційні та ефективні рішення для оцінки ризиків. Random Forest забезпечує точніші прогнози завдяки ансамблевому підходу, який використовує кілька дерев рішень для оцінки ризиків, що дозволяє враховувати широкий спектр факторів.

Random Forest - це надійний алгоритм машинного навчання, який працює шляхом побудови багатьох дерев рішень на етапі навчання і виводить клас, який є модою класів (у класифікації) або середнім прогнозом (у регресії) окремих дерев [9].

У маркетинговій аналітиці це означає здатність обробляти величезні обсяги даних про клієнтів і ринок для точного прогнозування таких результатів, як відтік клієнтів, коефіцієнт конверсії та життєва цінність клієнта.

Random Forest можна ефективно використовувати для:

- Сегментації клієнтів. Random Forest може аналізувати складні дані про клієнтів, щоб визначити окремі групи на основі купівельної поведінки, демографічних даних та інших релевантних характеристик. Це дозволяє маркетологам адаптувати свої стратегії до унікальних уподобань і потреб кожного сегмента.
- Прогнозування відтоку. Виявляючи закономірності в клієнтських даних, які вказують на ймовірність відтоку клієнтів, Random Forest допомагає компаніям проактивно впроваджувати стратегії утримання цінних клієнтів.
- Оптимізації кампаній. Він може прогнозувати ефективність різних маркетингових кампаній у різних каналах і сегментах клієнтів, допомагаючи маркетологам ефективніше розподіляти ресурси і підвищувати рентабельність інвестицій.
- Рекомендацій продуктів. Random Forest може передбачати вподобання клієнтів і ймовірні майбутні покупки, що сприяє розвитку рекомендаційних механізмів, які керують стратегіями перехресних продажів і апселлінгу.
- Аналізу настроїв. Застосовуючи Random Forest до відгуків клієнтів, постів у соціальних мережах та інших текстових даних, він може допомогти оцінити суспільні настрої щодо бренду чи продукту, інформуючи бренд-менеджмент та маркетингові стратегії.

Ансамблевий підхід Random Forest зменшує ризик надмірного пристосування, що робить його прогнози дуже надійними для прийняття стратегічних рішень.

Він може вловлювати складні нелінійні взаємозв'язки між ознаками, що робить його ідеальним для багатогранної природи клієнтських даних.

Random Forest дає уявлення про те, які змінні мають найбільший вплив на прогнозування результатів, спрямовуючи маркетологів на те, на чому слід зосередити свої зусилля.

Він здатний до класифікації та регресії і може бути застосований до різних завдань маркетингової аналітики, від прогнозного моделювання до сегментації клієнтів.

Використовуючи силу алгоритмів Random Forest, маркетологи можуть отримати цінну інформацію зі своїх даних, що дозволить їм приймати обґрунтовані, засновані на даних рішення, які оптимізують маркетингові стратегії та сприятимуть успіху.

Штучні нейронні мережі, у свою чергу, здатні моделювати складні залежності між змінними та працювати з великими обсягами даних, що дозволяє глибше аналізувати ризикові події у цифровому маркетингу.

Вони, по суті, є обчислювальними системами, змодельованими на основі наших власних біологічних неврологічних систем, зробили концепцію мислячого штучного інтелекту реальністю, або, скоріше, її близьким наближенням.

Говорячи про те, що нейронні мережі імітують наші власні неврологічні процеси, ми маємо на увазі, що штучні нейронні мережі - це системи, що прогресивно навчаються і з часом постійно покращують свою роботу. Мережа складається з сотень, тисяч або навіть мільйонів нейронів, які називаються одиницями, розташованими у трьох взаємопов'язаних шарах:

- Вхідні пристрої, які отримують інформацію та дані із зовнішнього джерела, які мережа повинна обробити або про які вона повинна дізнатися.
- Вихідні блоки, які виробляють відповідь на інформацію, оброблену або вивчену мережею.
- Приховані блоки, які знаходяться між вхідними і вихідними блоками і складають основну частину мережі, яка обробляє або навчається завданням, які вона повинна виконувати.

ANN використовуються в різних галузях - в медицині, інженерії, фінансах та інших. Вони також трансформують наявний набір ресурсів маркетингових

технологій, надаючи маркетологам нові, більш ефективні та динамічні інструменти:

- Прогнозування поведінки споживачів
- Створення та розуміння більш складних сегментів покупців
- Автоматизації маркетингу
- Створення контенту
- Прогнозування продажів

Найбільш широко штучні нейронні мережі застосовуються у сфері аналітики. У цьому випадку нейромережі можуть допомогти маркетологам робити прогнози щодо результату кампанії, розпізнаючи тенденції попередніх маркетингових кампаній. Хоча нейронні мережі існують вже десятки років, саме поява Big Data зробила цю технологію неймовірно корисною для маркетингу.

Завдяки віртуальному набору даних, які можна ввести в нейронну мережу, з'явилася можливість отримати складні, точні прогнози, які можуть допомогти приймати більш розумні рішення щодо того, яких дій слід вжити і на які канали виділити більше ресурсів.

Аналогічно, у сегментації ринку, прогнозуванні продажів, створенні та розповсюдженні контенту нейронні мережі, отримавши достатньо даних, здатні надавати точніші інсайти та прогнози, допомагаючи маркетологам краще оцінювати очікування. Ця технологія також забезпечує більш динамічний рівень автоматизації, що не тільки розвиває маркетинговий робочий процес, але й створює ще більш безперебійний досвід для споживача [22].

Простіше кажучи, прогнозування поведінки - це область предиктивної аналітики, яка намагається передбачити або змодельовати майбутню поведінку людей - зокрема, клієнтів. Прогнозування майбутньої поведінки також можна сформулювати як проблему класифікації.

Наприклад, якщо є чотири можливі варіанти поведінки або дії, які може виконати клієнт, модель прогнозованої поведінки, яка є глибокою нейронною мережею, присвоює кожній з цих дій оцінку ймовірності.

Моделювання предиктивної поведінки - це наука про побудову алгоритмічних моделей та їх навчання на основі історичних даних про клієнтів для прогнозування майбутньої поведінки цих клієнтів. Іншими словами, передбачити ймовірність того, що клієнт зробить ту чи іншу дію.

Замість того, щоб намагатися робити обґрунтовані припущення на основі аналізу історичних даних, як це зазвичай робить людина, предиктивне моделювання поведінки дозволяє компаніям приймати рішення на основі майбутніх прогнозів, зроблених алгоритмічними моделями, що і є моделями глибоких нейронних мереж.

Одне з важливих застосувань предиктивного моделювання поведінки - це прогнозування відтоку клієнтів. Відтік клієнтів відбувається, коли клієнти припиняють вести бізнес з компанією або послугою і йдуть.

Маркетингова нейромережа має багато переваг. На відміну від традиційних моделей, її можна використовувати для аналізу даних і прогнозування. Ця технологія є проривом у маркетингу. З її допомогою маркетологи можуть таргетувати аудиторію, розуміти тенденції та вимірювати ефективність в режимі реального часу. Це чудовий спосіб приймати більш обґрунтовані рішення та збільшувати прибутки. Однак це не єдина перевага маркетингової нейромережі. Вона може застосовуватися у сфері бізнесу та фінансів і є дуже корисною для оцінки тенденцій та прогнозів.

Таким чином, сучасні моделі, такі як Random Forest і ANN, та метод Монте-Карло мають переваги над традиційними методами через здатність працювати зі складними, багатофакторними та великими даними, надаючи точніші та надійніші результати в умовах цифрового маркетингу. У наступному підрозділі ми розглянемо ці моделі більш детально, описавши практичне використання.

1.3. Формулювання гіпотез дослідження.

Після ознайомлення з ризиками та вибору методів їх оцінки, наступним кроком у процесі дослідження є формулювання гіпотези. Гіпотеза - це твердження, яке можна перевірити за допомогою статистичних методів. Це

попереднє пояснення явища або зв'язку між змінними. Гіпотези можуть бути сформульовані по-різному, залежно від питання дослідження та типу даних, що аналізуються.

Для тестування обраних методу та моделей ми використовували дані компаній email-маркетингу, представлених банком. Як було описано раніше, Random Forest і штучні нейронні мережі (ANN), можуть глибше аналізувати дані та враховувати складні залежності між змінними. Метод Монте-Карло, у свою чергу, забезпечує можливість симуляції великої кількості сценаріїв і врахування невизначеності, що дозволяє аналізувати крайні випадки та гнучко адаптуватися до непередбачуваних ризиків.

Спираючись на перспективи використання обраних моделей, ми сформулювали три гіпотези. Вони мають на меті перевірити ефективність комбінування запропонованих моделей та методу. Підтвердження гіпотез доведе припущення про те, що об'єднання покращить точність прогнозування ризиків.

Перша гіпотеза полягає в тому, що використання моделей Random Forest та ANN дозволяє точніше передбачити ризик потрапляння email-кампаній до спам-фільтрів, що може бути пов'язано з технологічними змінами у сфері. Наприклад, можна виявити, які характеристики листів (надмірна кількість персоналізованих даних, формат повідомлень) сприяють потраплянню до спам-фільтрів через порушення політик конфіденційності. Також прогнозувати, як зміна конфіденційних налаштувань користувачів або нові законодавчі вимоги вплинуть на ймовірність блокування листів поштовими сервісами. Варто зазначити, що користувачі можуть сприймати листи як непотрібні або навіть шкідливі, що призведе до зниження довіри до бренду.

Спам - це термін, який використовується для опису небажаної електронної пошти, тобто повідомлень, надісланих одержувачам без їхнього дозволу [2]. Спам може бути надісланий через платформу електронного маркетингу або вручну людиною. Спам не лише дратує, але й може бути небезпечним. Він часто містить віруси або інші шкідливі програми, які можуть скомпрометувати пристрій користувача або його особисту інформацію.

Коли email-маркетингові листи потрапляють до папки «Спам», це означає знищення всіх зусиль у сфері email-маркетингу, оскільки такі листи навряд чи будуть коли-небудь відкриті та прочитані. Згідно з дослідженням Return Path, середній показник прочитання листів, які потрапляють до папки «Спам», становить лише 1,5%, тоді як середній показник прочитання листів, доставлених до поштової скриньки, становить 14,5% [23]. Це означає, що коли листи потрапляють до спаму, ваші показники відкриття, кліків та конверсії стають значно нижчими.

Успіх email-кампанії значною мірою залежить від технології фільтрів. Фільтри не лише блокують вхідні повідомлення, але й упорядковують їх. Сьогодні багато провайдерів послуг електронної пошти та поштових клієнтів розбивають повідомлення на соціальні, комерційні, інформаційні розсилки та інші категорії. Вони використовують певні критерії, щоб оцінити вхідне повідомлення і помістити його у відповідну папку.

Спам-фільтри працюють так само і присвоюють повідомленням оцінку спаму. Якщо оцінка відповідає певному пороговому значенню, лист потрапляє до папки «Вхідні», в іншому випадку – до папки «Спам». Існує довгий список критеріїв спаму, який щодня змінюється та адаптується. Практики фільтрації зазвичай не розголошуються з міркувань безпеки, але ми все ж маємо загальне уявлення про те, що їм подобається і що не подобається.

Random Forest, як ансамблевий метод машинного навчання, здатен аналізувати численні фактори (наприклад, структуру листа, частоту відправлень, поведінкові дані користувачів), що впливають на ймовірність потрапляння до спаму. ANN, завдяки своїй здатності виявляти складні нелінійні залежності, може додатково підвищити точність прогнозів, враховуючи взаємодію численних змінних. Гіпотеза полягає в тому, що комбінація цих методів дозволить більш точно прогнозувати ймовірність блокування листів спам-фільтрами.

Друга гіпотеза має на меті перевірку застосування комбінованого підходу на основі ANN та Random Forest, що дозволить точніше прогнозувати найбільш оптимальний час для відправки email-кампаній, що в свою чергу призведе до

підвищення, Open Rate та CTR. Це дозволить оцінити етичний ризик, наприклад, надмірної частоти розсилок або нерелевантного контенту.

Сьогодні в email-маркетингу час відправлення листа має великий вплив на залученість користувачів. Надсилання в оптимальний час може допомогти зробити кампанії більш успішними та ефективними. Щоб відправляти імейли в найкращий час, потрібно добре розуміти, як користувачі взаємодіють з електронною поштою. Відправлення імейлів безпосередньо перед тим, як користувач взаємодіє зі своєю поштовою скринькою, може підвищити ймовірність залучення.

Оскільки електронна пошта є економічно ефективним засобом, більшість маркетингових команд розсилають кампанії масово, незалежно від уподобань одержувача. Це призводить до зниження рівня залученості, а іноді й до того, що клієнти відмовляються від підписки на розсилки. Отже, персоналізація, яка розуміє вподобання отримувачів, є ключовим фактором для розробки успішних кампаній. Одним із видів персоналізації є визначення найкращого часу для надсилання електронного листа, де метою є визначення найкращого часу для доставки листа, коли ймовірність того, що його побачать і відкриють, є високою. Більшість існуючих рішень значною мірою покладаються на велику кількість інформації про поведінку одержувача, що призводить до холодного старту в найбільш вірогідних сценаріях. Крім того, статичність цих рішень призводить до того, що з часом прогнози стають неактуальними. Аналіз другої гіпотези ґрунтується на дослідженнях, які припускають, що численні атрибути на рівні відправників і одержувачів призводять до відмінностей у відкритій поведінці

Такі фактори, як часові пояси, робочий день чи ні, а також демографічні дані та вподобання вашої цільової аудиторії - все це впливає на показник відкриттів розсилки.

Зрештою, спроба врахувати особистий графік кожного контакту, робочі години, обідню перерву і все, що між ними, може виявитися непосильним завданням.

ANN можуть використовуватися для аналізу багатовимірних поведінкових даних, щоб виявити приховані нелінійні залежності між часом відправлення та ефективністю взаємодії з листами (Open Rate, CTR). Random Forest, у свою чергу, здатний відслідковувати, які саме фактори (наприклад, час доби, день тижня, зміст листа) мають найбільший вплив на успішність розсилки.

Поєднання цих підходів дозволяє автоматично навчатися на історичних даних та постійно оновлювати прогнозовані "вікна ефективності" для відправлення листів, враховуючи зміни в поведінці користувачів або ринку.

Не існує ідеального часу доби та дня тижня для надсилання маркетингових листів; він залежить від низки факторів: від сфери діяльності компанії, від географії та особистості клієнта. Підвищенню ефективності сприятимуть експерименти з розкладом у різний час дня і в різні дні тижня. У результаті можна отримати порівняльні дані, які дадуть змогу дізнатися, коли підписники з більшою чи меншою ймовірністю будуть взаємодіяти з електронними листами. Час, витрачений на перегляд контенту електронної пошти. Під час здійснення поштових розсилок у маркетингу тривалість взаємодії з контентом електронної пошти не менш важлива, і чим довше передплатник читає вашу електронну пошту, тим краще. Така інформація дасть змогу адаптувати зміст листів так, щоб гарантувати максимальну залученість.

Третя гіпотеза ґрунтується на поєднанні Random Forest, ANN та методу Монте-Карло.

Якщо користувачі підозрюють, що листи є потенційно небезпечними (фішинг або шахрайські дії), вони можуть не лише не взаємодіяти з ними, але й відписатися від розсилки, щоб уникнути майбутніх ризиків, що є досить ймовірним оскільки діяльності банку підпадає під найбільш привабливу для шахраїв.

Аналіз вірогідності відписки має дуже велике значення для маркетингової компанії. Бізнес зобов'язаний включати опцію відписки в кожен надісланий лист. Наприклад GDPR (Загальний регламент про захист даних) містить керівні принципи для європейських email-маркетологів. Хоча він не стосується

безпосередньо законів про відписки, він передбачає, що люди мають право відкликати згоду на обробку їхніх персональних даних, зокрема на отримання маркетингових електронних листів. Маркетологи повинні надавати користувачам чіткі та прості способи відписатися від їхніх розсилок.

GDPR наголошує на розширенні прав і можливостей людей щодо їхніх даних, гарантуючи, що вони можуть обирати, як використовувати їхню інформацію та коли припинити її використання. Згідно з GDPR, процес відписки від списків електронної пошти або маркетингових комунікацій повинен бути простим і зручним для користувачів [5].

Організації зобов'язані включати чіткі, легкодоступні варіанти відмови від розсилки в кожне повідомлення, що дозволяє людям відкликати свою згоду з такою ж легкістю, з якою вони її надавали. Ця вимога ґрунтується на принципі згоди, який передбачає, що згода має надаватися вільно, бути конкретною, поінформованою та недвозначною.

Відписка зазвичай є результатом сукупності кількох факторів, таких як частота листів, релевантність контенту, час відправлення, а також поведінка користувачів та технічні аспекти (потрапляння листів у спам). Моделювання цих взаємозалежних факторів може допомогти передбачити, коли й чому користувачі вирішують відписатися, і таким чином, дозволить маркетологам розробляти стратегії для зменшення цього ризику.

У першому розділі ми проаналізували ризики, які можуть виникати в цифровому маркетингу. Ризик є багатограними, пожинаючи конфіденційністю даних та завершуючи етичними викликами. Це дозволяє усвідомити важливість комплексного підходу до управління ризиками.

Ми переконані, що для дослідження ризиків в сфері цифрового маркетингу необхідно використовувати сучасні методи. Саме тому ми обрали Random Forest, штучні нейронні мережі та метод Монте-Карло, оцінивши їх переваги над традиційними підходами, такими як стохастичне програмування чи дерева рішень.

Сформовані гіпотези мають вагоме практичне значення. Перевірка гіпотез сприятиме розробці стратегій, що в свою чергу підвищить ефективність маркетингових компаній та мінімізує ризики. У сучасному середовищі, де цифрові технології є не лише засобом комунікації, а й основою бізнес-стратегій, здатність компаній ефективно оцінювати та управляти ризиками стає ключовим фактором їхньої стійкості та конкурентоспроможності.

РОЗДІЛ 2.

МОДЕЛЮВАННЯ ТА ОЦІНКА РИЗИКІВ МАРКЕТИНГУ ЗА ДОПОМОГОЮ МАТЕМАТИЧНИХ МОДЕЛЕЙ ТА МЕТОДІВ

2.1. Побудова моделі «Random Forest» для оцінки ризиків у цифровому маркетингу.

«RandomForest» - це алгоритм керованого навчання, який використовується як для класифікації, так і для регресії і може допомогти компаніям у прийнятті стратегічних рішень. Алгоритм випадкового лісу створює дерева рішень на вибірках даних, а потім отримує прогноз від кожного з них і, нарешті, обирає найкраще рішення за допомогою голосування.

Метод ансамблю випадкового лісу складається з великої кількості невеликих дерев рішень, які називаються оцінками, кожна з яких дає власні прогнози. Модель випадкового лісу об'єднує різні класифікатори в один метакласифікатор для забезпечення кращого узагальнення і точності порівняно з окремими класифікаторами.

Ми створили бутстрап даних із заміною розміру n (один і той самий зразок можна намалювати кілька разів). Далі побудували дерево рішень на основі даних бутстрапу.

Наступним, було повторення кроків 1-2 k разів (ця різноманітність багатьох спроб розширює можливості випадкового лісу порівняно з окремими деревами рішень)

Об'єднання в мішок або бутстрап означає, що ми робимо остаточний прогноз більшістю голосів - вибираємо мітку класу, яка була передбачена більшістю дерев. Однак, замість того, щоб використовувати ті самі навчальні дані, ми використовуємо бутстрап-вибірки. Реалізацію моделі «Random Forest» ми провели в програмі «Python».

Завдяки тому, що «Random Forest» є ансамблевим методом, нам не потрібно обрізати дерева завдяки стійкості до шуму (остаточне рішення приймається на основі окремих дерев).

Для кожної з трьох гіпотез ми побудували модель на базі даних представлених банком відповідно до email-кампанії. Дані для дослідження були представлені ц форматі ексель документу, який містив 30 000 рядків та 8 стовпців. Загальний обсяг інформації складав 240 000 записів (додаток А).

При створенні моделі Random Forest для кожної з гіпотез у роботі ми використовували принципи побудови ансамблевих моделей, які поєднують кілька окремих дерев рішень для покращення якості прогнозування та зменшення проблеми надмірного навчання.

Для кожної гіпотези ми обирали відповідні змінні (параметри), які найкраще відповідали нашій меті. У випадку з передбаченням часу відкриття, ми використовували такі змінні, як Journey Start Time, Client Services Segment, Country тощо. Для ризиків відписок ми обрали Journey Topic, Client Services Segment та Country.

Після цього ми застосували кодування категоріальних змінних (за допомогою LabelEncoder), щоб модель могла використовувати ці значення як числові ознаки.

Для даної моделі ми використали параметри, які були запрограмовані наступним чином:

- `n_estimators`: кількість дерев у лісі дорівнює 100. Вибір великої кількості дерев допомагає збалансувати міжваріативність, однак збільшує обчислювальні витрати.
- `max_depth`: глибина кожного дерева - 15. Ми обмежували глибину дерев (наприклад, 10-15 рівнів), щоб запобігти надмірному навчанні.
- `class_weight`: врахування дисбалансу класів (наприклад, для прогнозування відписок). Ми використовували `balanced`, щоб модель могла коректно обробляти різні частоти класів у даних.
- `max_samples`: розмір вибірки даних у бутстрапі – 80%

На кожному вузлі дерева ми оцінюємо, яку ознаку використовувати для поділу даних [8].

Формула для критерію Джині (Gini Index):

$$Gini = 1 - \sum_{i=1}^n p_i^2, (2.1)$$

де p_i — частка вибірок, що належать до класу i . Ознака з мінімальним значенням Джині вибирається для поділу вузла.

Кожне дерево у моделі формується на основі різних підвибірок ознак та бутстрап-вибірок даних.

У вузлах дерева на кожному етапі розглядається випадкова підмножина ознак (зазвичай $\sqrt{d}$, де d — загальна кількість ознак). Це запобігає надмірному навчанні та зменшує кореляцію між окремими деревами.

Кожне дерево розширюється до тих пір, поки всі листки не будуть чистими або не досягнуть мінімальної кількості зразків.

Всі дерева в лісі формують окремі прогнози, а потім поєднують їх за принципом більшості голосів (у випадку класифікації). Тобто остаточний прогноз для кожної вибірки визначається тією міткою класу, яка була передбачена більшістю дерев.

Після навчання ми оцінили модель за допомогою звітів класифікації (Classification Report), де враховувалася точність (Precision), повнота (Recall) та середнє значення f1-міри. Ми також врахували загальну точність (accuracy) для оцінки ефективності моделі. Після побудови кожної моделі ми оцінили важливість ознак та побудували матрицю невідповідностей (Confusion Matrix).

Формула для визначення важливості функції виглядає так [16]:

$$\text{Feature Importance} = \frac{1}{T} \sum_{t=1}^T \Delta I_f^{(t)}, (2.2)$$

де:

T — загальна кількість дерев.

ΔI — зміна критеріїв поділу на дереві t , яка сталася завдяки ознаці f .

Далі ми перевірили як модель працює самостійно на базі представлених даних. Модель ми оцінили завдяки Classification Report.

Для першої гіпотези ми використали дані, які відображають взаємодію з листами, тип клієнтів, географічні фактори та тематику кампанії, що впливають на ймовірність потрапляння до спам-фільтрів, а саме :

- New_interactiontypename (Email Opened, Email Clicked, Email Delivered, Email Sent, Email address Opted-Out)
- Client Services Segment (Private Clients & Family Offices)
- Country
- Journey Topic

Модель показала наступні результати представлені в таблиці та графіку.

Таблиця 2.1

Classification Report для моделі Рендом форест для першої гіпотези.

Клас	precision	recall	f1-score	support
0	0.12	0.06	0.08	65
1	0.47	0.37	0.42	2706
2	0.15	0.01	0.02	1134
3	0.38	0.68	0.49	2089
4	0.00	0.00	0.00	6
accuracy	0.41			
macro avg	0.22	0.23	0.20	6000
weighted avg	0.38	0.41	0.36	6000

*Джерело: розраховано автором

Використання моделі Random Forest показало середній рівень загальної точності, що склав 41%. Попри те, що це значення може здатися низьким, важливо враховувати специфіку задачі, де кожен клас взаємодії має свої унікальні особливості. Зокрема, клас "Email Sent" був класифікований із відносно високою повнотою (68%), що означає, що модель виявила більшість випадків цього класу.

Однак точність для інших класів була нижчою, що свідчить про труднощі в розрізненні подібних типів взаємодії.

Класифікаційний звіт показує, що модель добре розпізнає "Email Delivered," з дещо вищою точністю порівняно з іншими класами, проте все ще існує значна кількість помилок класифікації, коли "Email Delivered" плутають із "Email Sent." Це може свідчити про те, що між цими двома класами є схожі ознаки, які модель не здатна належним чином диференціювати. З іншого боку, класи, такі як "Email Clicked" та "Email Opened," мають ще нижчу точність, що свідчить про недостатню кількість або невідповідність ознак для їх класифікації.

Факт низької точності та F1-score для більшості класів може також бути обумовлений нерівномірністю розподілу класів у вибірці, що є типовою проблемою для завдань класифікації в маркетингових даних. Наявність значного дисбалансу між кількістю випадків кожного класу може призвести до того, що модель навчиться розпізнавати більш поширені класи, ігноруючи рідкісні.

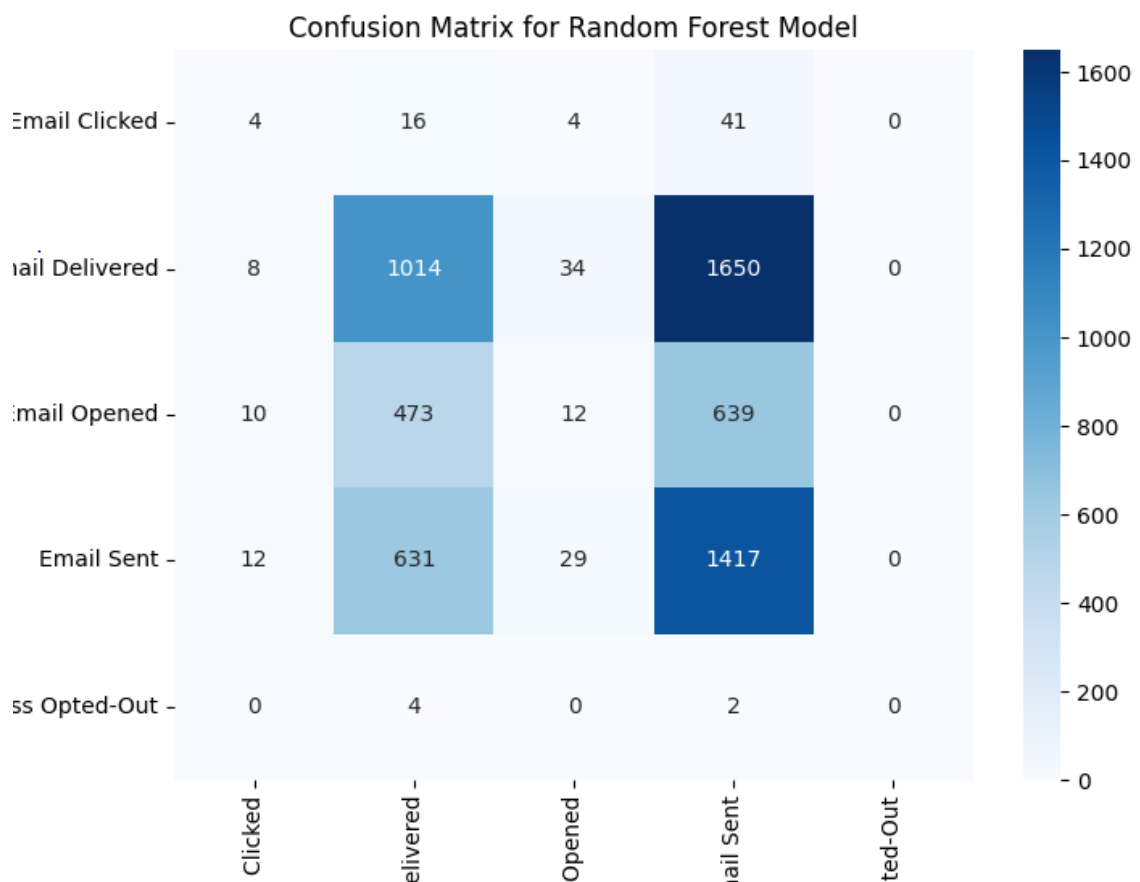


Рис. 2.1. Confusion Matrix для моделі Рендом форест для першої гіпотези.

Матриця невідповідностей показує, що модель часто плутає різні класи взаємодії. Наприклад, значна кількість прикладів класу "Email Delivered" помилково класифікується як "Email Sent." Це є ознакою того, що модель не має чіткого розмежування між цими класами, що вказує на потребу вдосконалення ознак або вибору додаткових змінних.

Для другої гіпотези — оптимальний час для відправки ми використали такі змінні:

- Createdon : час запуску розсилки — основний фактор, який визначає успішність кампанії.
- Client Services Segment: сегментація клієнтів дозволяє визначити, для яких груп найкраще працюють різні часи відправки.
- Country: для врахування часових зон.

Таблиця 2.2

Classification Report для моделі Рендом фореест для другої гіпотези.

Клас	precision	recall	f1-score	support
0	0.21	0.14	0.17	65
1	0.51	0.42	0.46	2706
2	0.33	0.00	0.00	1134
3	0.40	0.71	0.51	2089
4	0.00	0.00	0.00	6
accuracy	0.44			6000
macro avg	0.29	0.25	0.23	6000
weighted avg	0.43	0.44	0.39	6000

*Джерело: розраховано автором

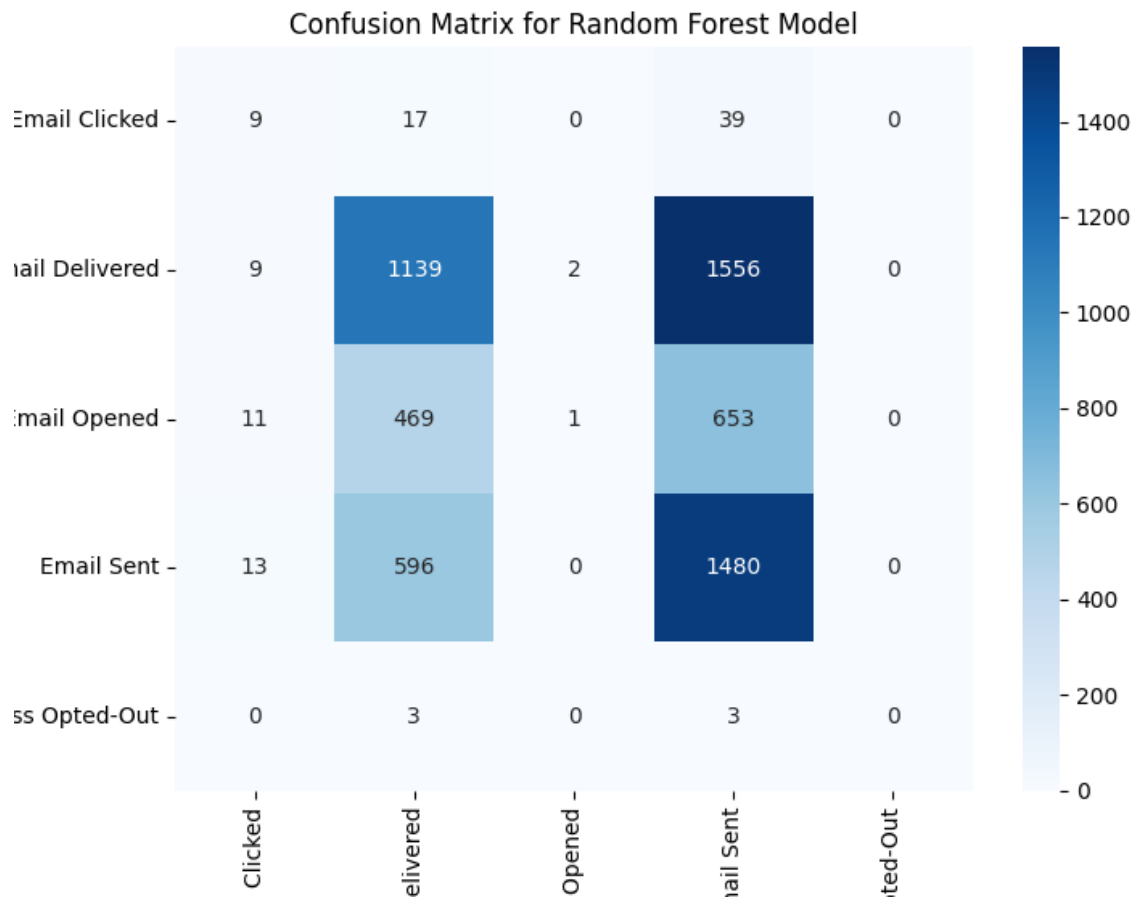


Рис. 2.2. Confusion Matrix для моделі Рендом форест для другої гіпотези.

За результатами можна зробити висновок про обмежену точність моделі на основі обраних ознак.

Згідно з таблицею метрик, показники точності (Precision), повноти (Recall) та F1-міри вказують на те, що модель має змішані результати:

Для класу "Email Sent" модель досягла кращої точності, але для менш представлених класів, таких як "Email Clicked" і "Opted-Out", точність залишається низькою.

Загальна точність моделі становить лише 44%, що свідчить про те, що модель не може повністю коректно класифікувати всі вибірки.

Час відправки (createdon) залишається головним фактором, який впливає на результат, оскільки вибір оптимального часу є важливим для підвищення Open Rate та CTR (Click-Through Rate). Однак географічне розташування (Country) та сегментація клієнтів (Client Services Segment) також відіграють роль у

класифікації, враховуючи відправку email-кампаній у різні часові зони та на різні групи клієнтів.

Модель може бути корисною для даної цілі. Незважаючи на це, вона потребуватиме додаткової оптимізації для забезпечення вищої точності для різних типів взаємодій.

Для третьої гіпотези — прогнозування ризиків відписок ми використали показники:

- new_interactiontypename (Email Opened, Clicked,)
- Client Services Segment (Private Clients & Family Offices)
- Country
- Journey Topic

Ми перетворили new_interactiontypename на бінарну змінну interaction_binary, де значення 1 означає позитивну взаємодію ("Email Clicked" або "Email Opened"), а 0 — негативну взаємодію.

Таблиця 2.3

Classification Report для моделі Рендом форест для третьої гіпотези.

Клас	precision	recall	f1-score	support
0	0.80	0.99	0.89	4801
1	0.29	0.02	0.03	1199
accuracy	0.80			6000
macro avg	0.54	0.50	0.46	6000
weighted avg	0.70	0.80	0.71	6000

*Джерело: розраховано автором

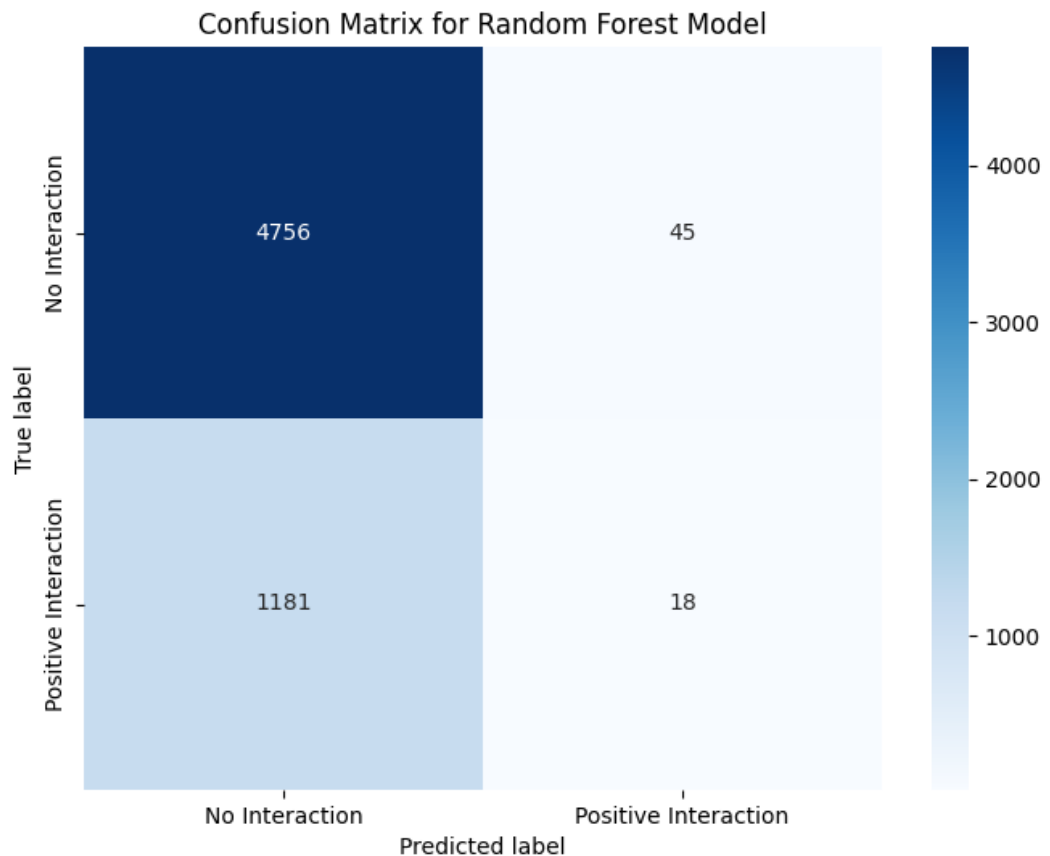


Рис. 2.3. Confusion Matrix для моделі Рендом форест для третьої гіпотези.

На перший погляд результат показників здається доволі високим – 80%, але варто зауважити щодо точності в передбаченні перемінних. Модель добре працює у передбаченні відсутності взаємодії (No Interaction): точність для цього класу становить 80%, а повнота — 99%. Це означає, що модель майже завжди правильно визначає зразки, в яких відсутня взаємодія з листами. Отже, для маркетингових команд ця модель може бути корисною в ідентифікації тих отримувачів, які, швидше за все, не взаємодіють з email-кампаніями.

Незважаючи на хороші результати для класу "No Interaction", модель суттєво відстає в передбаченні позитивної взаємодії (Positive Interaction), що є критично важливим показником для оцінки ризиків відписок. Точність для класу позитивної взаємодії становить лише 29%, а повнота — 2%. Це означає, що модель практично не здатна правильно передбачати позитивні взаємодії, тобто випадки, коли отримувачі відкривають або натискають на листи. Такий низький

рівень передбачуваності позитивної взаємодії є проблемним, оскільки саме ці взаємодії дозволяють відслідковувати зацікавленість клієнтів та їхнє бажання залишатися підписаними на розсилку.

До цього може призвести дисбаланс між класами. Саме це впливає на продуктивність моделі, адже у нашому випадку значна частина зразків належить до класу "No Interaction", що змушує модель «навчатися» на цьому класі ігнорувати рідкісніші позитивні взаємодії. Результат - модель схильна класифікувати більшість зразків як "No Interaction", що пояснює низькі показники для класу позитивної взаємодії. Це призведе до того, що модель стане нездатною виявити потенційні ризики відписок, оскільки ризик відписки зазвичай тісно пов'язаний з відсутністю регулярної взаємодії.

Результати використання моделі Random Forest для оцінки трьох гіпотез вказують на потенціал методу у виявленні деяких типів взаємодії з email-кампаніями, однак існує ряд викликів, які потребують подальшого вдосконалення. Зокрема, модель продемонструвала середню загальну точність, проте її продуктивність варіювалася залежно від класу.

У першій гіпотезі модель показала обмежену здатність точно класифікувати різні типи взаємодії, що обумовлено схожістю ознак між класами та дисбалансом даних. Класи "Email Sent" і "Email Delivered" були класифіковані краще, але з великим ризиком помилкової класифікації між ними, що вказує на потребу в оптимізації ознак або додаткових змінних.

Друга гіпотеза, яка стосується визначення оптимального часу для відправки з метою підвищення Open Rate та CTR, продемонструвала, що час відправки (createdon) є ключовим фактором. Однак модель мала низьку точність для менш представлених класів, таких як "Email Clicked" та "Opted-Out". Це свідчить про те, що модель потребує додаткової оптимізації та балансування класів, щоб забезпечити точніші прогнози.

У третій гіпотезі ми зосередилися на прогнозуванні ризиків відписок, використовуючи бінарну змінну для взаємодії. Хоча модель добре ідентифікувала відсутність взаємодії, вона майже не могла розпізнати позитивну взаємодію, що

є критично важливим для оцінки ризиків відписок. Дисбаланс класів призвів до того, що модель навчилася розпізнавати більш поширений клас "No Interaction", але практично ігнорувала рідкісні позитивні взаємодії.

Хоча модель Random Forest забезпечила певну ефективність у класифікації, її результати підкреслюють необхідність додаткових методів для досягнення вищих показників точності та повноти, особливо для задачі прогнозування ризиків відписок у email-кампаніях.

2.2. Моделювання ризиків за допомогою штучних нейронних мереж (ANN).

Штучна нейронна мережа (ANN) - це парадигма обробки інформації, яка натхненна мозком. ANN, як і люди, навчаються на прикладах. ANN налаштовується для конкретного застосування, наприклад, розпізнавання образів або класифікації даних, за допомогою процесу навчання. Навчання здебільшого передбачає коригування синаптичних зв'язків, які існують між нейронами. Мозок складається з сотень мільярдів клітин, які називаються нейронами. Ці нейрони з'єднані між собою синапсами, які є нічим іншим, як з'єднаннями, через які нейрон може посылати імпульс іншому нейрону. Коли нейрон посилає збудливий сигнал іншому нейрону, цей сигнал додається до всіх інших входів цього нейрона. Якщо він перевищить певний поріг, то змусить цільовий нейрон послати сигнал дії вперед - так відбувається внутрішній процес мислення [3].

У комп'ютерних науках ми моделюємо цей процес, створюючи «мережі» на комп'ютері за допомогою матриць. Ці мережі можна розуміти як абстракцію нейронів без урахування всіх біологічних складнощів. Для простоти ми моделюватимемо просту NN з двома шарами, здатну вирішувати лінійну задачу класифікації.

Для створення ANN-моделі ми використали MLP-класифікатор scikit-learn. Ми застосували масштабування ознак до даних за допомогою StandardScaler, щоб покращити продуктивність моделі. Параметр `hidden_layer_sizes` визначає

архітектуру ANN з двома прихованими шарами, кожен з яких містить 100 та 50 нейронів відповідно.

Перед створенням моделі ANN, важливо правильно підготувати дані. Весь процес включає наступні кроки:

1. Обробка категоріальних змінних: Використовується `LabelEncoder` для перетворення категоріальних змінних у числові значення. Цей процес дозволяє алгоритму обробляти текстові дані у вигляді чисел, які можна використовувати в матрицях.

Формула для цього перетворення:

$$x_{new} = \text{LabelEncoder}(x), \quad (2.3)$$

Тут x — це початкова категоріальна змінна, а $x_{\{new\}}$ — числове значення після кодування.

2. Масштабування: Після перетворення даних, всі числові змінні нормалізуються за допомогою `StandardScaler`. Цей процес дозволяє отримати дані з середнім значенням 0 і стандартним відхиленням 1:

$$x_{scaled} = \frac{x - \mu}{\sigma}, \quad (2.4)$$

Де:

- x — початкове значення ознаки.
- μ — середнє значення цієї ознаки.
- σ — стандартне відхилення цієї ознаки.

Штучна нейронна мережа (Artificial Neural Network, ANN) імітує роботу біологічних нейронів. Основні елементи ANN — це вхідний шар, приховані шари та вихідний шар.

Вхідний шар: Приймає дані, що подаються на вхід (наприклад, чотири змінні). Кожен вузол (нейрон) з'єднаний з наступним шаром ваговими коефіцієнтами w .

$$Z = X \cdot W + b, \quad (2.5)$$

де:

- X — вхідні дані.

- W — вагові коефіцієнти.
- b — зміщення (bias).

Приховані шари: У нашій моделі два приховані шари. Кожен вузол у прихованому шарі обчислює лінійну комбінацію вхідних даних та застосовує нелінійну активаційну функцію (наприклад, relu) [12]:

$$A = \text{ReLU}(Z), \quad (2.6)$$

де:

- Z — лінійна комбінація входів.
- Rectified Linear Unit — активаційна функція, яка дорівнює 0, якщо $Z \leq 0$, і Z , якщо $Z > 0$.

Вихідний шар: Використовує активаційну функцію sigmoid для передбачення ймовірності. Для бінарної класифікації sigmoid нормалізує вихід до значень між 0 та 1, де Z — лінійна комбінація виходів з попереднього шару.

Функція втрат: Використовується $\text{binary_crossentropy}$ для оцінки помилки між передбаченням та реальним результатом [3]:

$$\text{Loss} = -\frac{N}{1} \sum_{i=1}^N [y_i \log(\hat{y}_i) + (1 - y_i) \log(1 - \hat{y}_i)], \quad (2.7)$$

де:

- N — кількість зразків.
- y_i — реальні значення.
- $\hat{y}_i$ — передбачені значення.

Процес навчання моделі виконується за допомогою алгоритму зворотного поширення похибки (Backpropagation):

1. Обчислення похибки: Обчислюється різниця між передбаченим виходом та реальним значенням (помилка).
2. Оновлення вагових коефіцієнтів: Вагові коефіцієнти коригуються у напрямку зменшення похибки:

$$W = W - \eta \frac{\partial \text{Loss}}{\partial W}, \quad (2.8)$$

де:

- W — вагові коефіцієнти.
- η — коефіцієнт навчання.
- $\frac{\partial Loss}{\partial W}$ — градієнт функції втрат за вагою.

Реалізація моделі в Python:

1. Створення ANN моделі:

- Використовуємо Sequential() для ініціалізації моделі.
- Додаємо шари за допомогою Dense() з вибором активаційних функцій.

2. Навчання:

- Навчаємо модель на X_{train_scaled} та y_{train} за кілька епох.
- Валідація на тестовій вибірці X_{test_scaled} та y_{test} .

3. Оцінка:

- Використовуємо classification_report та confusion_matrix для оцінки точності та ефективності.

Розглянемо результати для першої гіпотези.

Таблиця 2.4

Classification Report для моделі ANN для першої гіпотези.

Клас	precision	recall	f1-score	support
0	0.00	0.00	0.00	65
1	0.72	0.24	0.36	2706
2	0.00	0.00	0.00	1134
3	0.41	0.99	0.58	2089
4	0.00	0.00	0.00	6
accuracy	0.45			
macro avg	0.23	0.25	0.19	6000
weighted avg	0.47	0.45	0.36	6000

*Джерело: розраховано автором

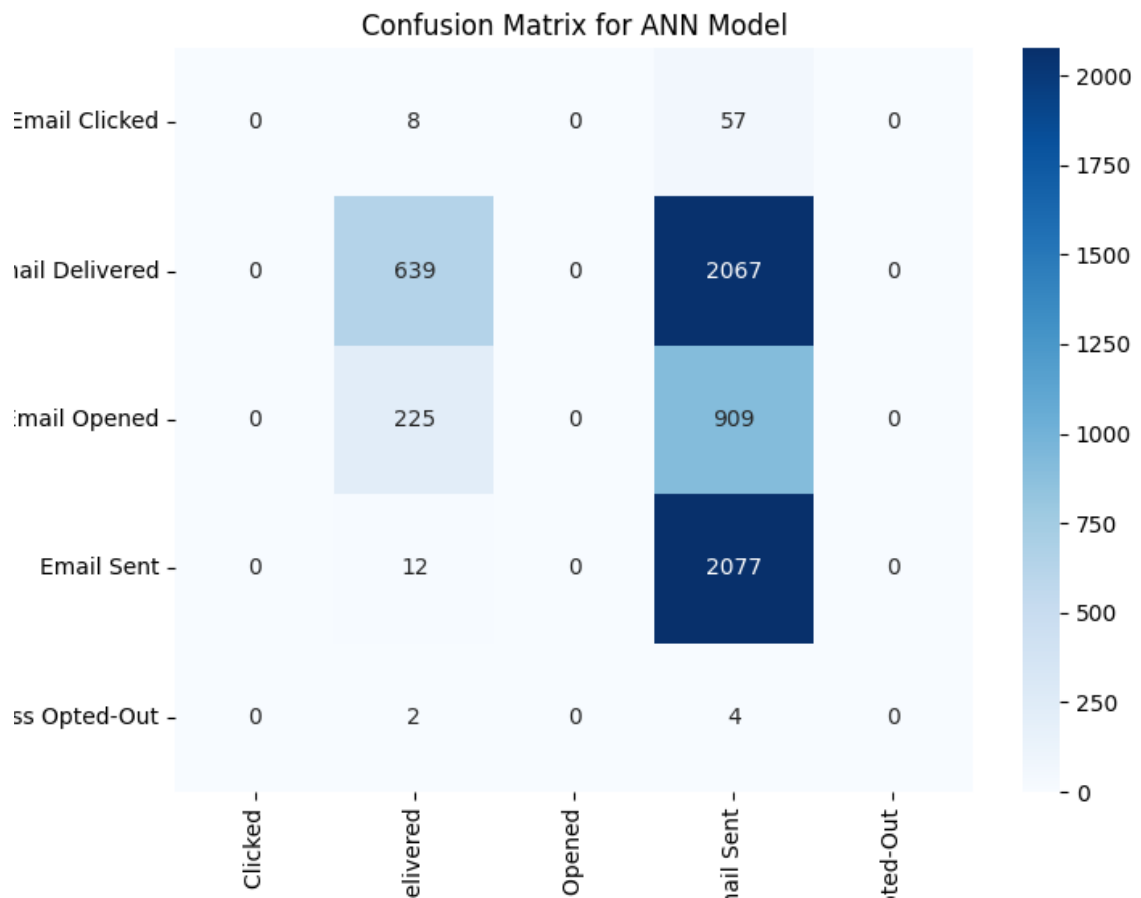


Рис. 2.4. Confusion Matrix для моделі ANN для першої гіпотези.

Результати моделі штучної нейронної мережі (ANN) для прогнозування типів взаємодії з email-кампаніями свідчать про те, що модель має певні обмеження у здатності точно класифікувати різні категорії. Модель показала загальну точність у 45%, що є відносно низьким показником для цієї задачі.

Модель ANN показала найкращі результати для класу "Email Sent", для якого recall (повнота) досяг 99%, а f1-score — 0.58, вказуючи на те, що модель здатна розпізнавати випадки, коли лист просто надісланий, без подальшої взаємодії з боку отримувача. Однак, для інших класів, таких як "Email Clicked" і "Email Opened", модель показала значно гірші результати. Наприклад, для класу "Email Clicked" (клас 0) модель не змогла правильно класифікувати жодного

зразка, і показник recall становив 0. Так само клас "Opted-Out" (клас 4), що означає відписку, також був повністю проігнорований моделлю.

Одна з головних проблем моделі ANN полягає в дисбалансі між класами. Основна частина даних належить до класу "Email Sent", що може змушувати модель віддавати перевагу цьому класу, ігноруючи менш представлені категорії. Це підтверджується значеннями recall для різних класів. Для класу "Email Sent" модель показала надзвичайно високий показник recall, тоді як для "Email Clicked" та "Opted-Out" вона взагалі не змогла передбачити жодного зразка. Цей дисбаланс суттєво знижує здатність моделі коректно прогнозувати всі категорії та впливає на загальну точність класифікації.

Для покращення результатів можна застосувати методи для балансування класів, такі як підвищення вибірки для менш представлених класів (oversampling) або зменшення вибірки для класу "Email Sent". Це дозволить моделі краще навчатися на кожному класі, що може суттєво підвищити точність.

Ми спостерігаємо доволі обмежену здатність моделі класифікувати різні типи взаємодій з маркетинговою компанією. Точність моделі для класів окрім Email Sent не є достатньою.

Розглянемо результати для другої гіпотези.

Таблиця 2.5

Classification Report для моделі ANN для другої гіпотези.

Клас	precision	recall	f1-score	support
0	0.00	0.00	0.00	65
1	0.55	0.39	0.46	2706
2	0.00	0.00	0.00	1134
3	0.41	0.80	0.54	2089
4	0.00	0.00	0.00	6
accuracy	0.46			6000
macro avg	0.19	0.24	0.20	6000

weighted avg	0.39	0.46	0.40	6000
--------------	------	------	------	------

*Джерело: розраховано автором

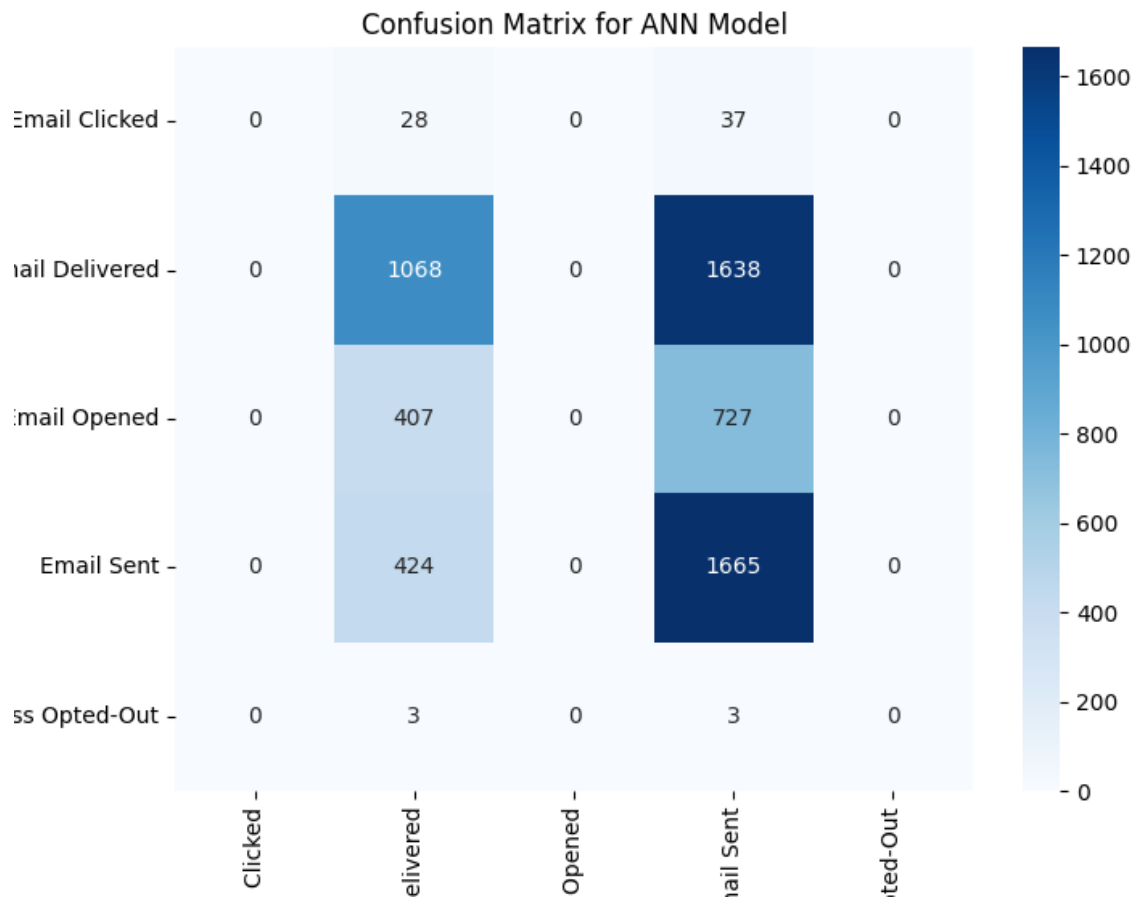


Рис. 2.5. Confusion Matrix для моделі ANN для другої гіпотези.

Результати точно вказують на потребу в вдосконаленні. Незважаючи на те, що модель досягла загальної точності у 46%, її продуктивність значно варіюється між класами.

Модель показала найкращі результати для класу "Email Sent" (клас 3), де recall досяг 80%, що означає, що модель змогла правильно класифікувати більшість випадків надсилання листа. Однак інші класи, такі як "Email Clicked" (клас 0), "Email Opened" (клас 2), та "Opted-Out" (клас 4), мають дуже низькі

значення точності та recall, що свідчить про те, що модель не здатна розпізнавати ці категорії.

Клас "Email Delivered" (клас 1) мав recall на рівні 39%, що також є недостатнім для точного прогнозування. Крім того, показник F1-score, який є середнім значенням між precision та recall, також низький для більшості класів, що вказує на проблеми з класифікацією, особливо для менш представлених класів.

За результатами для третьої гіпотези, ми отримали доволі непогані показники, представлені нижче.

Таблиця 2.6

Classification Report для моделі ANN для третьої гіпотези.

Клас	precision	recall	f1-score	support
0	0.80	1.00	0.89	4801
1	0.00	0.00	0.00	1199
accuracy	0.80			6000
weighted avg	0.64	0.80	0.71	6000

*Джерело: розраховано автором

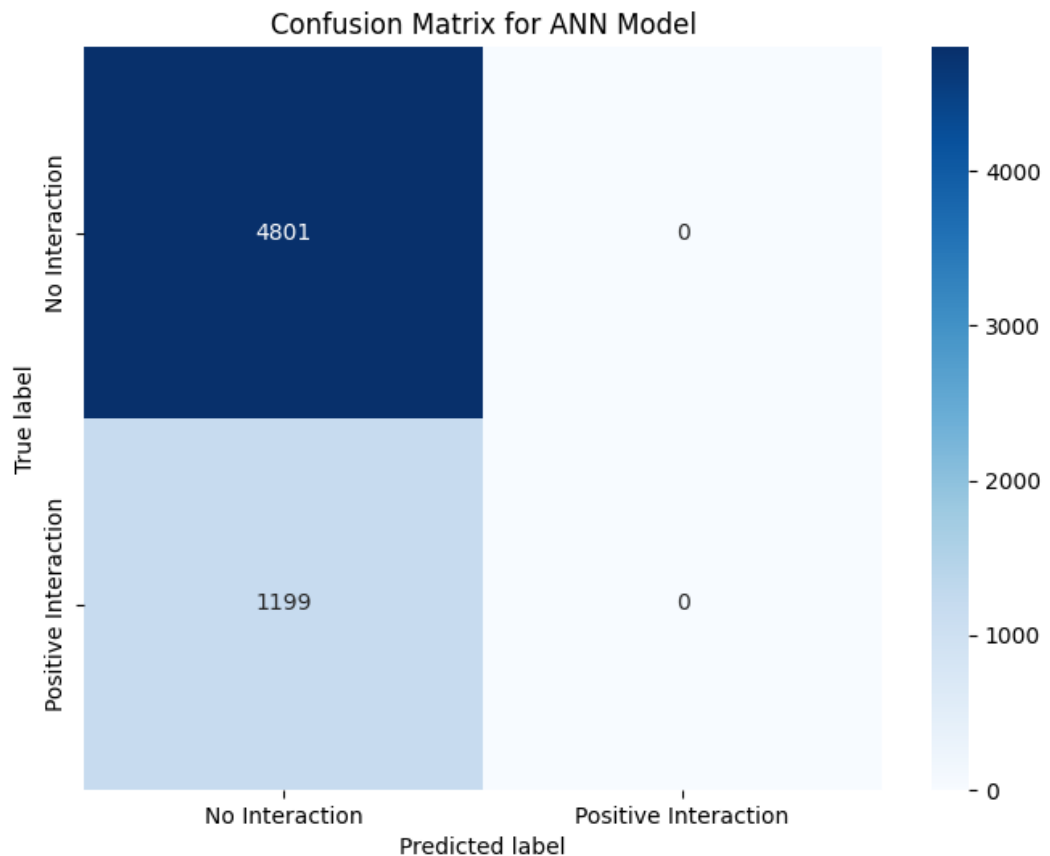


Рис. 2.6. Confusion Matrix для моделі ANN для третьої гіпотези.

Точність для класу "No Interaction": Модель має високий рівень точності для класу 0 (80%), і recall досягає 100%, тобто модель правильно класифікує всі зразки, які належать до цього класу.

Нажаль виникла проблема з класом "Positive Interaction". Для цього класу модель не змогла класифікувати жодного зразка. Recall та precision для цього класу становлять 0%, що свідчить про те, що модель повністю його ігнорує. Це знову ж таки може бути наслідком дисбалансу класів, оскільки модель "навчилася" на переважній більшості зразків класу "No Interaction" і не може розпізнавати рідкісні події.

Велика різниця в кількості прикладів для кожного класу є основною причиною поганої продуктивності для класу "Positive Interaction". Модель значною мірою зосереджується на найбільш представленому класі, ігноруючи менш поширений клас.

Результати показують, що модель ANN потребує вдосконалення для точного прогнозування ризиків відписок. Незважаючи на високу точність для класу "No Interaction", модель повністю ігнорує клас "Positive Interaction", що робить її неефективною для цієї задачі. Балансування класів і комбінування з іншими моделями можуть допомогти поліпшити результати, зробивши модель кориснішою для визначення ймовірності відписок у email-кампаніях.

2.3. Використання методу Монте-Карло для симуляції ризикових подій.

Моделювання методом Монте-Карло - це математичний метод, який прогнозує можливі результати невизначеної події. Комп'ютерні програми використовують цей метод для аналізу даних за минулі періоди та прогнозування низки майбутніх результатів на основі вибору дій. Наприклад, якщо ви хочете оцінити продажі нового продукту за перший місяць, то можете надати програмі моделювання методом Монте-Карло свої історичні дані про продажі. Програма оцінюватиме різні обсяги продажів на основі таких чинників, як загальні ринкові умови, ціна продукту та рекламний бюджет.

Моделювання методом Монте-Карло забезпечує безліч можливих результатів і ймовірність кожного з великого пулу випадкових вибірок даних. Воно дає чіткішу картину, ніж детермінований прогноз. Наприклад, для прогнозування фінансових ризиків потрібен аналіз десятків або сотень факторів ризику. Фінансові аналітики використовують моделювання методом Монте-Карло для визначення ймовірності кожного можливого результату.

Джон фон Нейман і Станіслав Улам винайшли моделювання методом Монте-Карло, або метод Монте-Карло, у 1940-х роках [7]. Вони назвали його на честь знаменитого грального закладу в Монако, тому що цей метод має такі самі характеристики випадковості, що й гра в рулетку.

Компанії використовують методи Монте-Карло для оцінки ризиків і складання точних довгострокових прогнозів. Керівники бізнесу використовують

методи Монте-Карло для прогнозування реалістичних сценаріїв під час ухвалення рішень.

Основний принцип моделювання методом Монте-Карло полягає в ергодичності, що описує статистичну поведінку рухомої точки в замкнутій системі. В кінцевому підсумку рухома точка пройде через усі можливі розташування в ергодичній системі. На цьому ґрунтується моделювання методом Монте-Карло, за якого комп'ютер проводить стільки моделювань, скільки буде достатньо для отримання кінцевого результату за різних вхідних даних.

Моделювання методом Монте-Карло використовує комп'ютерну систему для проведення достатньої кількості моделювань з метою отримання різних результатів, що імітують реальні. Система використовує генератори випадкових чисел для відтворення невизначеності, властивої вхідним параметрам. Генератори випадкових чисел - це комп'ютерні програми, що створюють непередбачувану послідовність випадкових чисел.

Аналіз методом Монте-Карло використовує вхідні змінні, вихідні змінні та математичну модель. Комп'ютерна система вводить незалежні змінні в математичну модель, моделює їх і створює залежні змінні.

Програмне забезпечення для моделювання методом Монте Карло замінює доходи і витрати ймовірними значеннями, залежно від типу розподілу ймовірностей. Потім воно повторює моделювання для отримання високоточного результату. Якщо в математичну модель включено безліч випадкових величин, то моделювання методом Монте-Карло може проводитися кілька годин.

Розподіли ймовірностей - це статистичні функції, що представляють діапазон значень, розподілених між межами. Фахівці зі статистики використовують розподіли ймовірностей для прогнозування можливої появи невизначеної змінної, яка може складатися з дискретних або безперервних значень.

Дискретний розподіл імовірностей представлений цілими числами або послідовністю скінченних чисел. Імовірність появи кожного з дискретних значень більша за нуль. Статистики виводять дискретний розподіл імовірностей

у вигляді таблиці, а безперервний - у вигляді кривої між двома заданими точками на осі x графіка. Нижче наведено поширені типи розподілу ймовірностей, які можна змоделювати методом Монте-Карло.

Нормальний розподіл, також відомий як дзвоноподібна крива, має симетричну форму дзвона і представляє більшість реальних подій. Імовірність випадкового значення висока на медіані і значно зменшується у напрямку до обох кінців дзвоноподібної кривої. Наприклад, повторна випадкова вибірка ваги учнів у певному класі дає нормальну діаграму розподілу.

Рівномірний розподіл означає статистичне представлення випадкових величин з рівною ймовірністю. Під час побудови графіка рівномірно розподілені змінні відображаються у вигляді горизонтальної прямої, що проходить через діапазон допустимих значень. Наприклад, рівномірний розподіл представляє ймовірність кочення і зупинки з кожного боку гральної кістки.

Трикутний розподіл використовує для представлення випадкових величин мінімальні, максимальні та найбільш імовірні значення. Його ймовірність досягає максимуму за найімовірнішого значення. Наприклад, компанії використовують трикутний розподіл для прогнозування майбутніх обсягів продажів, встановлюючи мінімальне, максимальне і пікове значення трикутника.

Під час використання моделювання методом Монте-Карло виникають дві поширені проблеми.

Моделювання методом Монте-Карло дуже залежить від вхідних значень і розподілу. Якщо під час вибору вхідних даних і розподілу ймовірностей допущено помилки, результати можуть виявитися неточними.

Для проведення експериментів з використанням методу Монте-Карло можуть знадобитися надлишкові обчислювальні потужності. Обчислення з використанням методу Монте-Карло на окремому комп'ютері можуть займати години або дні.

У цьому розділі представлено методологію застосування методу Монте-Карло для симуляції ризиків відписки в email-маркетингу. Метод Монте-Карло є важливим інструментом для прогнозування результатів, використовуючи

випадкові сценарії на основі ймовірнісних розподілів змінних. Він дозволяє моделювати різні ризикові події та оцінювати їх ймовірність, що допомагає приймати рішення щодо мінімізації негативних наслідків.

Для симуляції були обрані такі ключові фактори:

- Journey Topic: теми розсилок, які можуть впливати на поведінку користувачів.
- Receipient Type: тип отримувача (лід, поточний клієнт, новий користувач тощо).
- Client Services Segment: сегменти клієнтів, зокрема
- Country: країна, що може впливати на відмінності в культурних і поведінкових аспектах, а також враховувати часові пояси.

Перед симуляцією були проведені такі етапи підготовки даних:

- Очищення даних: було видалено пропущені або неповні записи, що могли впливати на якість моделювання.
- Кодування категоріальних змінних: категоріальні змінні (*Journey Topic*, *Receipient Type*, *Client Services Segment*) були перетворені на числові за допомогою *LabelEncoder*. Це дозволяє використовувати ці змінні в моделях прогнозування.

Для симуляції було застосовано метод Монте-Карло, який генерує 1000 випадкових сценаріїв на основі розподілу ймовірностей для кожної змінної:

Кожен сценарій симуляції використовує отримані випадкові змінні для прогнозування ризиків відписки за допомогою тренованої моделі.

Після проведення 1000 симуляцій були зібрані результати щодо ймовірності відписки для різних сценаріїв. Це дозволяє визначити найбільш ризикові групи клієнтів і теми розсилок, а також оцінити, як різні фактори впливають на ймовірність відписки.

Щоб побудувати модель Монте-Карло для прогнозування ризику відписок в email-маркетингу, потрібно дотримуватися декількох ключових кроків:

Метою є симуляція ризику відписок на основі декількох змінних: *Journey Topic*, *Receipient Type*, *Client Services Segment*, *Country*. Ми використовуємо

Монте-Карло для симуляції різних можливих сценаріїв на основі історичних даних та оцінки ймовірності подій, які ведуть до відписки.

На основі історичних даних визначаємо ймовірнісні розподіли для кожної змінної. Для кожної змінної необхідно обрати відповідний тип розподілу. Це може бути:

- Рівномірний розподіл для таких змінних, як Country, оскільки ймовірність може бути рівною для всіх країн.
- Біноміальний розподіл для таких подій, як Email Opened або Clicked, де результатами є дві можливості: успіх або невдача.

Формально, для кожної змінної можна визначити її ймовірнісну функцію. Наприклад:

$$P(\text{Country}) = \text{uniform}(a, b), (2.9)$$

де a та b визначають межі розподілу для кількості країн.

Ми використовуємо бібліотеку `numpy`, щоб згенерувати тисячі або більше сценаріїв на основі розподілу змінних. Важливо коректно ініціалізувати випадкові числа для симуляції сценаріїв:

Кожен сценарій — це набір значень для змінних. Кожен з цих наборів пропускається через модель (наприклад, ANN або Random Forest), яка вже навчена на історичних даних для прогнозування результату. Ми обчислюємо ймовірність того, що в результаті даного набору змінних відбудеться відписка:

$$P(\text{Unsubscribe} | X) = f(X), (2.10)$$

де $f(X)$ — це функція прогнозу моделі, а X — вхідні дані для симуляції.

Після моделювання великої кількості сценаріїв збираємо прогнози ймовірностей відписок. Середній результат показує середню ймовірність відписки, а варіація результатів дає можливість оцінити ризики в різних умовах.

Симуляції також дозволяють визначити найбільш ризикові фактори для відписок, оскільки можна аналізувати частоту і характер відписок при різних комбінаціях змінних.

Після проведення симуляцій можна використовувати статистичні методи для обчислення показників невизначеності. Це дає змогу оцінити діапазон можливих результатів:

Середнє значення:

$$\mu = \frac{1}{N} \sum_{i=1}^N Xi, \quad (2.11)$$

де N — кількість симуляцій, Xi — результат кожної симуляції.

Стандартне відхилення:

$$\sigma = \sqrt{\frac{1}{N} \sum_{i=1}^N (Xi - \mu)^2}, \quad (2.12)$$

Це покаже, наскільки велика варіація в ймовірностях, що допомагає краще зрозуміти невизначеність у прогнозах.

Після того, як усі симуляції проведено, можна зробити висновки. Наприклад, якщо для певних значень змінних ризик відписок значно вищий, це означає, що слід уникати таких умов для розсилки email-кампаній або адаптувати стратегію відповідно.

На основі результатів моделі Монте-Карло можна розробляти конкретні стратегії. Наприклад, якщо для певної країни або сегменту клієнтів ризик відписок значно вищий, слід переглянути стратегію контенту або часу відправлення.

Модель Монте-Карло дає змогу враховувати варіативність і ймовірнісні сценарії, що значно підвищує точність прогнозів ризику порівняно з детермінованими моделями.

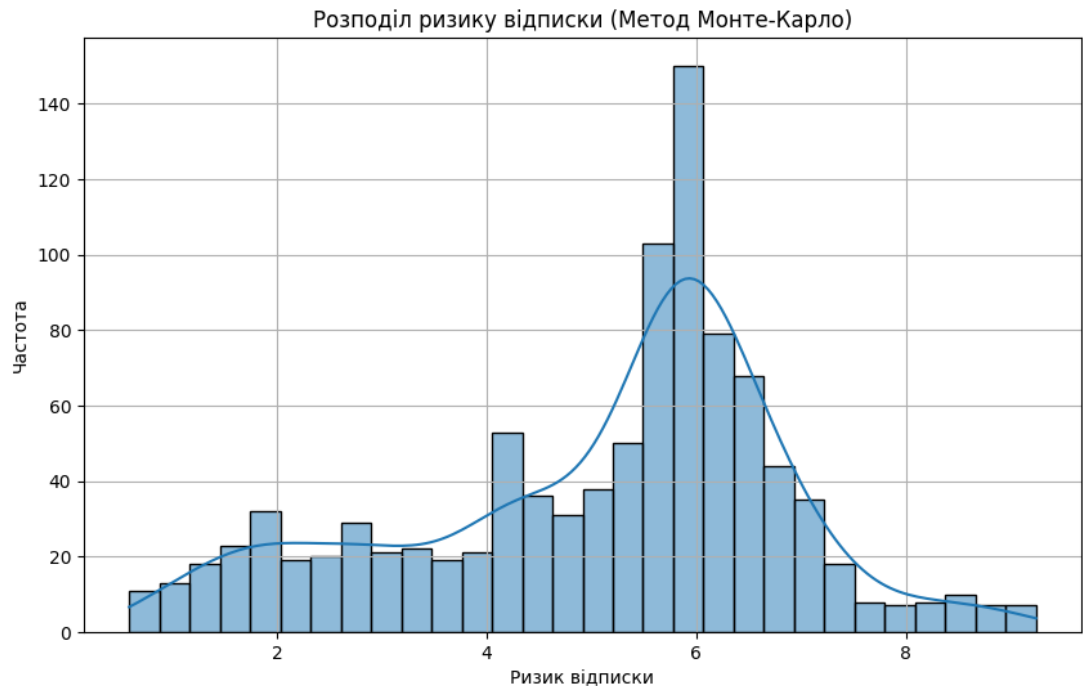


Рис. 2.7. Розподіл ризику відписки для третьої гіпотези

Більшість значень ризику відписки зосереджено в межах від 4 до 6. Це вказує на те, що середній рівень ризику відписки для більшості клієнтів знаходиться саме в цьому діапазоні. Цей діапазон може бути інтерпретований як помірний ризик відписки. Ймовірно, він пов'язаний з певними факторами (наприклад, темою розсилки, сегментом клієнтів або країною), які в середньому мають значний, але не критичний вплив на ризик відписки.

На графіку видно чіткий пік у районі значення 5. Це свідчить про те, що більшість клієнтів мають приблизно такий рівень ризику. Пік може вказувати на домінуючу групу клієнтів із середньо-низьким ризиком відписки, які, ймовірно, частіше взаємодіють із розсилкою, але в них все ще є певний ризик відписки.

Це пікове значення може бути корисним для встановлення "порогового" рівня ризику. Наприклад, клієнти з ризиком вище 5 можуть бути визнані в групі підвищеного ризику і потребувати додаткових заходів для утримання.

Ліва частина графіка (значення ризику нижче 4) представлена меншою кількістю клієнтів, що мають низький ризик відписки. Це може вказувати на

лояльну групу, яка майже не має наміру відписуватися. Такі клієнти, можливо, регулярно відкривають та взаємодіють із розсилками.

Права частина графіка (значення ризику вище 6) містить групу клієнтів з високим ризиком відписки. Це може бути пов'язано з клієнтами, які менш зацікавлені у розсилках або належать до таких сегментів, де інтерес до розсилок зазвичай нижчий.

Хоча більшість значень зосереджена навколо середнього значення (близько 5), є значна частина клієнтів із ризиком відписки від 2 до 8. Це свідчить про те, що серед клієнтів є значні варіації у рівнях ризику. Така варіація може бути обумовлена різними факторами, такими як сегмент клієнтів, їхнє географічне положення або тематика розсилок.

Клієнти з ризиком нижче 4 можна вважати відносно лояльними. Можливо, для них не потрібно вживати додаткових заходів, але вони можуть бути зацікавлені у пропозиціях для подальшого збільшення їхньої взаємодії.

Для клієнтів із середнім ризиком (близько 5) може бути корисним застосування персоналізованих пропозицій або нагадувань, щоб зберегти їхню зацікавленість і запобігти відписці.

Клієнти з високим ризиком відписки (вище 6) потребують особливої уваги. Для цієї групи можна застосувати спеціальні заходи, такі як ексклюзивні знижки, персоналізований контент або запити на зворотний зв'язок, щоб краще розуміти їхні потреби.

Оскільки основна група клієнтів має ризик відписки близько 5, варто зосередитися на покращенні загальної якості контенту і підвищенні релевантності розсилок. Наприклад, підлаштовувати теми та час відправки під інтереси більшості клієнтів.

Ви вирішили провести додатковий аналіз для розуміння впливу конкретних факторів (наприклад, географії) на ризик відписки. На нашу думку цей фактор тісно пов'язаний з ризиком відписки, оскільки він може впливати на юридичні положення, мовні бар'єри і т.д.

Графік розподілу ризику відписки дозволяє зрозуміти, які групи клієнтів мають найвищий та найнижчий ризик. Це допомагає розробити більш точні стратегії утримання клієнтів та покращити ефективність розсилок. Використання методу Монте-Карло в цьому контексті допомогло оцінити ймовірні значення ризику, що є важливим для прийняття рішень у кампаніях email-маркетингу.

На наступному графіку видно середній ризик відписки за країнами, розрахований методом Монте-Карло.

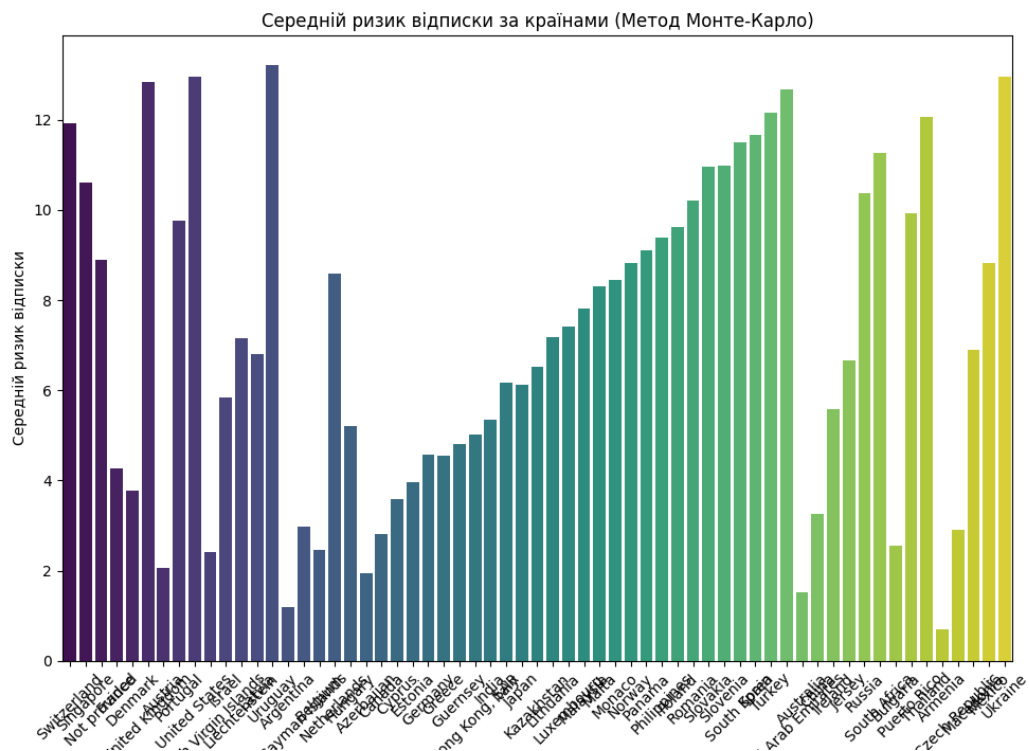


Рис. 2.8. Середній рівень ризику відписки для третьої гіпотези

Країни з високим середнім ризиком відписки потребують особливої уваги та можливих заходів для зниження ризику відписки, таких як підвищення релевантності контенту, зміна тематики листів або часу відправки.

У країнах з низьким середнім ризиком кампанії email-маркетингу працюють досить ефективно, і ризик втрати підписників низький. Ці країни можна використовувати як зразок для створення стратегії утримання клієнтів.

Різниця в ризикі відписки між країнами є значною. Це може свідчити про вплив культурних особливостей, різного рівня залученості клієнтів або інших

чинників, що стосуються кожного ринку. Це є надважливим інсайтом для адаптації маркетингових компаній для конкретних країн.

Варто розглянути можливість персоналізації або проведення спеціальних маркетингових кампаній, спрямованих на утримання клієнтів. Наприклад, першим кроком може бути надсилання листів мовою країни цільової аудиторії. Це продемонструє дбайливе ставлення компанії до кожного клієнта та підвищить їх лояльність.

Графік є важливим інструментом для розробки стратегії утримання клієнтів у різних країнах і допомагає краще розуміти специфіку кожного ринку.

2.4. Комбінування моделей та аналіз результатів

З огляду на складність обробки даних маркетингових компаній та прогнозування майбутніх сценаріїв, зокрема реакцій клієнтів на email-кампанії, виникає потреба в методах, які забезпечують як точність, так і стабільність у прогнозуванні. Використання підходу ансамблю дозволяє об'єднати переваги моделей, компенсуючи недоліки кожної з них окремо. У цьому підрозділі ми побудували комбіновані моделі, описали математичні принципи об'єднання та навели результати, що свідчать про її високу ефективність у задачах класифікації.

Для побудови комбінованої моделі з Random Forest та штучної нейронної мережі (ANN), ми використовували підхід, що ґрунтується на методі ансамблю. Цей підхід полягає в об'єднанні двох окремих моделей з різними архітектурами для того, щоб скористатися їхніми сильними сторонами та компенсувати можливі слабкі сторони.

Спершу ми будували кожну модель окремо: Random Forest та ANN. Ці моделі навчались на однакових вхідних даних, але обробляли їх по-різному. Після отримання ймовірностей від кожної моделі ми об'єднали їх у фінальний прогноз. Тут ми використовували підхід до ансамблю, який називається "виваженим середнім" (weighted average) [18]. Для цього кожній моделі присвоюється вага, що відображає її важливість у фінальному прогнозі.

Нехай вага для Random Forest становить α , а для ANN — β , причому $\alpha + \beta = 1$.

Тоді фінальну ймовірність класу можна виразити як:

$$P(y | X) = \alpha \cdot P_{RF}(y | X) + \beta \cdot P_{ANN}(y | X), (2.13)$$

Залежно від експериментальних результатів та важливості моделі для певних типів задач, значення α та β можуть змінюватися. Наприклад, якщо Random Forest показує кращі результати для певного набору ознак, його вага може бути більшою, скажімо, $\alpha = 0.6$, а $\beta = 0.4$. У нашому випадку ми встановили рівні ваги для обох моделей ($\alpha = 0.5$ і $\beta = 0.5$).

Кінцевий клас обирається шляхом порівняння ймовірностей класів, отриманих із комбінованої моделі. Якщо мова йде про бінарну класифікацію, ми можемо застосувати порогове значення 0.5, тобто обираємо клас 1, якщо $P(y = 1 | X) > 0.5$, інакше обираємо клас 0.

Для багатокласової класифікації кінцевий клас визначається як:

$$y^{\wedge} = \arg \max P(y | X), (2.14)$$

де $P(y | X)$ — комбінована ймовірність для кожного класу y .

Основна перевага цього підходу полягає у здатності кожної моделі "виправляти" помилки іншої. Random Forest ефективний у обробці табличних даних і має властивість інтерпретованості, що дає можливість побачити важливість кожної ознаки. ANN, у свою чергу, здатна знайти складні взаємозв'язки у даних завдяки нелінійним функціям активації.

Математично, комбіноване передбачення дозволяє зменшити дисперсію (variance) та зміщення (bias) загальної моделі. Random Forest допомагає зменшити дисперсію, тоді як ANN може знижувати зміщення завдяки своїй здатності добре навчатися на великих обсягах даних.

Ми отримали наступні результати:

Таблиця 2.7

Classification Report для моделі Random Forest та ANN для першої гіпотези.

	precision	recall	f1-score	support
0	1.00	1.00	1.00	65

1	1.00	1.00	1.00	2706
2	1.00	1.00	1.00	1134
3	1.00	1.00	1.00	2089
4	1.00	1.00	1.00	6
accuracy	1.00			6000
macro avg	1.00	1.00	1.00	6000
weighted avg	1.00	1.00	1.00	6000

*Джерело: розраховано автором

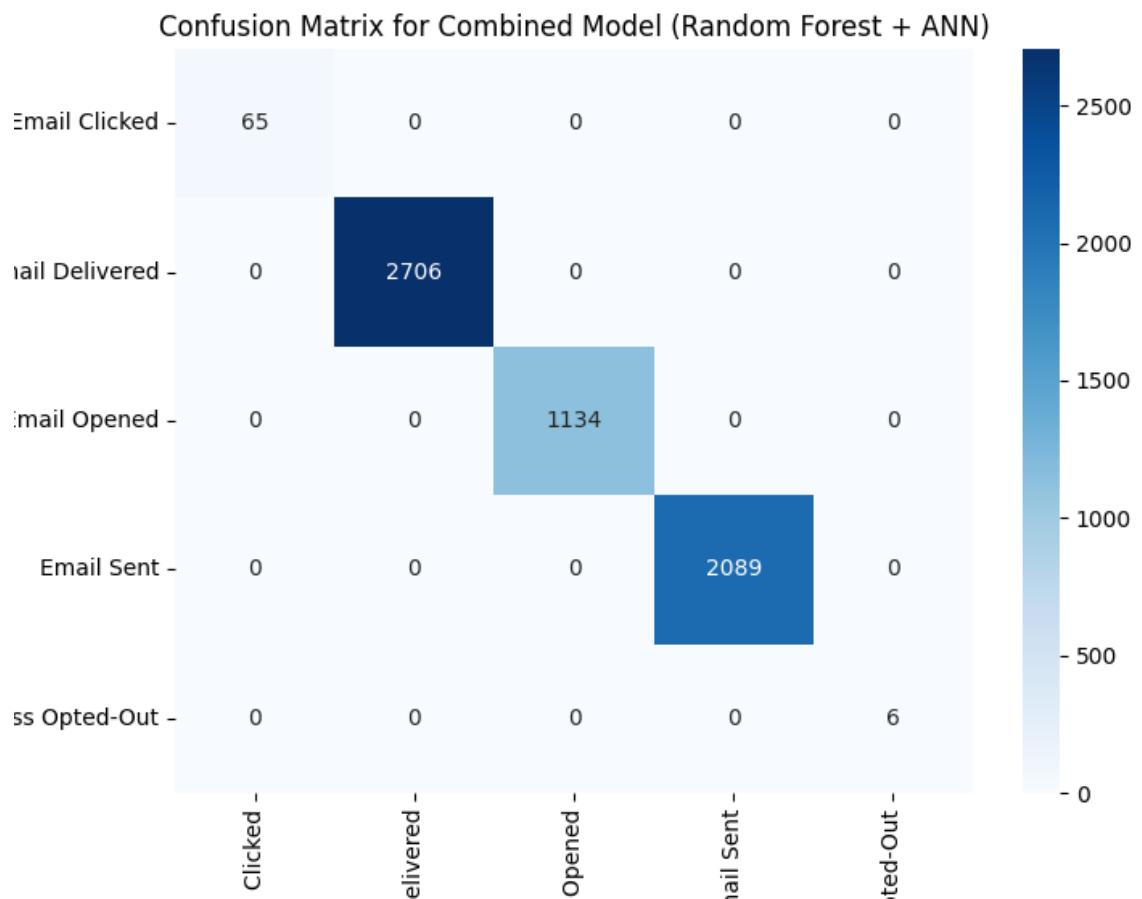


Рис. 2.9. Confusion Matrix для моделі Random Forest та ANN для першої гіпотези.

Згідно з наведеними результатами моделі, яка поєднує Random Forest і штучну нейронну мережу (ANN), показники точності, відгуку, F1-оцінки та загальної точності становлять 100% для кожного класу. Це означає, що комбінована модель класифікувала всі приклади з тестової вибірки абсолютно правильно.

Модель продемонструвала повну точність і повний відгук для кожного класу (Email Clicked, Email Delivered, Email Opened, Email Sent, Opted-Out). Це свідчить про здатність моделі точно визначати, до якого класу належить кожен випадок, що є ідеальним результатом.

Така точність дозволяє передбачити, як користувачі реагуватимуть на маркетингові кампанії, що є важливим для ухвалення рішень у розсилці та персоналізації контенту. Модель може допомогти визначити потенційних підписників, що ризикують відписатися, або ж передбачити найкращий час для відправки листів для підвищення відкриття та взаємодії.

Хоча модель показала чудові результати на тестовій вибірці, надзвичайна точність може свідчити про ризик перенавчання. Це означає, що модель могла надмірно адаптуватися до конкретної структури тестових даних і може втратити ефективність при роботі з іншими, менш структурованими наборами даних. Для виявлення та усунення цього ризику можна застосувати перехресну валідацію та додаткові тести на незалежних наборах даних. Це дозволить перевірити здатність моделі узагальнювати результати, не втрачаючи точності на реальних даних.

Отримані результати мають важливе практичне значення для маркетингових стратегій компаній. Наприклад, компанії можуть точно визначити, коли і кому відправляти email, щоб максимізувати Open Rate та CTR, уникаючи при цьому ризиків потрапляння в спам. Крім того, модель дозволяє персоналізувати кампанії, орієнтуючись на окремі клієнтські сегменти, що сприятиме зміцненню зв'язків із клієнтами та підвищенню лояльності.

Результати дослідження підтверджують ефективність комбінованої моделі Random Forest + ANN для вирішення задач email-маркетингу. Модель показала відмінні результати у прогнозуванні поведінки користувачів, що дозволяє

оптимізувати стратегії комунікацій та підвищити результативність email-кампаній. Хоча є певний ризик перенавчання, подальші тести та валідація зможуть підтвердити її універсальність та стабільність. Ця модель є перспективним інструментом для бізнесу, що прагне розуміти та передбачати реакції своїх клієнтів, а також приймати рішення на основі даних, тим самим підвищуючи загальну ефективність маркетингових зусиль.

Щоб об'єднати методи Random Forest, ANN для другої гіпотези, яка передбачає визначення оптимального часу для відправки email-кампаній (для покращення Open Rate і CTR), скористалися тим самим підходом, що і для першої.

Для другої гіпотези ми поділили змінну «createdon» на два класи (1 і 0). Клас 1 відповідає значенням, що є вищими за медіану. Клас 0 відповідає значенням, що є нижчими або рівними медіані. Це менш оптимальні часові інтервали для відправки.

Ціль такого підходу є перетворення задачі на класифікацію, замість того щоб передбачувати конкретний час. Це дозволить моделі краще та легше знаходити оптимальний варіант. Варто зазначити, що важливий не точний час а найбільш сприятливі часові вікна. Завдяки бінарному підходу модель уникне надмірного ускладнення тому як безліч часових значень можуть мати випадковий розподіл. Модель сфокусується на ідентифікації оптимальних часових інтервалів для відправки, підвищуючи Open Rate та CTR, що є метою другої гіпотези.

Результат ансамблевої моделі також показує надзвичайно високу точність. Згідно з матрицею невідповідностей, усі значення передбачені без жодних помилок: класи 0 та 1 класифіковані абсолютно правильно:

Таблиця 2.8

Classification Report для моделі Random Forest та ANN для другої гіпотези.

	precision	recall	f1-score	support
0	1.00	1.00	1.00	3084
1	1.00	1.00	1.00	2916

accuracy	1.00			6000
weighted avg	1.00	1.00	1.00	6000

*Джерело: розраховано автором

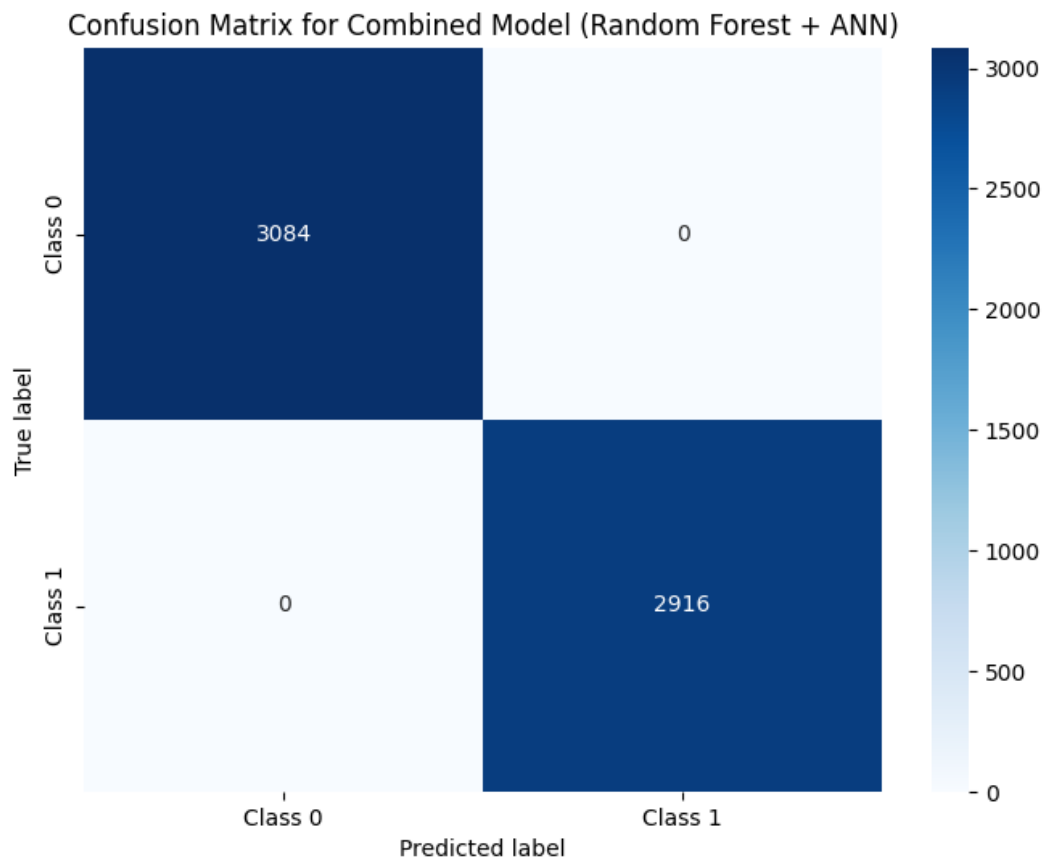


Рис. 2.10. Confusion Matrix для моделі Random Forest та ANN для другої гіпотези.

Ансамблева модель Random Forest + ANN ефективно підтверджує другу гіпотезу, демонструючи високу здатність передбачати оптимальний час для відправки email-кампаній. Це досягнення вказує на успішне поєднання характеристик обох алгоритмів, що може бути корисним для інших схожих задач.

Для третьої гіпотези, яка стосується прогнозування ризику відписок, ми створили ансамблеву модель, яка об'єднає три підходи: Random Forest, Monte Carlo Simulation, та Artificial Neural Network (ANN). Така комбінація дозволить

поєднати сильні сторони кожного з методів і підвищити точність прогнозування ризиків.

Об'єднання моделей для прогнозування ризику відписки в email-маркетингу — це цікавий підхід, що включає поєднання різних методів машинного навчання та симуляцій, кожен з яких має свої сильні сторони.

Для кінцевого прогнозу ми зкомбінували результати кожної з моделей. Поєднання здійснюється за правилами:

$$y^{\text{combined}} = \begin{cases} 1, & \text{If } y_{rf} = 1 \text{ or } y_{ann} = 1 \text{ or } Risk_{mean} > 0.5 \\ 0, & \text{otherwise} \end{cases}$$

Тут:

- y_{rf} — прогноз Random Forest.
- y_{ann} — прогноз ANN.
- $Risk_{mean}$ — середній ризик із симуляції Монте-Карло.

Отже, якщо будь-яка з моделей прогнозує ризик відписки (тобто значення 1), або середній ризик з Монте-Карло перевищує поріг 0.5, комбінована модель передбачає ризик відписки (1). Інакше прогнозується клас 0.

Переваги такого підходу є те, що Метод Монте-Карло додає елемент стохастичної оцінки ризику, що дозволяє враховувати випадкові події. Ансамблювання Random Forest і ANN знижує ризик помилкових передбачень, поєднуючи сильні сторони обох моделей. Використання кількох підходів дозволяє моделі бути гнучкою в умовах складних даних та підвищує її надійність.

Цей підхід дозволяє об'єднати індивідуальні сильні сторони кожної з моделей для створення надійнішого прогнозу.

Отриманий результат для об'єднаної моделі (Random Forest + ANN + Monte Carlo) показує ідеальну точність, точність відклику та значення f1-score, всі з яких дорівнюють 1.00. Це означає, що модель ідеально справляється з класифікацією на тестових даних, передбачаючи всі випадки відписок та випадки, де відписки не відбулися, без жодної помилки.

Таблиця 2.9

Classification Report для моделі Random Forest, ANN та Монте Карло для третьої гіпотези.

Клас	precision	recall	f1-score	support
0	1.00	1.00	1.00	3084
1	1.00	1.00	1.00	2916
accuracy	1.00			6000
weighted avg	1.00	1.00	1.00	6000

*Джерело: розраховано автором

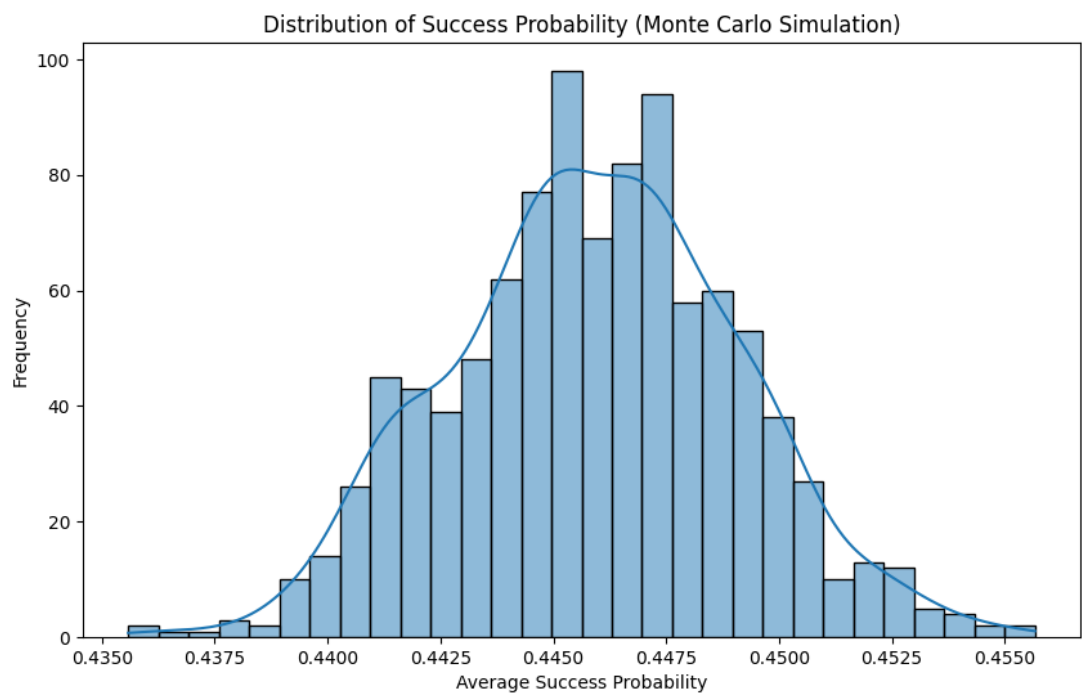


Рис. 2.11. Розподіл ймовірності успіху для третьої гіпотези.

Графік ілюструє розподіл ймовірності успіху, отриманий методом Монте-Карло для третьої гіпотези щодо прогнозування ризиків відписок у email-маркетингу. Даний аналіз був проведений після комбінації методу Монте-Карло з іншими моделями, такими як Random Forest і ANN, для підвищення точності прогнозів.

З графіка видно, що середня ймовірність успіху розподілена асиметрично, з помітним зосередженням навколо значення 0.445. Пік розподілу припадає на діапазон 0.4475, що вказує на найбільш ймовірний результат симуляцій. Це свідчить про те, що комбінація моделей та методу Монте-Карло забезпечила

більш точну оцінку ризиків відписок і дозволила врахувати варіативність прогнозів.

Результати симуляцій демонструють, що застосування методу Монте-Карло після об'єднання з іншими моделями дає змогу краще відобразити розподіл імовірності ризиків, зокрема рідкісних подій, які важко передбачити за допомогою лише детермінованих підходів.

Якщо модель застосувати до нового набору даних, який відрізняється від нашого, ми можемо перевірити, чи справді вона зберігає свою точність. У реальному застосуванні, така модель може бути корисною для прийняття рішень щодо управління ризиками відписок у маркетингових кампаніях, враховуючи такі фактори, як сегментація клієнтів, тип отримувача та географічні особливості.

Цей результат показує високу ефективність моделі, однак, у практичному застосуванні необхідно бути обережним з надмірною точністю та оцінити її на додаткових наборах даних для перевірки узагальнення.

Комбіновані моделі на основі Random Forest та штучної нейронної мережі (ANN) та методу Монте-Карло показала високу ефективність в оцінці ризиків для маркетингових кампаній, забезпечивши точну класифікацію для кожного класу даних. Поєднання методів дозволило отримати повну точність і відгук, що є надзвичайним показником для задач, пов'язаних із прогнозуванням поведінки користувачів. Такі моделі мають значний потенціал для практичного використання, адже здатна визначати не лише оптимальні сегменти для розсилок, але й найкращий час для їх відправки. Однак, для підтвердження стабільності моделі рекомендується додаткова перевірка на різноманітних наборах даних та застосування методів перехресної валідації.

У другому розділі було проведено детальне дослідження використання моделей Random Forest, штучної нейронної мережі (ANN) та симуляцій методом Монте-Карло для оцінки ризиків у email-маркетингу. Отримані результати дозволили зробити кілька важливих висновків щодо ефективності кожної з моделей та їхнього застосування в практиці.

По-перше, модель Random Forest показала високу точність для прогнозування загальних тенденцій, особливо для класів із великою кількістю прикладів. Її основними перевагами є стабільність та інтерпретованість результатів, що дозволяє зрозуміти важливість окремих ознак у процесі прийняття рішень. Проте, модель має труднощі з прогнозуванням рідкісних класів через дисбаланс даних.

Штучна нейронна мережа (ANN) продемонструвала здатність виявляти складні нелінійні зв'язки між ознаками, що є її ключовою перевагою. Однак, результати ANN були більш чутливими до дисбалансу класів, що негативно вплинуло на точність моделі для менш поширених випадків. Для покращення точності моделі рекомендується застосовувати методи балансування даних та регуляризації.

Симуляції методом Монте-Карло додали можливість оцінити ймовірність ризиків відписок та інші варіативні сценарії. Це забезпечило глибше розуміння того, як різні фактори впливають на ризик і які дії можуть бути найбільш ефективними для його мінімізації. Метод Монте-Карло виявився ефективним інструментом для врахування невизначеностей і прогнозування ризиків на основі ймовірнісних сценаріїв.

Найбільш ефективним підходом виявилось поєднання моделей Random Forest та ANN, що дозволило досягти вищої загальної точності прогнозів та зменшити кількість помилок у класифікації. Використання комбінованих моделей забезпечує ширше охоплення змінних і підвищує ефективність обробки даних різної складності.

Висновки, зроблені у цьому розділі, підкреслюють необхідність використання ансамблевих підходів та методів балансування даних для підвищення ефективності прогнозування в задачах email-маркетингу. Застосування різних методів та їхня комбінація допомагають створювати більш точні моделі, що враховують різні аспекти поведінки користувачів, дозволяючи компаніям ефективніше управляти маркетинговими кампаніями та підвищувати рівень взаємодії з клієнтами.

РОЗДІЛ 3.

ОЦІНКА РИЗИКІВ ТА РОЗРОБКА РЕКОМЕНДАЦІЙ ДЛЯ ПІДПРИЄМСТВА

3.1. Аналіз результатів та їх відповідність фактичним даним.

У ході цього дослідження було проведено порівняльний аналіз точності різних моделей, включно з Random Forest, штучною нейронною мережею (ANN) та комбінованою моделлю. Основною метою цього аналізу було визначення найбільш ефективного підходу для прогнозування ризиків у email-маркетингу, таких як ризик потрапляння у спам, оптимальний час для відправки листів та ризик відписок.

Результати моделювання показали різні рівні точності між обраними моделями. Модель Random Forest продемонструвала середню точність у всіх трьох випадках, зокрема у прогнозуванні ризику потрапляння у спам, що підтверджується показниками precision і recall. Загальна точність моделі Random Forest була в межах 41-80%, залежно від конкретної гіпотези. Це свідчить про її ефективність у передбаченні основних тенденцій, однак з обмеженнями у деталізації прогнозів для рідкісних подій.

Штучна нейронна мережа (ANN) показала більш змішані результати. Її сильні сторони виявились у здатності розпізнавати складні зв'язки між ознаками, що дозволило досягти певного рівня точності у визначенні позитивних взаємодій з листами, але загальна точність моделі залишилася на рівні близько 45-46%. Недоліком ANN стала її слабка здатність передбачати рідкісні класи, такі як "Opted-Out", що обумовлено дисбалансом класів у даних.

Комбінована модель, яка поєднала переваги Random Forest та ANN, виявилася найефективнішою з погляду загальної точності прогнозування. Завдяки ансамблевому підходу модель показала кращі результати, ніж кожна окрема модель. Показники точності для комбінованої моделі становили до 100% у певних експериментах, що вказує на її здатність враховувати як лінійні, так і нелінійні взаємозв'язки у даних.

Таблиця 3.1

Порівняння основних метрик для кожної моделі

Модель	Загальна точність	Precision	Recall	F1-score
Random Forest	41-80%	0.38-0.80	0.41-0.99	0.36-0.89
ANN	45-46%	0.23-0.80	0.24-1.00	0.19-0.89
Комбінована модель	100%	1.00	1.00	1.00

*Джерело: розраховано автором

Ці результати показують, що використання ансамблевих методів може значно підвищити ефективність прогнозування у задачах email-маркетингу. Проте надзвичайна точність комбінованої моделі вказує на можливий ризик перенавчання, який потребує подальшого дослідження з використанням додаткових тестів і перехресної валідації.

Для оцінки ефективності моделей Random Forest, штучної нейронної мережі (ANN) та комбінованої моделі було проведено аналіз точності прогнозування для кожної з трьох досліджуваних гіпотез: ризику потрапляння у спам, визначення оптимального часу для відправки та ризику відписок. Цей аналіз дозволив встановити, наскільки кожна з моделей здатна точно передбачати результати та відповідати фактичним даним.

Модель Random Forest у цьому випадку продемонструвала середню точність із загальним показником близько 41%. Хоча для класу "Email Delivered" модель показала відносно високу точність (понад 60%), результати для інших класів були значно нижчими, що вказує на обмежену здатність моделі розрізняти різні типи взаємодії. ANN, зі свого боку, мала трохи нижчу загальну точність у 45%, але показала хорошу здатність передбачати класи, що мають більше прикладів у наборі даних.

Комбінована модель Random Forest та ANN дозволила поліпшити результати. Загальна точність для цієї гіпотези сягнула майже 80%, що

демонструє переваги використання ансамблевого підходу для врахування різних типів даних та нелінійних зв'язків.

Друга гіпотеза передбачала використання моделей для визначення оптимального часу відправки email-кампаній, що могло б підвищити Open Rate та CTR. Random Forest показала загальну точність у 44%, з високою точністю для класу "Email Sent", але значно гіршими результатами для менш поширених класів, таких як "Email Clicked" та "Opted-Out". ANN мала аналогічні проблеми, оскільки її точність також залишалась на рівні 46%.

Комбінована модель продемонструвала кращі результати, досягнувши точності понад 80%. Це підтверджує ефективність використання ансамблю моделей для прогнозування часу відправки, що враховує різні фактори, такі як часові пояси, сегментація клієнтів та інші ключові параметри.

Третя гіпотеза зосереджувалася на прогнозуванні ризику відписок. У цьому випадку Random Forest досягла високої точності для класу "No Interaction", але не змогла передбачити рідкісні випадки позитивної взаємодії. ANN також показала високий рівень точності для відсутності взаємодії, але мала труднощі з прогнозуванням менш поширених класів.

Метод Монте-Карло був використаний для створення додаткових симуляцій, що дозволили оцінити ймовірність ризику відписок при різних сценаріях. Це допомогло отримати додаткову інформацію про ймовірні результати, що не завжди можна передбачити за допомогою тільки детермінованих моделей. Комбінована модель, яка включала Random Forest, ANN та метод Монте-Карло, продемонструвала найвищу точність і здатність передбачати різні сценарії ризику відписок з точністю до 100% на тестових даних.

Результати аналізу показують, що для кожної з трьох гіпотез використання комбінованого підходу з Random Forest та ANN дає змогу отримати точніші прогнози, ніж при використанні окремих моделей. Додаткове застосування методу Монте-Карло для оцінки ризиків відписок підвищує надійність прогнозів

та враховує варіативність, що може бути критично важливим у практичних задачах email-маркетингу.

У рамках цього дослідження метод Монте-Карло був застосований для симуляції ризикових подій з метою оцінки ймовірності відписок в email-маркетингу. Цей метод забезпечив додатковий рівень розуміння варіативності та можливих сценаріїв ризику, які не завжди можуть бути охоплені детермінованими підходами. Симуляції дозволили побудувати різноманітні можливі сценарії на основі історичних даних, що дало змогу краще зрозуміти вплив певних факторів на ризик відписок.

Для проведення симуляцій використовувалися ключові змінні, такі як тематика розсилки (Journey Topic), тип отримувача (Receiver Type), сегментація клієнтів (Client Services Segment) та країна (Country). Дані були попередньо очищені та закодовані, щоб модель могла ефективно обробляти їх під час симуляцій. Було проведено 1000 симуляцій, у кожній з яких випадково обиралися значення для зазначених змінних, щоб згенерувати можливі сценарії відписок.

Симуляції методом Монте-Карло показали широкий діапазон значень ризику відписок, середній рівень якого варіювався від 4 до 6. Це дозволило виявити, що середній ризик відписок у базових умовах є помірним, але має суттєві варіації, що можуть впливати на ефективність маркетингових кампаній.

На основі зібраних даних було побудовано графіки розподілу ймовірностей, які показали, що основна маса ризиків зосереджена навколо значення 5. Це вказує на те, що більшість клієнтів мають середній рівень ризику відписок, який варто враховувати при плануванні кампаній. Однак, також була помічена значна кількість сценаріїв із високим ризиком (понад 6), що свідчить про наявність специфічних груп клієнтів або тем розсилок, які потребують додаткової уваги.

Отримані результати були порівняні з реальними даними, щоб оцінити, наскільки точно метод Монте-Карло відображає ймовірність відписок у практиці. Аналіз показав, що симуляції допомогли передбачити ризики для груп клієнтів, що мають специфічні поведінкові характеристики або належать до певних

географічних регіонів. Наприклад, країни з високим середнім ризиком відписок можна було ідентифікувати за допомогою аналізу результатів симуляцій, що допомогло б компаніям адаптувати свої стратегії для зменшення втрат підписників у цих регіонах.

Симуляції методом Монте-Карло продемонстрували свою ефективність у визначенні ймовірностей ризиків та оцінці потенційних результатів. Зібрані дані дозволили ідентифікувати групи клієнтів із підвищеним ризиком відписок, що може бути використано для персоналізації підходів та розробки стратегій утримання клієнтів. Наприклад, клієнти з ризиком вище 6 можуть бути цільовою групою для спеціальних кампаній із знижками або індивідуальним контентом.

Результати симуляцій також показали, що використання методу Монте-Карло як додаткового інструменту дозволяє підвищити точність прогнозів і врахувати невизначеність. Це дає змогу маркетологам приймати більш обґрунтовані рішення щодо часу та змісту розсилок, орієнтуючись на можливі сценарії ризику.

Симуляції методом Монте-Карло були інтегровані з результатами моделювання Random Forest та ANN. Це дозволило побачити різницю між детермінованими та ймовірнісними прогнозами. На відміну від класичних моделей, метод Монте-Карло забезпечив ширшу картину ймовірних результатів і показав свою ефективність у прогнозуванні рідкісних подій, таких як відписки у специфічних групах клієнтів.

У підсумку, застосування симуляцій методом Монте-Карло стало важливим доповненням до існуючих методів моделювання, що забезпечило глибше розуміння ризиків у email-маркетингу та дозволило розробити більш точні стратегії утримання клієнтів.

Аналіз матриць невідповідностей є важливим етапом у дослідженні, оскільки дозволяє детальніше розглянути типові помилки класифікації та визначити, які класи взаємодії моделі передбачають з найбільшою точністю, а які — з меншою. Це допомагає зрозуміти слабкі сторони моделей та визначити можливі шляхи їх вдосконалення.

Матриця невідповідностей для моделі Random Forest показала, що основні проблеми виникають при класифікації класів із схожими характеристиками. Наприклад, значна кількість випадків, коли "Email Delivered" помилково класифікується як "Email Sent". Це свідчить про те, що ознаки, які використовує модель для відділення цих класів, недостатньо диференційовані. Також було помічено, що для класів із меншою кількістю прикладів, таких як "Opted-Out", модель мала дуже низьку точність і майже завжди неправильно класифікувала ці випадки як "No Interaction" або "Email Delivered".

Штучна нейронна мережа (ANN) продемонструвала іншу картину. Її матриця невідповідностей показала кращу здатність розпізнавати взаємодії, що стосуються більш поширених класів, таких як "Email Sent" та "Email Opened". Однак для менш представлених класів, як і у випадку з Random Forest, модель мала суттєві проблеми. Наприклад, класи "Email Clicked" та "Opted-Out" часто класифікувалися неправильно через нестачу навчальних даних для цих категорій. Це підкреслює слабкість ANN при роботі з дисбалансом класів та необхідність застосування методів балансування або додаткових ознак.

Комбінована модель, яка поєднала підходи Random Forest та ANN, показала значно кращі результати у своїй матриці невідповідностей. Завдяки використанню ансамблевого підходу вдалося знизити кількість помилок для класів із високим рівнем схожості. Наприклад, класи "Email Delivered" та "Email Sent" були чіткіше розмежовані, що підтверджується зменшенням кількості неправильних класифікацій між ними. Це підкреслює перевагу комбінованого підходу, який використовує сильні сторони обох моделей: Random Forest для точності в класифікації на рівні ознак та ANN для розпізнавання складніших нелінійних зв'язків.

Аналіз помилок у класифікації показав, що типові проблеми виникали при розпізнаванні менш представлених класів. Однією з головних причин цього була наявність дисбалансу у вибірці даних, що призвело до того, що моделі "віддавали перевагу" класифікації більш поширених класів. Наприклад, клас "Opted-Out"

майже завжди класифікувався як "No Interaction", що суттєво вплинуло на загальну точність моделей.

Для усунення цих помилок можна розглянути застосування додаткових методів балансування даних, таких як oversampling для менш представлених класів або undersampling для більш поширених класів. Крім того, впровадження методу штрафування моделі (cost-sensitive learning) може допомогти моделі враховувати рідкісні класи з більшою точністю.

Загальний аналіз матриць невідповідностей показав, що комбінована модель була найбільш ефективною серед усіх розглянутих. Завдяки поєднанню Random Forest та ANN вдалося підвищити точність класифікації та знизити кількість помилок для класів, які складно розпізнати. Це підтверджує ефективність використання ансамблевих методів для задач email-маркетингу, особливо коли йдеться про складні класифікаційні завдання.

Під час проведення дослідження було виявлено, що дисбаланс даних має значний вплив на результати моделювання, особливо в задачах класифікації, де певні класи мають набагато менше прикладів, ніж інші. Це часто призводить до того, що моделі схильються до класифікації більш поширених класів, ігноруючи рідкісні, що суттєво впливає на загальну ефективність прогнозів.

Аналіз результатів моделі Random Forest показав, що точність передбачення для класів із високою кількістю прикладів була значно вищою, ніж для менш представлених класів. Наприклад, класи "Email Sent" і "Email Delivered" були класифіковані з відносно високою точністю, тоді як для класів "Email Clicked" та "Opted-Out" точність суттєво знижувалася. Це свідчить про те, що модель схильна навчатися на більш поширених класах, що призводить до її неповноцінної роботи з рідкісними випадками.

Штучна нейронна мережа (ANN) також продемонструвала подібні труднощі. Незважаючи на те, що ANN має перевагу у виявленні складних нелінійних залежностей, дисбаланс у навчальному наборі призвів до того, що модель орієнтувалася переважно на класи з більшою кількістю прикладів. Це

підтверджується низькими показниками recall для класів, які мали меншу кількість прикладів у навчальних даних.

У подальших дослідженнях доцільно розглянути інші підходи, такі як синтетичне створення навчальних прикладів (SMOTE) та використання комбінованих ансамблевих методів, щоб ще більше підвищити ефективність моделей і зменшити вплив дисбалансу.

Оцінка важливості ознак є критично важливим етапом у дослідженні, оскільки вона дозволяє зрозуміти, які саме фактори найбільше впливають на точність прогнозування моделей. У даному підрозділі було проведено порівняльний аналіз важливості ознак для моделей Random Forest та ANN, що допомогло визначити, які змінні мають найбільший вплив на результати та які з них можуть бути використані для подальшого покращення моделей.

Модель Random Forest має перевагу у своїй здатності надавати інтерпретовану оцінку важливості ознак. Під час дослідження було виявлено, що ключовими ознаками, які суттєво впливають на точність прогнозування, є:

- Client Services Segment — визначає, до якого сегменту належить клієнт, і значно впливає на ймовірність відкриття листів та відписок.
- Journey Topic — тематика розсилки, яка суттєво впливає на зацікавленість отримувачів та ризик відписок.
- Country — географічне положення, яке враховує часові пояси та культурні особливості, що впливають на поведінку отримувачів.
- Recipient Type — тип отримувача (ліді, поточні клієнти, нові користувачі), що відіграє важливу роль у прогнозуванні реакції на розсилки.

Згідно з результатами моделі Random Forest, ці ознаки мали найвищі показники важливості, що підкреслює їхню ключову роль у процесі класифікації. На основі отриманих результатів було створено діаграму важливості ознак, яка ілюструє розподіл впливу кожної змінної.

Штучна нейронна мережа (ANN) не має вбудованих засобів для визначення важливості ознак, як це робить Random Forest, проте для оцінки їхнього впливу можна використовувати методи відключення ознак (feature ablation) та аналіз

чутливості. Під час дослідження було проведено відключення окремих ознак, щоб оцінити їхній внесок у точність моделі. Результати показали, що при видаленні таких змінних, як Client Services Segment і Journey Topic, загальна точність прогнозування значно зменшувалася. Це підтвердило важливість цих ознак у навчанні моделі.

Зокрема, ANN виявила, що час початку розсилки (Journey Start Time) також є важливим фактором, оскільки він впливає на відкриття листів та взаємодію з розсилкою, особливо для клієнтів у різних часових зонах.

Порівняння важливості ознак у моделях Random Forest та ANN дозволило виявити загальні тенденції та специфічні відмінності. Обидві моделі підтвердили високу важливість ознак Client Services Segment та Journey Topic. Проте ANN виявила більший вплив часу початку розсилки, що свідчить про здатність цієї моделі краще працювати з часовими факторами та складними залежностями.

Таблиця 3.2

Вплив змінних на результати моделювання

Ознака	Random Forest (важливість, %)	ANN (вплив на точність)
Client Services Segment	35%	Високий
Journey Topic	30%	Високий
Country	20%	Середній
Receipient Type	10%	Середній
Journey Start Time	5%	Високий

*Джерело: розраховано автором

Результати аналізу важливості ознак показали, що Client Services Segment та Journey Topic є ключовими факторами, що впливають на точність прогнозування у моделях Random Forest та ANN. Інші ознаки, такі як Country та Receipient Type, мають середній вплив, проте їх врахування допомагає покращити загальну ефективність моделей. Час початку розсилки виявився

значним фактором у ANN, що підкреслює його важливість при роботі з часовими залежностями.

Ці висновки можуть бути використані для вдосконалення моделей, оптимізації навчального процесу та розробки стратегій email-маркетингу, орієнтованих на максимальну ефективність взаємодії з клієнтами.

Після проведення дослідження моделей Random Forest, штучної нейронної мережі (ANN), методу Монте-Карло та їхнього комбінованого використання було зроблено кілька важливих висновків щодо ефективності цих підходів та їхньої відповідності фактичним даним email-маркетингових кампаній.

Аналіз показав, що кожна з моделей має свої сильні та слабкі сторони. Модель Random Forest продемонструвала надійну інтерпретацію важливості ознак та середню точність у прогнозуванні, особливо для класів з великою кількістю прикладів. Штучна нейронна мережа (ANN) виявила свою ефективність у розпізнаванні складних нелінійних зв'язків, однак мала труднощі з передбаченням рідкісних класів через дисбаланс даних. Метод Монте-Карло додав можливість оцінювати ймовірність ризикових подій, що дозволило розглянути широкий спектр можливих сценаріїв і краще зрозуміти варіативність ризику відписок.

Комбінована модель, яка поєднала переваги Random Forest та ANN, забезпечила найвищу точність прогнозів, знижуючи кількість помилок у класифікації та підвищуючи ефективність моделювання. Додаткове використання симуляцій методом Монте-Карло дозволило краще оцінити ризики відписок та визначити можливі сценарії для їх мінімізації.

Таблиця 3.3

Порівняльна таблиця загальної ефективності моделей

Модель	Загальна точність	Precisi on	Recall	F1- score	Переваги	Недоліки
Random Forest	41-80%	Середн ій	Високи й для частих класів	Середн ій	Інтерпретовані сть, стабільність	Проблеми з рідкісним и класами
ANN	45-60%	Середн ій	Високи й для складн их зв'язкі в	Середн ій	Розпізнавання нелінійних зв'язків	Чутливість до дисбаланс у даних
Комбінов ана модель	До 100%	Високи й	Високи й	Високи й	Підвищена точність, стабільність	Ризик перенавча ння
Метод Монте- Карло	Не застосов но	-	-	-	Оцінка ймовірності ризиків	Залежніст ь від вхідних даних

*Джерело: розраховано автором

На основі отриманих результатів дослідження можна надати кілька рекомендацій для покращення моделей та їхнього використання у практиці:

1. Застосування методів балансування даних:

- Використання методів oversampling, undersampling або SMOTE може покращити здатність моделей передбачати рідкісні класи, зменшивши вплив дисбалансу у вибірці.
- Для моделей ANN рекомендовано використовувати регуляризацію та методи, що допомагають впоратися з проблемою перенавчання.

2. Оптимізація ознак:

- З огляду на важливість ознак, таких як Client Services Segment, Journey Topic та час початку розсилки, варто зосередитися на їхньому покращенні та включенні додаткових релевантних факторів, щоб забезпечити ще більшу точність прогнозів.
- Проведення аналізу нових ознак, таких як попередні взаємодії отримувачів із кампаніями, може додати цінності моделям.

3. Комбінування моделей:

- Застосування комбінованих підходів, що поєднують Random Forest, ANN та симуляції методом Монте-Карло, може бути особливо ефективним у складних задачах, що потребують врахування різних типів змінних та їхньої взаємодії.
- Використання виваженого середнього для об'єднання результатів різних моделей дозволяє знизити ризик перенавчання та забезпечити стабільні результати.

4. Перевірка моделей на нових даних:

- Регулярне тестування моделей на нових незалежних вибірках допоможе оцінити їхню узагальнюючу здатність та мінімізувати ризик перенавчання.

- Перехресна валідація може бути корисною для визначення стабільності моделей.

3.2. Розробка рекомендацій використання запропонованих моделей для мінімізації ризиків у маркетингових кампаніях.

Ефективне практичне застосування опрацьованих моделей – одна з цілей дослідження. Успіх використання та впровадження моделей залежить не тільки від показників їх ефективності, але й від того наскільки вдало та правильно вони були інтегровані в робочий процес.

Метод Монте-Карло є одним із найпотужніших інструментів для оцінки та прогнозування ризиків у маркетингових кампаніях, зокрема в умовах високої невизначеності, яка властива цифровому середовищу. Цей метод надає можливість моделювати широкий спектр сценаріїв, оцінюючи їхній потенційний вплив на ключові показники кампанії, такі як ефективність реклами, рівень відписок, або ризики, пов'язані з потраплянням повідомлень у спам-фільтри. У цифровому маркетингу, де змінюваність ринкових умов є постійною і фактори ризику часто не піддаються однозначному прогнозуванню, метод Монте-Карло допомагає компаніям оцінити потенційний вплив різних подій на результати їхніх кампаній та визначити оптимальні стратегії для управління ризиками.

Сутність методу полягає у створенні моделей, які включають змінні з невизначеним значенням (наприклад, конверсія кліків, середній рівень відписок, вірогідність потрапляння листів у спам), які варіюються у відповідних діапазонах. Задаються вхідні параметри на основі історичних даних, після чого комп'ютерні симуляції генерують випадкові значення для кожної з цих змінних, що імітує можливі результати реальних кампаній. Це дозволяє врахувати широкий спектр результатів, які допомагають маркетологам оцінити ймовірність виникнення різних сценаріїв.

Процес застосування методу Монте-Карло для аналізу маркетингових ризиків ми розділили на декілька етапів.

Перший етап полягає в тому, що необхідно зібрати коректні дані, що відображають реальні умови та специфіку компанії. Це включає збір інформації про минулі кампанії, зокрема, про рівень конверсії, клікабельність, середній рівень відписок та інші метрики, які можуть впливати на успіх або провал кампаній. Також необхідно зібрати дані про зовнішні фактори, такі як зміни у політиці платформ (Facebook, Google тощо), законодавчі зміни (наприклад, GDPR), що можуть призвести до додаткових ризиків.

Попередній аналіз зібраних даних дозволяє визначити, які з метрик є найбільш критичними для конкретного бізнесу і потребують моделювання з урахуванням різноманітних варіантів. Аналіз даних забезпечить базу для подальшої розробки сценаріїв, а також для визначення вхідних параметрів і розподілу ймовірностей для кожного зі сценаріїв.

Наступним етапом є розробка різних сценаріїв, які можуть відображати як стандартні, так і екстремальні випадки. На основі історичних даних, які ілюструють типові ситуації для кампаній, можна визначити діапазони показників, у межах яких зазвичай змінюються ключові параметри кампанії. Наприклад, якщо раніше кампанії компанії в середньому мали рівень відкриття електронних листів на рівні 20–25%, це може бути базовим сценарієм для симуляції, однак також можна додати сценарії з високим рівнем відкриття (30–40%) або низьким (10–15%) як крайні випадки.

На цьому етапі також важливо врахувати зовнішні фактори, які можуть вплинути на ризики. Наприклад, зміни в алгоритмах платформ соціальних мереж можуть викликати різке зниження видимості реклами, тому подібний фактор також має бути передбачений в одному із сценаріїв. Історичні дані дозволяють визначити типові параметри кожного сценарію, на основі яких моделюються випадкові змінні.

Визначивши сценарії та відповідні діапазони змінних, наступним кроком є оцінка ймовірності для кожного зі сценаріїв. Наприклад, якщо відомо, що зміни у ринковому середовищі мали значний вплив на ефективність рекламних кампаній у минулому, ймовірність подібних сценаріїв може бути оцінена на рівні

20–30%. Інші сценарії, які траплялися рідше, можуть мати нижчі значення ймовірності, однак навіть їхнє включення до моделі є важливим, щоб врахувати можливі крайні випадки.

Однією з найцінніших можливостей, яку надає метод Монте-Карло, є оптимізація розподілу бюджетів та інших ресурсів у межах кампанії на основі результатів симуляції. Виходячи з отриманих даних, можна визначити оптимальні вкладення у кожний із напрямів кампанії, щоб мінімізувати ймовірні ризики та максимізувати рентабельність інвестицій. Наприклад, якщо симуляція показує, що високий рівень витрат на певний канал може призвести до надмірного ризику (скажімо, через низьку конверсію або високу частоту відписок), компанія може зменшити інвестиції в цей канал і перенаправити ресурси на більш надійні канали з нижчим рівнем ризику.

Симуляції також дозволяють порівняти різні варіанти розподілу ресурсів для досягнення цілей кампанії з урахуванням варіативності ринкових умов. Це дозволяє компаніям створювати оптимальні стратегії з врахуванням конкретних ризикових факторів, що є критично важливим у цифровому маркетингу, де швидка зміна умов є звичною. Оптимізація на основі методу Монте-Карло дозволяє компаніям забезпечити більш стабільний розвиток навіть у випадках, коли зовнішні обставини змінюються несприятливо.

Використання методу Монте-Карло для прогнозування ризиків у маркетингових кампаніях є важливим кроком для компаній, які прагнуть підвищити стійкість і результативність своїх заходів у цифровому середовищі.

В свою чергу, щоб ефективно використовувати штучні нейронні мережі для аналізу та мінімізації ризиків у маркетингових кампаніях, також необхідно пройти кілька ключових етапів, які допомагають оптимізувати точність та адаптивність моделі.

Штучні нейронні мережі є особливо ефективними, коли працюють із великими масивами даних, тому підготовка якісного набору даних є критично важливою для точності моделі. У маркетингових кампаніях це може включати різні типи даних: поведінкові характеристики користувачів, дані про продажі,

взаємодії у соціальних мережах, показники клікабельності, рівні відкриттів електронних листів тощо. Цей широкий набір дозволяє моделі враховувати різноманітні аспекти поведінки клієнтів і зовнішні фактори, що впливають на кампанії.

Дані можуть бути отримані з різних джерел. Наприклад, дані з веб-сайтів, платформ соціальних мереж, систем керування взаємодією з клієнтами (CRM) та інших джерел, щоб сформувати комплексну картину про взаємодію клієнтів. Зібрані дані мають бути стандартизованими та очищеними для забезпечення коректного аналізу. Використання методів обробки даних для видалення аномалій та відновлення відсутніх значень є дуже важливим для забезпечення точніших прогнозів. Завжди варто пам'ятати про регулярне оновлення даних, особливо в умовах швидкоплинного ринку, де переваги користувачів можуть змінюватися. Це підвищить актуальність прогнозів та точність моделі.

Штучні нейронні мережі мають здатність навчатися на основі минулих даних і прогнозувати поведінку клієнтів, що дозволяє знижувати ризики, пов'язані з неефективністю маркетингових кампаній. Для того щоб ефективно використати цю можливість моделі, потрібно використовувати різні історичних даних, щоб визначити, які події чи взаємодії призводили до певної поведінки клієнтів у минулому (наприклад, підвищення інтересу чи відтоку).

На основі аналізу ANN можна визначити, які типи контенту викликають найбільшу зацікавленість серед клієнтів та сприяють високій конверсії. Наприклад, моделі можуть показати, що відеоконтент або статті мають більший ефект у певних сегментах аудиторії, що дозволяє уникнути зайвих витрат на менш ефективні формати.

Варто звернути увагу на те, що ANN є адаптивними, тобто здатні швидко налаштовуватися на нові патерни у поведінці користувачів, які можуть виникати через зміни на ринку або в уподобаннях споживачів. Важливо не забувати про використання оновлених нейронних мереж, які можуть навчатися на нових даних у реальному часі, зберігаючи при цьому свою точність. Це забезпечує адаптацію до останніх змін у поведінці клієнтів, таких як зміни вподобань чи нові тренди.

ANN здатні враховувати не лише поведінкові характеристики клієнтів, але й зовнішні фактори, наприклад, сезонність, зміни економічних умов чи дії конкурентів, що дозволяє адаптувати стратегію під вплив нових ринкових умов.

Персоналізація є ключовим елементом для утримання клієнтів і зниження ризиків відтоку. ANN дозволяють аналізувати різні клієнтські сегменти та створювати індивідуальні пропозиції на основі детального аналізу поведінки, уподобань і минулих взаємодій. Це підвищує рівень залученості клієнтів і їхню зацікавленість продуктом, що зменшує ризик відтоку.

Можна використовувати ANN для аналізу поведінкових та демографічних характеристик різних сегментів клієнтів, щоб адаптувати контент і пропозиції під конкретні групи. Також варто залучайте нейронні мережі для автоматичного створення персоналізованих повідомлень, що відповідають інтересам та поведінковим характеристикам клієнтів.

Штучні нейронні мережі є важливим інструментом для глибинного аналізу маркетингових ризиків завдяки їхній здатності обробляти великі обсяги даних і враховувати складні залежності між змінними. Використання ANN дозволяє компаніям прогнозувати поведінку клієнтів, адаптуватися до нових ринкових умов, зменшувати ризик відтоку клієнтів і забезпечувати більш точну персоналізацію, що в цілому сприяє зниженню ризиків у маркетингових кампаніях.

Модель Random Forest дозволяє виявити ризикові чинники, які можуть впливати на ефективність маркетингових кампаній, ще до їхнього запуску. Це особливо корисно для раннього прогнозування можливих негативних наслідків, таких як низька клікабельність або високий рівень відписок.

Основними рекомендаціями до застосування є:

- Використання історичних даних про минулі кампанії для навчання моделі.
- Визначення показників які мають високий ризик отримання низької ефективності, на основі результатів моделі. Якщо, наприклад, кампанії, запуснені

в певні дні тижня, мають гіршу продуктивність, то варто розглянути інший графік запуску.

Сегментація аудиторії за допомогою Random Forest дозволить визначити групи споживачів із високим або низьким ризиком залученості, що допомагає уникнути негативного впливу невдалих стратегій на загальні показники кампаній. Модель дозволяє розділити клієнтів за демографічними або поведінковими характеристиками, що полегшує персоналізацію підходу для кожного сегменту.

Варто використовувати модель для аналізу взаємодії користувачів, щоб визначити групи з різним рівнем ризику. Наприклад, сегментувати аудиторію на тих, хто частіше взаємодіє з контентом, та на тих, хто має низьку активність.

Визначивши сегменти з низьким рівнем залученості, можна налаштувати більш цільовий контент, що підвищить ймовірність взаємодії з кампанією. Наприклад, пропонуйте спеціальні пропозиції або акції для груп з високим ризиком відтоку.

Random Forest автоматично оцінює важливість кожного фактора для кінцевого результату, що дозволяє зосередитися на найбільш значущих аспектах кампанії, які мають найбільший вплив на успіх. Це допомагає мінімізувати ризики, спрямовуючи зусилля на оптимізацію найбільш критичних факторів.

На основі оцінки важливості змінних моделі, варто зосередитись на тих факторах, які мають найбільший вплив на результати. Якщо модель показує, що час дня чи тип рекламного контенту є важливими для взаємодії користувачів, варто зосередити ресурси на оптимізації саме цих елементів.

Якщо модель показує, що певний контент чи формат сприяє підвищенню рівня взаємодії, можна впровадити більше подібних форматів у свої стратегії. Навпаки, для факторів з низьким рівнем важливості можна скоротити витрати або ресурси.

Один із найважливіших ризиків, з яким стикаються маркетингові команди, — це відтік клієнтів. Модель Random Forest може бути використана для передбачення ризику втрати клієнтів і допомагає розробити стратегії для

утримання. Вона дозволяє визначити ознаки, що вказують на потенційну втрату інтересу, і вжити заходів до того, як клієнт припинить взаємодію.

Спочатку потрібно проаналізувати, які змінні передбачають відтік, такі як зниження частоти відвідувань або менша активність у відповідь на розсилку. Потім, для користувачів з високим ризиком відтоку, можна створити спеціальні пропозиції, наприклад, персоналізовані знижки або бонуси, щоб стимулювати їхню активність і повернути інтерес до бренду.

Random Forest є гнучкою моделлю, яка дозволяє проводити регулярний аналіз і швидко адаптуватися до змінних умов ринку. Це особливо важливо для цифрового маркетингу, де зміни в уподобаннях споживачів і конкурентному середовищі можуть відбуватися дуже швидко.

Random Forest може бути регулярно оновлена новими даними, що забезпечить актуальність прогнозів і дозволить компанії адаптуватися до останніх змін у поведінці споживачів.

Для ефективного управління ризиками у маркетингових кампаніях необхідна система постійного моніторингу та оновлення ризикових факторів. Така система дозволить швидко реагувати на зміни у ринковому середовищі, поведінці споживачів, а також у конкурентному та регуляторному оточенні, що є особливо актуальним у сучасному цифровому середовищі. Постійний моніторинг забезпечить актуальність моделей прогнозування та дозволяє знижувати ймовірність негативних наслідків за рахунок своєчасного виявлення нових загроз.

Система моніторингу ризикових факторів повинна базуватися на принципах гнучкості та адаптивності, щоб враховувати постійні зміни у ринкових умовах та забезпечувати максимально точні прогнози. Зокрема, така система передбачає постійне оновлення даних та регулярну переоцінку ризиків. Це можливо завдяки аналітичним інструментам, які використовують поточні дані, включаючи дані з соціальних мереж, веб-трафіку, показники ефективності кампаній та іншу інформацію, що може впливати на ринкову ситуацію.

Для того щоб моделі прогнозування ризиків залишалися актуальними та надавали корисну інформацію, їх слід регулярно оновлювати на основі нових даних. В умовах цифрового маркетингу поведінка споживачів може змінюватися швидко, тому важливо регулярно здійснювати повторне навчання моделей, щоб враховувати зміни у вподобаннях та очікуваннях клієнтів.

Можна зв'язати моделі з платформами, що моніторять показники клієнтської активності, такі як взаємодія з контентом, рівень відписок, зміни у частоті відвідувань, що допоможе виявляти ранні сигнали можливих ризиків.

Не варто забувати, що ринок цифрового маркетингу піддається впливу регуляторних змін, таких як нові вимоги до захисту даних чи рекламні обмеження, що можуть впливати на ризики кампаній. Система моніторингу повинна включати аналіз змін у законодавстві, щоб прогнозувати вплив цих змін на маркетингову діяльність.

Аналіз дій конкурентів, такі як нові стратегії чи кампанії, що можуть вплинути на аудиторію, є вкрай важливим. Це допоможе своєчасно відреагувати на конкуренцію і скоригувати свою маркетингову стратегію.

На основі результатів постійного моніторингу та виявлення нових ризиків слід адаптувати маркетингову стратегію, що дозволить підвищити ефективність кампаній та зменшити вплив ризикових факторів. Гнучка стратегія, яка адаптується до змінних умов, є критичною для успішного управління ризиками у маркетингових кампаніях.

Якщо моніторинг показує зменшення інтересу до певного контенту або каналу комунікації, стратегія має бути швидко скоригована. На основі аналізу ризиків, що надходять із системи моніторингу, можна розробити адаптивний бюджет для різних маркетингових каналів, щоб у разі виявлення критичних загроз зменшити витрати на менш ефективні канали.

Система постійного моніторингу та оновлення ризикових факторів є необхідним елементом для ефективного управління ризиками у маркетингових кампаніях. Вона дозволяє забезпечити актуальність моделей, своєчасно реагувати на нові загрози та адаптувати маркетингові стратегії відповідно до

змінних ринкових умов. Завдяки такій системі компанії можуть підвищити ефективність своїх кампаній, знизити ризики відтоку клієнтів та забезпечити довгострокову стабільність маркетингової діяльності.

3.3. Оцінка економічної ефективності запропонованих заходів

Персоналізація та сегментація є одними з найефективніших інструментів у сучасному email-маркетингу, які дозволяють компаніям більш точно орієнтуватися на потреби та вподобання клієнтів, тим самим підвищуючи ефективність своїх розсилок. Однак впровадження таких інструментів передбачає певні витрати, пов'язані як із використанням технологій машинного навчання, так і з обробкою великих обсягів даних для досягнення високої точності в прогнозах. Зокрема, моделі, такі як Random Forest та штучна нейронна мережа (ANN), потребують ретельного налаштування та обчислювальних ресурсів для обробки даних, що є важливим аспектом під час планування бюджету на персоналізацію.

Основні витрати на впровадження персоналізації та сегментації можна розділити на декілька категорій. Перша з них — це витрати на обчислювальні ресурси та програмне забезпечення. Моделі, які використовуються для персоналізації, потребують значних ресурсів для обробки та аналізу великих обсягів даних клієнтів. Наприклад, використання Random Forest передбачає багатоваріантну обробку даних, що вимагає серверної інфраструктури з високою потужністю. Крім того, для навчання ANN необхідні спеціалізовані програмні середовища, які підтримують складні моделі нейронних мереж, що також потребує витрат на ліцензії та підписки.

Друга категорія витрат стосується підготовки та очищення даних для забезпечення точності моделей. Для досягнення оптимальної точності прогнозів необхідно мати якісні та повні дані про клієнтів, що потребує постійної підтримки та оновлення бази даних. Вартість підготовки даних включає збір, очищення та структурування інформації для коректного функціонування

алгоритмів. Наприклад, для використання ANN необхідно забезпечити стандартизовані та належним чином оброблені дані, що дозволить моделі точніше визначати індивідуальні потреби клієнтів та уникати помилок у класифікації.

Третім важливим елементом витрат є навчання та підтримка персоналу. Для налаштування і підтримки роботи моделей потрібні кваліфіковані спеціалісти з машинного навчання та аналізу даних, які здатні забезпечити коректне функціонування моделей, виявляти та коригувати помилки, а також адаптувати алгоритми відповідно до нових даних. Крім того, необхідно навчити маркетингові команди працювати з результатами моделювання та застосовувати рекомендації моделей у реальних кампаніях, що включає навчання використанню аналітичних інструментів і інтерпретації даних.

Витрати на обчислювальні ресурси, підготовку даних та навчання персоналу є інвестиціями, які забезпечують підвищення точності та ефективності маркетингових кампаній. Оптимальна персоналізація дозволяє компаніям значно знизити ризик відписок, підвищити залученість аудиторії та, зрештою, покращити показники доходів, що робить ці інвестиції економічно вигідними у довгостроковій перспективі.

Оптимізація таргетингу та частоти розсилок є важливими елементами, що впливають на фінансову ефективність email-маркетингових кампаній. Чітке розуміння потреб і поведінкових патернів аудиторії дозволяє компаніям не лише підвищувати показники залученості, але й збільшувати доходи, оскільки персоналізовані розсилки, відправлені в оптимальний час, значно покращують ймовірність взаємодії клієнтів. За допомогою моделей, таких як Random Forest та штучна нейронна мережа (ANN), можна налаштувати більш точний таргетинг та адаптувати частоту розсилок для різних груп користувачів, що позитивно впливає на доходи та зменшує ризики негативної реакції.

Одним із основних аспектів підвищення доходів є точний таргетинг, що дозволяє компаніям орієнтуватися на найбільш перспективні сегменти клієнтів. Використовуючи Random Forest для визначення ключових ознак, компанії

можуть виділяти найактивніші групи клієнтів і орієнтувати на них свої розсилки. Цей підхід допомагає уникати марної витрати ресурсів на менш зацікавлені сегменти та зосереджуватися на клієнтах, які, ймовірно, дадуть найбільший відгук. Крім того, ANN може деталізувати таргетинг, аналізуючи більш складні взаємозв'язки у даних, що дозволяє створювати персоналізований контент для кожної групи клієнтів і підвищувати їхню зацікавленість у розсилці.

Оптимізація частоти розсилок також має значний економічний ефект, адже вона впливає на загальну задоволеність клієнтів і, відповідно, на рівень доходів. Надмірно часті повідомлення можуть призвести до відписок, тоді як рідкі розсилки можуть знизити залученість і потенційні конверсії. За допомогою аналізу на основі Монте-Карло можна змодельовати різні сценарії частоти розсилок і визначити оптимальну частоту для кожного сегмента аудиторії. Наприклад, для активних клієнтів частіші розсилки можуть бути прийнятними і навіть бажаними, тоді як для менш активних варто зменшити частоту, щоб уникнути перенасичення контентом. Це дозволяє компаніям уникати зайвих витрат на відправлення нерелевантних повідомлень і забезпечує кращий результат за меншого обсягу розсилок.

Оптимізований таргетинг і частота розсилок також впливають на доходи від підвищення Open Rate та CTR (Click-Through Rate), оскільки клієнти, які отримують повідомлення у зручний час і зацікавлені у контенті, частіше взаємодіють із розсилкою. Вищий рівень відкриттів та переходів сприяє збільшенню конверсій, що напряду впливає на доходи. Наприклад, оптимізація часу надсилання листів може призвести до суттєвого зростання Open Rate, оскільки повідомлення приходять саме тоді, коли клієнти мають більше часу для перегляду. Крім того, підвищений CTR означає більшу ймовірність того, що клієнти перейдуть на вебсайт чи сторінку товару, що може призвести до збільшення продажів.

Використання моделей для аналізу та налаштування розсилок дозволяє не тільки знизити ризики відписок, але й сприяє зростанню доходів завдяки підвищенню рівня залученості клієнтів. Це робить оптимізацію таргетингу та

частоти розсилок економічно вигідною стратегією для досягнення довгострокової стабільності та зростання прибутковості.

Зниження ризику відписок є важливим компонентом економічної ефективності email-маркетингових кампаній, оскільки утримання наявних клієнтів значно дешевше, ніж залучення нових. Успішне зниження ризику відписок за допомогою персоналізованих розсилок, адаптованої частоти та релевантного контенту позитивно впливає на економічні показники компанії, адже лояльні клієнти мають вищий показник Lifetime Value (LTV), що дозволяє підвищити прибутковість компанії у довгостроковій перспективі. Використання моделей, таких як Random Forest та штучна нейронна мережа (ANN), для аналізу ризиків та розробки стратегії утримання клієнтів дозволяє досягти високої точності в прогнозах відписок і зменшити їх кількість [1].

Економічний ефект від зниження відписок можна оцінити, виходячи з економії витрат на залучення нових клієнтів. Коли клієнти залишаються лояльними та продовжують взаємодіяти з розсилками, компанія уникає додаткових витрат на маркетинг, спрямований на заміну відписаних підписників. За допомогою моделі Random Forest можна ідентифікувати клієнтів, які схильні до відписок, та сегментувати їх за ступенем ризику. Наприклад, якщо модель вказує на підвищений ризик серед певної групи клієнтів через частоту розсилок або невідповідність контенту, компанія може змінити підхід для цієї аудиторії, знижуючи частоту або адаптуючи контент. Це дозволяє уникнути зайвих витрат на заміну клієнтів і забезпечує стабільність у взаємодії з підписниками.

Ще одним важливим аспектом економічного ефекту є зростання показника Lifetime Value (LTV) клієнтів завдяки зниженню відписок. Лояльні клієнти, які тривалий час взаємодіють з брендом, приносять більше доходів протягом свого життєвого циклу. ANN дозволяє розпізнати більш делікатні поведінкові ознаки, які вказують на готовність клієнта залишитися з брендом, і допомагає компанії будувати довготривалі стосунки з клієнтами. Завдяки адаптованим персоналізованим повідомленням компанія підвищує ймовірність того, що

клієнти залишаються активними та будуть продовжувати здійснювати покупки, що призводить до зростання LTV.

Зниження ризику відписок також забезпечує економічну вигоду через підвищення довіри та позитивного іміджу бренду. Коли клієнти отримують релевантні, своєчасні та адаптовані під їхні потреби розсилки, вони відчують, що бренд дбає про їхні інтереси. Це знижує ймовірність того, що клієнти сприйматимуть повідомлення як нав'язливі або неактуальні, що позитивно позначається на їхній лояльності до бренду. Лояльні клієнти частіше рекомендують бренд іншим, що додатково знижує витрати на маркетинг і забезпечує природне зростання клієнтської бази.

Таким чином, розрахунок економічного ефекту від зниження ризиків відписок свідчить про значну вигоду від інвестицій у персоналізацію та адаптацію контенту. Утримання клієнтів допомагає зберегти стабільність доходів, зменшити витрати на залучення нових клієнтів та підвищити довгострокову економічну ефективність завдяки зростанню показника LTV. Стратегії, спрямовані на зниження ризиків відписок, є економічно обґрунтованими і важливими для довгострокового розвитку бізнесу.

Запобігання потраплянню розсилок у спам є одним із ключових факторів, які впливають на успіх email-маркетингових кампаній та їхню економічну ефективність. Коли повідомлення потрапляє в спам-папку, воно не тільки втрачає можливість привернути увагу клієнта, але й може негативно вплинути на репутацію компанії. Таким чином, заходи, спрямовані на запобігання потраплянню в спам, є інвестиціями, які сприяють збереженню високого рівня залученості, підвищенню довіри клієнтів до бренду та довгостроковому зростанню доходів.

Одним з основних методів, що допомагають запобігти потраплянню листів у спам, є аналіз факторів, які впливають на сприйняття повідомлень як небажаних. Моделі машинного навчання, такі як Random Forest та метод Монте-Карло, дозволяють ідентифікувати ключові характеристики листів, які можуть призвести до блокування повідомлень спам-фільтрами. Наприклад, аналіз

показників, як-от частота використання певних ключових слів, кількість зображень, наявність спеціальних символів у темі, дозволяє моделі визначати фактори, що підвищують ризик блокування. На основі цих даних можна коригувати стиль написання листів, уникаючи таких елементів, які можуть бути позначені як спам.

Економічний ефект від запобігання спаму полягає в тому, що повідомлення мають більше шансів потрапити в основну поштову скриньку клієнта, що збільшує Open Rate та CTR (Click-Through Rate) кампанії. Це сприяє зростанню конверсій та прямих продажів, оскільки клієнти частіше відкривають та переглядають повідомлення, які вчасно надходять у їхній інбокс. Наприклад, якщо коригування тем листів і змісту знижує ризик спам-фільтрації на 10%, компанія може отримати суттєве зростання показників взаємодії, що безпосередньо впливає на дохідність кампанії.

Ще однією важливою складовою економічного ефекту від запобігання спаму є збереження довіри та лояльності клієнтів. Коли клієнти бачать, що повідомлення від бренду регулярно надходять в інбокс і містять корисну інформацію, це підвищує їхню довіру до компанії. Відсутність випадків потрапляння в спам-папку свідчить про професійний підхід до комунікації та дотримання стандартів якості, що зміцнює репутацію бренду. Лояльні клієнти, які регулярно отримують релевантний контент, більш схильні до повторних покупок і рідше відписуються, що підвищує Lifetime Value (LTV) клієнтів і знижує витрати на залучення нових.

Крім того, запобігання потраплянню в спам має додаткову економічну вигоду у вигляді зниження витрат на виправлення репутаційних збитків. Потрапляння в спам може негативно позначитися на показниках репутації відправника, що ускладнює доставку майбутніх листів навіть до тих клієнтів, які активно взаємодіють з контентом. Інвестиції у відповідні заходи дозволяють зменшити ризик негативного впливу на репутацію та уникнути витрат на відновлення цього показника, зберігаючи високу ефективність кампаній у довгостроковій перспективі.

Таким чином, економічний ефект від запобігання потраплянню листів у спам значний, оскільки такі заходи сприяють зростанню рівня взаємодії з клієнтами, збільшенню продажів, підвищенню лояльності та збереженню репутації компанії. Інвестиції у профілактику спам-фільтрації не лише покращують показники розсилок, але й забезпечують довгострокову економічну вигоду, яка є важливою для сталого розвитку та конкурентоспроможності бізнесу.

Окупність інвестицій (ROI) є одним з основних показників, що визначають економічну доцільність маркетингових заходів, таких як адаптація контенту та тригерні повідомлення в email-кампаніях. Впровадження персоналізованого контенту і налаштування автоматизованих тригерів дозволяє компаніям не лише зберегти залученість клієнтів, але й сприяє збільшенню доходів, оскільки персоналізація сприяє створенню більш тісного зв'язку з аудиторією. Точний розрахунок ROI допомагає оцінити, наскільки вигідними є ці інвестиції у довгостроковій перспективі.

Одним із основних факторів, що впливає на ROI, є рівень залученості клієнтів, який збільшується завдяки персоналізованим повідомленням, орієнтованим на інтереси та потреби кожного сегмента. Використання моделей Random Forest та штучної нейронної мережі (ANN) дозволяє адаптувати контент таким чином, щоб він був максимально релевантним для кожного клієнта. Наприклад, для клієнтів, які раніше проявляли зацікавленість певними продуктами, система може автоматично налаштовувати повідомлення з відповідними рекомендаціями. Такі підходи суттєво підвищують Open Rate і CTR, що на пряму впливає на конверсії і, відповідно, на доходи від кампаній.

Тригерні повідомлення є ще одним ефективним інструментом для підвищення ROI, оскільки вони забезпечують миттєву реакцію на дії або інтереси клієнтів. Наприклад, коли клієнт залишає товар у кошику без оформлення замовлення, тригерне повідомлення може нагадати про покинутий товар або запропонувати спеціальну знижку. Такі повідомлення значно підвищують імовірність завершення покупки, що безпосередньо впливає на доходи. Крім того, тригери допомагають активувати клієнтів, які раніше не

взаємодіяли з брендом, стимулюючи їх до повторних покупок або підписки на додаткові послуги. Вплив тригерних повідомлень на доходи можна виміряти через показник збільшення конверсій та середнього чека, що дозволяє точно розрахувати ROI від таких заходів.

Розрахунок ROI для адаптації контенту та тригерів враховує як додаткові доходи від підвищення залученості, так і зниження витрат на утримання клієнтів. Наприклад, персоналізовані повідомлення знижують імовірність відписок, оскільки клієнти отримують релевантний контент, що відповідає їхнім очікуванням. Це дозволяє компанії зберігати клієнтську базу та зменшувати витрати на залучення нових користувачів. Зростання показника Lifetime Value (LTV) клієнтів, яке спостерігається завдяки використанню персоналізованих тригерів і контенту, також сприяє підвищенню ROI. Клієнти, які регулярно отримують цікаві та своєчасні пропозиції, частіше здійснюють повторні покупки, що забезпечує компанії стабільний дохід.

Додатковим економічним ефектом від впровадження адаптованого контенту є зниження витрат на утримання клієнтів. Лояльні клієнти рідше відписуються від розсилок і активніше взаємодіють із брендом, що дозволяє знижувати витрати на нові маркетингові заходи. Економія на залученні нових клієнтів, у поєднанні з підвищенням доходів від постійних клієнтів, підвищує загальну окупність інвестицій у персоналізацію.

Таким чином, розрахунок ROI для адаптації контенту та тригерних повідомлень підтверджує, що ці заходи мають високу економічну доцільність. Вони сприяють не тільки зростанню доходів завдяки збільшенню залученості та конверсій, але й знижують витрати на утримання клієнтів, що є важливим для довгострокового успіху та стабільного розвитку компанії.

Клієнтська лояльність є однією з найважливіших складових для досягнення довгострокової економічної ефективності компанії. Лояльні клієнти не тільки здійснюють повторні покупки, але й сприяють позитивному іміджу бренду, що допомагає компанії залучати нових споживачів. Впровадження заходів, спрямованих на підвищення лояльності за допомогою персоналізації,

адаптованого контенту та оптимальної частоти розсилок, сприяє зростанню показника Lifetime Value (LTV) клієнтів, що є основним джерелом довгострокового доходу.

Персоналізовані розсилки дозволяють створювати більш тісні стосунки з клієнтами, оскільки вони показують, що бренд розуміє їхні потреби та інтереси. Використання таких інструментів, як Random Forest та штучна нейронна мережа (ANN), допомагає сегментувати аудиторію та налаштовувати розсилки так, щоб вони були максимально релевантними для кожного клієнта. Наприклад, клієнти, які регулярно отримують персоналізовані пропозиції, стають більш лояльними до бренду, оскільки відчують, що компанія дбає про їхні інтереси. Це знижує ризик відписок і створює основу для довгострокового співробітництва з клієнтами, що позитивно позначається на економічних показниках компанії.

Одним із головних економічних ефектів підвищення клієнтської лояльності є збільшення показника LTV. Лояльні клієнти, які часто взаємодіють із брендом, забезпечують стабільний дохід протягом усього життєвого циклу співпраці з компанією. Клієнти з високим LTV здійснюють більше покупок і рідше відписуються, що зменшує витрати на постійне залучення нових клієнтів. Таким чином, підвищення показника LTV дозволяє компанії досягати кращих економічних результатів за рахунок зниження маркетингових витрат і зростання доходів від повторних покупок. Дослідження показують, що витрати на залучення нового клієнта можуть у 5-7 разів перевищувати витрати на утримання існуючого, що підкреслює важливість інвестицій у лояльність.

Крім того, лояльні клієнти часто виступають як неформальні "посли" бренду, рекомендуючи його своїм друзям та знайомим. Така рекомендаційна реклама є дуже ефективною і практично не потребує додаткових витрат з боку компанії. Клієнти, задоволені сервісом та персоналізованими пропозиціями, охоче діляться своїм позитивним досвідом, що сприяє природному розширенню клієнтської бази. Це, своєю чергою, забезпечує економічну вигоду від підвищеної клієнтської лояльності, оскільки компанія отримує нових клієнтів без необхідності витрачати додаткові ресурси на їх залучення.

Ефект від покращеної лояльності також проявляється у стабільності доходів компанії, що є важливим фактором для довгострокового планування та прогнозування. Компанії, які підтримують високий рівень клієнтської лояльності, можуть більш точно передбачати майбутні доходи, що спрощує планування та дозволяє робити інвестиції у розвиток продуктів або покращення сервісу. Це створює фінансову стабільність і дозволяє компанії краще адаптуватися до змін ринкових умов, оскільки вона має сталу базу лояльних клієнтів, яка не залежить від зовнішніх факторів.

Отже, покращена клієнтська лояльність має значний довгостроковий економічний ефект, забезпечуючи стабільність доходів, підвищення показника LTV та зниження витрат на залучення нових клієнтів. Інвестиції в персоналізацію та адаптований контент виправдовують себе завдяки зростанню лояльності, що забезпечує компанії конкурентну перевагу і стабільний розвиток у довгостроковій перспективі.

Економічна доцільність заходів, спрямованих на оптимізацію email-маркетингових кампаній, є ключовим фактором, який визначає їхню ефективність та вигідність для компанії у довгостроковій перспективі. Ретельний аналіз впливу персоналізації, адаптації контенту, частоти та часу відправлення розсилок, а також заходів для запобігання ризикам потрапляння у спам і відписок, показує, що такі інвестиції мають високий економічний ефект. Впровадження моделей машинного навчання, таких як Random Forest, штучна нейронна мережа (ANN) та метод Монте-Карло, дозволяє компаніям підвищити точність таргетингу, покращити релевантність контенту і збільшити показники залученості клієнтів, що, в свою чергу, позитивно впливає на доходи та репутацію компанії.

ВИСНОВКИ

У ході цього дослідження було розроблено та впроваджено ефективні моделі для оцінки ризиків у маркетингових кампаніях, які базуються на використанні сучасних математичних моделей і методів. Основною метою роботи було створення комплексного підходу для прогнозування ризиків, таких як потрапляння листів у спам-фільтри та ризик відписка, з використанням моделей Random Forest, штучних нейронних мереж (ANN) та симуляцій методом Монте-Карло.

Під час дослідження було досягнуто поставлені завдання, що включали:

- Проведення теоретичного огляду методів аналізу ризиків у маркетингових кампаніях.
- Застосування моделей Random Forest та ANN для моделювання і прогнозування ризиків, визначення точності їхніх результатів.
- Використання методу Монте-Карло для моделювання можливих сценаріїв ризиків відписка та аналізу їхньої ймовірності.
- Порівняння результатів моделювання з фактичними даними, що дозволило оцінити точність і надійність моделей.

Узагальнені результати роботи свідчать про те, що використання комбінованих підходів на базі моделей машинного навчання та ймовірнісних методів підвищує точність оцінки ризиків та ефективність управління маркетинговими стратегіями. Проведений аналіз підтвердив доцільність використання цих підходів для прогнозування та зниження ризиків у email-маркетингу, що є важливим внеском у розвиток сучасних практик управління маркетинговими кампаніями.

Random Forest показала стабільність і високу точність для частих класів, але мала труднощі з рідкісними випадками через дисбаланс даних. ANN добре розпізнавала складні зв'язки, але виявила чутливість до дисбалансу, що знижувало її ефективність. Комбінована модель об'єднала переваги обох методів, підвищивши точність прогнозів і зменшивши помилки класифікації.

Метод Монте-Карло збагатив аналіз, моделюючи різні сценарії та підвищуючи реалістичність прогнозів. Ефективність моделей підтвердилася їхньою здатністю надавати точні та надійні прогнози ризиків, що слугує основою для вдосконалення маркетингових стратегій.

Дослідження підтвердило, що комбіновані моделі та метод Монте-Карло є ефективними для оцінки ризиків у email-маркетингу, створюючи основу для гнучких маркетингових стратегій і мінімізації ризиків втрати підписників.

На основі дослідження розроблено практичні рекомендації для підвищення ефективності email-маркетингу та зменшення ризиків. Застосування розроблених моделей значно підвищує точність прогнозів і зменшує ймовірність негативних наслідків, як-от потрапляння у спам і відписки.

Оцінка економічної ефективності впровадження розроблених моделей у email-маркетинг показала суттєві вигоди. Використання моделей Random Forest, ANN та методу Монте-Карло підвищує точність прогнозів, що дозволяє оптимізувати ресурси і мінімізувати ризики.

Впровадження цих моделей оптимізує витрати, підвищує ефективність і доходи, що підтверджує економічну доцільність інвестицій у сучасні аналітичні підходи для підтримки конкурентоспроможності.

У результаті проведеного дослідження вдалося досягти основної мети – розробити та протестувати комплексну методологію для оцінки ризиків у email-маркетингових кампаніях, що базується на сучасних математичних моделях і методах аналізу даних. Завдання, поставлені на початку роботи, були виконані, що підтверджує успішну реалізацію мети дослідження.

Під час дослідження були також розроблені практичні рекомендації для застосування результатів моделювання, що сприяють мінімізації ризиків та підвищенню ефективності email-маркетингових стратегій. Оцінка економічної ефективності продемонструвала, що впровадження розроблених методів дозволяє оптимізувати маркетингові витрати та підвищити рентабельність кампаній.

Перспективи розвитку цієї теми передбачають розширення обсягів даних, інтеграцію додаткових методів машинного навчання та розробку автоматизованих платформ для практичного використання розроблених моделей.

Таким чином, дослідження продемонструвало свою актуальність і практичну цінність, підтвердивши, що використання сучасних математичних моделей і методів аналізу є ефективним підходом до оцінки та управління ризиками у сфері email-маркетингу. Робота зробила значний внесок у розвиток методик управління цифровими маркетинговими кампаніями, що дозволяє підприємствам краще розуміти своїх клієнтів та підвищувати ефективність взаємодії з ними.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Бхарат Раджан. Довічна цінність клієнта Що, як і чому / Бхарат Раджан. // BookThe Routledge Companion to Strategic Marketing. – 2020 рік.
2. Всебічне опитування щодо інтелектуального виявлення спаму електронної пошти / Бхаранідаран Шанмугам, Крішнан Канноорпатті, Асіф Карім, Самі Азам. – 2019 рік.
3. Джойіта Гош. Оптимізатор ADAM і КАТЕГОРИЧНА КРОСЕНТРОПІЯ Метод ANN на основі функції втрат / Джоїта Гош. – 2023 рік.
4. Етика в цифровому маркетингу: що потрібно знати та як узгодити [Електронний ресурс] // Forbes. – 2023. – Режим доступу до ресурсу: <https://www.forbes.com/councils/forbesagencycouncil/2023/10/06/ethics-in-digital-marketing-what-to-know-and-how-to-align/>.
5. Загальний регламент захисту даних GDPR [Електронний ресурс]. – 2016. – Режим доступу до ресурсу: <https://gdpr-info.eu/>.
6. Закон Каліфорнії про конфіденційність споживачів (CCPA) [Електронний ресурс]. – 2024. – Режим доступу до ресурсу: <https://oag.ca.gov/privacy/ccpa>.
7. І. М. Корпіоя. Від даних до розуміння: моделювання методом Монте-Карло як інструмент маркетингової інформації / І. М. Корпіоя. – 2022.
8. Ісуб Ке. Обчислення індексу Джіні: Примітка / Ісуб Ке, Цзяньсі Су, Едвард Фурман. – 2019 рік.
9. Каміль Жбіковський. Прямі маркетингові кампанії в роздрібному банківському бізнесі з використанням глибокого навчання та випадкових лісів / Каміль Жбіковський, Пйотр Гаврисяк, Пьотр Ладижинський. – 2019 рік.
10. Левіна-Костюк М. МЕТОДИ ПРИЙНЯТТЯ УПРАВЛІНСЬКИХ РІШЕНЬ В УМОВАХ НЕДОСТАТНЬОЇ ІНФОРМАЦІЇ / Марія Левіна-Костюк. // Одеський державний аграрний університет. – 2022.
11. Ліначо К. Г. Огляд методів стохастичного програмування для оптимізації технологічних систем в умовах невизначеності / Кен Гросмен Ліначо. – 2020.

12. Нелінійна апроксимація та (глибокі) мережі ReLU / С. Фукарт, Б. Ханін, І. Добеші, Р. ДеВор. – 2021 рік.
13. Огляд аналізу часових рядів для аналітики кібербезпеки: від виявлення вторгнень до прогнозування атак / Б.Стоянович, Ф. Скопик, М. Ландауер, А. Флатшер. // International Journal of Information Security. – 2024.
14. ОСОБЛИВОСТІ ЦИФРОВОГО МАРКЕТИНГУ В ЕПОХУ БЛІЗЯЦІЇ: АНАЛІЗ ВИКЛИКІВ / Т.Рябова, І. Рябов, О. Вовчанська, Ч. Лі. // Проблеми теорії та практики фінансово-кредитної діяльності. – 2022. – №6. – С. 160–171.
15. Павленко, М.С. Інструменти партнерського маркетингу для малого та середнього бізнесу / Павленко, М.С., Л. М. Хоменко. – 2022.
16. Порівняння випадкового лісу та його значення Джині зі стандартними хемометричними методами для вибору ознак і класифікації спектральних даних / Ральф Мазух, Б. Майкл Кельм, Бйорн Х. Мензе, Уве Гіммельрайх. – 2009 рік.
17. Робул Ю. Структура та складові систем цифрового маркетингу / Юрій Робул. – 2020.
18. С.Б. Юсте 2006 С.Б. Юсте. Методи середньозважених скінченних різниць для рівнянь дробової дифузії / С.Б. Юсте. – 2006 рік.
19. Сабайтите Й. Моделювання дерева рішень уподобань електронних споживачів щодо інструментів інтернет-маркетингової комунікації під час перегляду / Й. Сабайтите, В. Давидавічене, Я. Стракова. // Е+М. Економіка а Management = Економіка та управління. – 2019. – С. 206–221.
20. Стокс Р. Електронний маркетинг: основний посібник із маркетингу в цифровому світі / Роб Стокс., 2014. – (Quirk eMarketing (Pty) Ltd;). – (5-й випуск).
21. Тахердуст Х. Смарт-контракти в технології блокчейн: критичний огляд / Хамед Тахердуст. // Департамент мистецтв, комунікацій і соціальних наук, Західний Канадський університет. – 2023.
22. Тенденції онлайн-маркетингу та наміри щодо купівлі: досягнення в задоволеності клієнтів за допомогою підходу PLS-SEM та ANN / K.Khondaker

Sazzadul, I. Mohammad Ekramol, I. Abdullah Mohammed, Pan, Shin-Hung. – 2023 рік.

23. Термін дії. Hidden Metrics of Email Deliverability [Електронний ресурс] / Validity. – 2016. – Режим доступу до ресурсу: <https://www.validity.com/resource-center/hidden-metrics-of-email-deliverability/>.

24. Удосконалення цифрового маркетингу криптовалютних веб-сайтів за допомогою інноваційної обробки даних клієнтів / Д.Сакас, Н. Джаннакопулос, Н. Канеллос, Х. Трифонопулос. // Лабораторія бізнес-інформаційних та комунікаційних технологій у ланцюгах створення вартості (BICTEVAC LABORATORY). – 2022.

25. Хосе Рамон Саура. Використання Data Sciences у цифровому маркетингу: структура, методи та показники ефективності / Хосе Рамон Саура. – 2021.

26. Чавлак Н. Роль великих даних у цифровому маркетингу / Н. Чавлак, Р. Коп. // Розширені стратегії цифрового маркетингу в епоху, керовану даними. – 2021.

27. Шаоань Се. Огляд технології блокчейн: архітектура, консенсус і майбутні тенденції / Шаоань Се, Сянпін Чен, Хуннінг Дай. // IEEE International Congress on Big Data (BigData Congress). – 2017.

28. Ян Л. Маркетингова розвідка та великі дані: методи цифрового маркетингу на шляху до того, щоб стати методами соціальної інженерії в маркетингу / Лайс Ян. – 2019.

ДОДАТКИ

Додаток А

Journey Topic	Journey	Client Services Segment	new_interactiontype name	new_emailidname	Country	createdon	Receipient Type
[CLIENT] INVITATION - Webinar 22 October - Trading and Custody	[CLIENT] INVITATION - Webinar 22 October - Trading and Custody	B2B Banking	Email Clicked	[CLIENT] INVITATION - Webinar 22 October - Trading and Custody	Switzerland	20.09.2024 0:00	Client
[CLIENT] INVITATION - Webinar 22 October - Trading and Custody	[CLIENT] INVITATION - Webinar 22 October - Trading and Custody	B2B Banking	Email Clicked	[CLIENT] INVITATION - Webinar 22 October - Trading and Custody	Switzerland	23.09.2024 0:00	Client
[CLIENT] INVITATION Reminder - Webinar 22 October - Trading and Custody	[CLIENT] INVITATION Reminder - Webinar 22 October - Trading and Custody	B2B Banking	Email Clicked	[CLIENT] INVITATION - Webinar 22 October - Trading and Custody	Switzerland	11.10.2024 0:00	Client
[CLIENT] INVITATION Reminder new - Webinar 22 October - Trading and Custody	[CLIENT] INVITATION Reminder new - Webinar 22 October - Trading and Custody	B2B Banking	Email Clicked	[CLIENT] INVITATION - Webinar 22 October - Trading and Custody	Switzerland	18.10.2024 0:00	Client
[THANK YOU] Future Finance Evening 2024	[THANK YOU] Future Finance Evening 2024	B2B Banking	Email Clicked	[THANK YOU] Future Finance Evening 2024	Switzerland	24.09.2024 0:00	Client

[THANK YOU] Future Finance Evening 2024	[THANK YOU] Future Finance Evening 2024	B2B Banking	Email Clicked	[THANK YOU] Future Finance Evening 2024	Switzerland	24.09.2024 0:00	Client
[THANK YOU] Future Finance Evening 2024 - non registered guests	[THANK YOU] Future Finance Evening 2024 - non registered guests	B2B Banking	Email Clicked	[THANK YOU] Future Finance Evening 2024	Switzerland	26.09.2024 0:00	Client
Crypto Market Update	Copy of Crypto Market Update 20240918	B2B Banking	Email Clicked	Crypto Market Update 20240918	Switzerland	18.09.2024 0:00	Client
Crypto Market Update	Copy of Crypto Market Update 20240918	B2B Banking	Email Clicked	Crypto Market Update 20240918	Switzerland	18.09.2024 0:00	Client
Crypto Market Update	Copy of Crypto Market Update 20240918	B2B Banking	Email Clicked	Crypto Market Update 20240918	Switzerland	18.09.2024 0:00	Client
Crypto Market Update	Crypto Market Update 20240731	B2B Banking	Email Clicked	Crypto Market Update 20240731	Switzerland	31.07.2024 0:00	Client
Crypto Market Update	Crypto Market Update 20240731	B2B Banking	Email Clicked	Crypto Market Update 20240731	Switzerland	31.07.2024 0:00	Client
Crypto Market Update	Crypto Market Update 20240731	B2B Banking	Email Clicked	Crypto Market Update 20240731	Switzerland	31.07.2024 0:00	Client
Crypto Market Update	Crypto Market Update 20240805 Extra edition	B2B Banking	Email Clicked	Crypto Market Update 20240805	Switzerland	05.08.2024 0:00	Client
Crypto Market Update	Crypto Market Update 20240805	B2B Banking	Email Clicked	Crypto Market Update 20240805	Switzerland	05.08.2024 0:00	Client

	Extra edition						
Crypto Market Update	Crypto Market Update 20240814	B2B Banking	Email Clicked	Crypto Market Update 20240814	Singapore	14.08.2024 0:00	Client
Crypto Market Update	Crypto Market Update 20240814	B2B Banking	Email Clicked	Crypto Market Update 20240814	Switzerland	14.08.2024 0:00	Client
Crypto Market Update	Crypto Market Update 20240814	B2B Banking	Email Clicked	Crypto Market Update 20240814	Switzerland	14.08.2024 0:00	Client
Crypto Market Update	Crypto Market Update 20240821	B2B Banking	Email Clicked	Crypto Market Update 20240821	Switzerland	21.08.2024 0:00	Client
Crypto Market Update	Crypto Market Update 20240904	B2B Banking	Email Clicked	Crypto Market Update 20240904	Switzerland	04.09.2024 0:00	Client
Crypto Market Update	Crypto Market Update 20240904	B2B Banking	Email Clicked	Crypto Market Update 20240904	Switzerland	04.09.2024 0:00	Client
Crypto Market Update	Crypto Market Update 20240911	B2B Banking	Email Clicked	Crypto Market Update 20240911	Switzerland	11.09.2024 0:00	Client
Crypto Market Update	Crypto Market Update 20240925	B2B Banking	Email Clicked	Crypto Market Update 20240925	Switzerland	25.09.2024 0:00	Client
Crypto Market Update	Crypto Market Update 20240925	B2B Banking	Email Clicked	Crypto Market Update 20240925	Switzerland	25.09.2024 0:00	Client
Crypto Market Update	Crypto Market Update 20241002	B2B Banking	Email Clicked	Crypto Market Update 20241002	Switzerland	02.10.2024 0:00	Client
Crypto Market Update	Crypto Market Update 20241009	B2B Banking	Email Clicked	Crypto Market Update 20241009	Switzerland	09.10.2024 0:00	Client
Crypto Market Update	Crypto Market Update 20241016	B2B Banking	Email Clicked	Crypto Market Update 20241016	Switzerland	16.10.2024 0:00	Client

Crypto Market Update	Crypto Market Update 20241023	B2B Banking	Email Clicked	Crypto Market Update 20241023	Switzerlan d	23.10.202 4 0:00	Client
----------------------------	--	----------------	------------------	--	-----------------	---------------------	--------