

Міністерство освіти і науки України
Харківський національний університет радіоелектроніки

Факультет Комп'ютерних наук
(повна назва)

Кафедра Інформаційних управляючих систем
(повна назва)

КВАЛІФІКАЦІЙНА РОБОТА
Пояснювальна записка

рівень вищої освіти другий (магістерський)

Дослідження та використання методів аналізу поведінки клієнтів
при розробці інформаційної системи для планування діяльності
підприємства
(тема)

Виконав:

здобувач 2 року навчання, групи ІУСТм-24-1
Потапенко Анна Олександрівна
(прізвище, ім'я, по батькові)

Спеціальність 122 Комп'ютерні науки
(код і повна назва спеціальності)

Тип програми освітньо-професійна
(освітньо-професійна або освітньо-наукова)

Освітня програма Інформаційні управляючі
системи та технології
(повна назва освітньої програми)

Керівник зав. каф. ІУС Костянтин ПЕТРОВ
(посада, власне ім'я, прізвище)

Допускається до захисту

Зав. кафедри ІУС



(підпис)

Костянтин ПЕТРОВ

(власне ім'я, прізвище)

2025 р.

Харківський національний університет радіоелектроніки

Факультет _____ Комп'ютерних наук _____

Кафедра _____ Інформаційних управляючих систем _____

Рівень вищої освіти _____ другий (магістерський) _____

Спеціальність _____ 122 Комп'ютерні науки _____
(код і повна назва)

Тип програми _____ освітньо-професійна _____
(освітньо-професійна або освітньо-наукова)

Освітня програма _____ Інформаційні управляючі системи та технології _____
(повна назва)

ЗАТВЕРДЖУЮ:

Зав. кафедри _____

(підпис)

“ 24 ” листопада 2025 р.

ЗАВДАННЯ НА КВАЛІФІКАЦІЙНУ РОБОТУ

здобувачеві _____ Потапенко Анні Олександрівні _____
(прізвище, ім'я, по батькові)

1. Тема роботи Дослідження та використання методів аналізу поведінки клієнтів при розробці інформаційної системи для планування діяльності підприємства

затверджена наказом по університету від “ 24 ” листопада 2025 р. № 1055Ст

2. Термін подання здобувачем роботи до екзаменаційної комісії “ 17 ” грудня 2025 р.

3. Вихідні дані до роботи наукові публікації та інтернет-джерела з тематики аналізу поведінки клієнтів, матеріали передатестаційної практики, інформація про методи прогнозування та оцінки CLV, транзакційні дані підприємства електронної торгівлі, додаткові джерела інформації.

4. Перелік питань, що потрібно опрацювати в роботі аналіз предметної області, аналіз існуючих методів аналізу поведінки клієнтів, дослідження проблем використання методів аналізу поведінки клієнтів в ІС, формулювання постановки задачі, дослідження методу когортного аналізу, CLV та survival-моделювання, розробка комбінованого методу аналізу поведінки клієнтів, експериментальна перевірка працездатності методу, аналіз отриманих результатів, формування висновків.

КАЛЕНДАРНИЙ ПЛАН


№	Назва етапів роботи	Термін виконання етапів роботи	Примітка
1	Аналіз предметної області	24.11.2025 – 26.11.2025	Виконано
2	Аналіз існуючих методів прогнозування та підходів до оцінки поведінки клієнтів	27.11.2025 – 29.11.2025	Виконано
3	Формулювання постановки задачі	30.11.2025	Виконано
4	Обґрунтування вибору методів та формування структури комбінованого методу	01.12.2025 – 03.12.2025	Виконано
5	Розробка комбінованого методу аналізу поведінки клієнтів	04.12.2025 – 06.12.2025	Виконано
6	Експериментальна перевірка працездатності запропонованого комбінованого методу аналізу поведінки клієнтів	07.12.2025 – 09.12.2025	Виконано
7	Аналіз отриманих результатів та інтерпретація поведінкових показників	10.12.2025	Виконано
8	Оформлення пояснювальної записки	11.12.2025 – 13.12.2025	Виконано
9	Оформлення графічного матеріалу	13.12.2025	Виконано
10	Захист кваліфікаційної роботи	17.12.2025	Виконано

Дата видачі завдання 24 листопада 2025 р.

Здобувач


(підпис)

Керівник роботи


(підпис)

зав. каф. ІУС Петров К. Е.

(посада, прізвище, ініціали)

РЕФЕРАТ

Пояснювальна записка кваліфікаційної роботи: 86 с., 8 рис., 11 табл., 1 дод., 37 джерел.

ЖИТТЄВА ЦІННІСТЬ КЛІЄНТА, КОГОРТНИЙ АНАЛІЗ, КОМБІНОВАНИЙ МЕТОД, МАРКЕТИНГОВІ СТРАТЕГІЇ, ПОВЕДІНКА КЛІЄНТА, ПРИЙНЯТТЯ УПРАВЛІНСЬКИХ РІШЕНЬ.

Об'єктом дослідження кваліфікаційної роботи є процес аналізу поведінки клієнтів у сфері електронної комерції.

Предметом дослідження є методи аналізу поведінки клієнтів, що використовуються для підтримки процесів планування діяльності в інформаційній системі підприємства.

Метою кваліфікаційної роботи є розробка комбінованого методу аналізу поведінки клієнтів на основі використання існуючих методів з подальшою реалізацією у вигляді аналітичного модуля інформаційної системи, що призначена для підтримки управлінських рішень щодо планування діяльності підприємства.

Наукова новизна дослідження полягає у розробці комбінованого методу аналізу поведінки клієнтів у контексті підтримки процесів планування діяльності підприємства. На відміну від більшості існуючих підходів, що зосереджені на окремих аспектах поведінки, пропонується поєднання методів когортного аналізу, оцінювання життєвої цінності клієнта та прогнозування відтоку. Таке поєднання дозволить формалізувати поведінкові патерни клієнтів і адаптувати їх до цілей інформаційної системи для планування.

ABSTRACT

Master's thesis: 86 pages, 8 figures, 11 tables, 1 appendix, 37 sources.

COHORT ANALYSIS, COMBINED METHOD, CUSTOMER BEHAVIOR, CUSTOMER LIFETIME VALUE, DECISION-MAKING, MARKETING STRATEGIES.

The object of research of the qualification work is the process of analyzing customer behavior in the field of e-commerce.

The subject of the research is the methods of customer behavior analysis applied to support enterprise planning processes within an information system.

The aim of this thesis is to develop a combined method for customer behavior analysis based on the integration of existing modern approaches, with its subsequent implementation as an analytical module of an information system designed to support managerial decision-making in enterprise planning.

The scientific novelty of the research lies in the development of a combined method for analyzing customer behavior in the context of supporting enterprise planning processes. Unlike most existing approaches that focus on individual aspects of customer behavior, the proposed approach integrates cohort analysis, customer lifetime value assessment, and churn prediction. This combination makes it possible to formalize customer behavioral patterns and align them with the objectives of the enterprise planning information system.

ЗМІСТ

Скорочення та умовні позначки	8
Вступ.....	9
1 Аналіз предметної області та постановка задачі дослідження.....	11
1.1 Сучасний стан проблеми аналізу поведінки клієнтів при плануванні діяльності підприємства	11
1.2 Аналіз існуючих методів аналізу поведінки клієнтів	13
1.2.1 Когортний аналіз.....	14
1.2.2 Методи прогнозування відтоку клієнтів.....	17
1.2.3 Методи, що використовують підпискові моделі	19
1.2.4 Метод розрахунку життєвої цінності клієнта	21
1.2.5 Методи оцінювання ефективності маркетингових каналів	22
1.2.6 Порівняння методів аналізу поведінки клієнтів	24
1.3 Проблеми використання методів аналізу поведінки клієнтів при розробці інформаційної системи для планування діяльності підприємства.....	26
1.4 Постановка задачі дослідження.....	27
2 Дослідження методів аналізу поведінки клієнтів	29
2.1 Вибір та обґрунтування груп методів	29
2.2 Класичні методи аналізу поведінки клієнтів	31
2.3 Прогнозні методи	35
2.4 Порівняльний аналіз методів	37
3 Розробка комбінованого методу аналізу поведінки клієнтів	39
3.1 Вимоги до розроблюваного методу	39
3.2 Основні етапи методу	40
3.3 Формальний опис етапів комбінованого методу	44
3.4 Інтеграція методу в ІС та очікувані результати	56

4 Експериментальна перевірка працездатності розробленого комбінованого методу аналізу поведінки клієнтів	58
4.1 Приклад застосування розробленого комбінованого методу.....	58
4.2 Аналіз отриманих результатів	64
Висновки	67
Перелік джерел посилання	69
Додаток А Графічний матеріал кваліфікаційної роботи.....	

СКОРОЧЕННЯ ТА УМОВНІ ПОЗНАКИ

ІС – інформаційна система

CLC – Customer Lifecycle

CLV – Customer Lifetime Value

IDEF0 – Integration Definition for Function Modeling

RFM – Recency, Frequency, Monetary

ВСТУП

Зростання потреби підприємств в аналізі даних про поведінку клієнтів і використання його результатів для прийняття ефективних управлінських рішень у сфері планування діяльності є актуальною проблемою для сучасного бізнесу. В умовах цифрової трансформації та високої конкуренції на ринку аналіз поведінки клієнтів та їх вподобань стає ключовим чинником у підвищенні прибутковості та сталого розвитку бізнесу.

Тема дослідження є актуальною через недостатній рівень інтеграції сучасних методів, наприклад, таких як когортний аналіз, прогнозування відтоку, підпискові моделі, моделі життєвої цінності клієнта (customer lifetime value, CLV) та омніканальні стратегії в інформаційні системи (ІС), які використовуються бізнесом для аналізу поведінки клієнтів. Існуючі дослідження, як правило, зосереджені на аналізі окремих аспектів поведінки клієнтів, тоді як нагальною потребою підприємства є створення єдиного аналітичного середовища для підтримки прийняття управлінських рішень.

Об'єктом дослідження кваліфікаційної роботи є процес аналізу поведінки клієнтів у сфері електронної комерції.

Предметом дослідження є методи аналізу поведінки клієнтів, що використовуються для підтримки процесів планування діяльності в ІС підприємства.

Метою кваліфікаційної роботи є розробка комбінованого методу аналізу поведінки клієнтів на основі використання існуючих методів з подальшою реалізацією у вигляді аналітичного модуля ІС, що призначена для підтримки управлінських рішень щодо планування діяльності підприємства.

Методика проведення дослідження передбачає використання методів системного аналізу для вивчення предметної області та постановки задачі дослідження, розробку комбінованого методу аналізу поведінки клієнтів на основі використання сучасних методів з подальшою реалізацією у вигляді

аналітичного модуля ІС; проведення серії комп'ютерних експериментів (експериментального дослідження) для визначення його працездатності при вирішенні практичних завдань, аналіз отриманих результатів, формулювання висновків і рекомендацій щодо практичного використання розробленого методу.

Наукова новизна дослідження полягає у розробці комбінованого методу аналізу поведінки клієнтів у контексті підтримки процесів планування діяльності підприємства. На відміну від більшості існуючих підходів, що зосереджені на окремих аспектах поведінки, запропоновано поєднання методів когортного аналізу, оцінювання CLV та прогнозування відтоку. Таке поєднання дозволить формалізувати поведінкові патерни клієнтів.

Практична цінність дослідження полягає у можливості використання отриманих результатів для створення аналітичного модуля ІС підприємства, що забезпечить:

- прогнозування ймовірності відтоку клієнтів та формування рекомендацій щодо їх утримання;
- виокремлення когорт та ціннісних сегментів клієнтів для підтримки прийняття управлінських рішень у сфері планування продажів;
- уточнення прогнозу доходів у розрізі сегментів клієнтської бази на основі поєднання показників CLV та динаміки утримання клієнтів;
- інформаційну підтримку вибору варіантів планів діяльності підприємства (зокрема, планів продажів і заходів взаємодії з клієнтами) з урахуванням виявлених поведінкових патернів.

1 АНАЛІЗ ПРЕДМЕТНОЇ ОБЛАСТІ ТА ПОСТАНОВКА ЗАДАЧІ ДОСЛІДЖЕННЯ

1.1 Сучасний стан проблеми аналізу поведінки клієнтів при плануванні діяльності підприємства

В умовах цифрової економіки та глобальної конкуренції підприємства дедалі більше усвідомлюють важливість аналізу поведінки клієнтів як ключового елементу стратегічного управління. Якщо раніше цей напрям розглядався здебільшого як інструмент підтримки маркетингу, то сьогодні він трансформувався у фундаментальну частину системи планування діяльності.

Сучасне стратегічне планування все більше орієнтується на концепцію клієнтоорієнтованості (customer-centricity), де клієнт стає центральним елементом всіх бізнес-процесів. У дослідженні [1] підкреслюється, що інтеграція аналітики клієнтської поведінки в системи корпоративного планування дозволяє створювати більш точні фінансові моделі, основані на прогнозах майбутньої цінності клієнтської бази, а не лише на історичних показниках. Це особливо важливо для компаній із підписковими моделями, де вартість бізнесу безпосередньо залежить від життєвої вартості клієнтів.

Причини такої трансформації пов'язані з низкою чинників. По-перше, цифровізація бізнесу призвела до різкого зростання обсягів клієнтських даних: транзакційних, поведінкових, соціальних. Кожна взаємодія клієнта з компанією (у вебсайті, мобільному додатку чи соціальних мережах) формує цифровий слід, який може бути використаний для аналітики. По-друге, сучасні моделі економіки, зокрема електронна комерція, підпискові сервіси, ґрунтуються на довгострокових відносинах із клієнтами, де головним завданням стає утримання, а не лише залучення. По-третє, підвищена конкуренція робить прогнозованість доходів і точність фінансових планів неможливими без глибокого розуміння клієнтської поведінки.

У результаті аналіз поведінки клієнтів набуває стратегічного значення,

оскільки забезпечує:

- прогнозування доходів, оскільки моделі розвитку клієнтської бази дозволяють будувати точні фінансові сценарії;
- оптимізацію ресурсів, адже розподіл маркетингового бюджету без точних прогнозів часто є неефективним;
- зростання прибутковості, завдяки своєчасному виявленню клієнтів із високим ризиком відтоку;
- підтримку стратегічних рішень, що базуються не лише на історичних даних, а й на прогнозних оцінках майбутньої цінності клієнтів.

Сучасні наукові дослідження [2] підтверджують, що аналіз поведінки клієнтів стає наріжним каменем у процесах планування діяльності підприємств. Він формує основу для прийняття обґрунтованих управлінських рішень, прогнозування динаміки доходів та розробки маркетингових стратегій, орієнтованих на довгострокову взаємодію з клієнтами.

Інтеграція модуля аналізу поведінки клієнтів у системи планування створює комплексний ефект у кількох напрямках. По-перше, вона забезпечує безперервний цикл зворотного зв'язку між операційною діяльністю та стратегічними цілями. Дані про клієнтів у режимі реального часу дозволяють оперативно коригувати рішення та переглядати стратегічні орієнтири. По-друге, використання єдиного джерела поведінкових даних сприяє розвитку крос-функціональної взаємодії: відділи маркетингу, продажів, обслуговування клієнтів і розробки продуктів працюють узгоджено, спираючись на спільну інформаційну базу. Це усуває традиційні розбіжності в інтерпретації ринкової ситуації. По-третє, аналітика дозволяє реалізувати принцип динамічного планування, коли фінансові прогнози, маркетингові бюджети та операційні показники постійно оновлюються на основі актуальних даних.

Важливим є і багаторівневий ефект від використання поведінкової аналітики:

- на операційному рівні: дозволяє оптимізувати щоденні бізнес-процеси від персоналізації маркетингових повідомлень до налаштування роботи кол-

центрів;

- на тактичному рівні: забезпечує основу для розподілу ресурсів між різними каналами продажів, сегментами клієнтів та географічними ринками;
- на стратегічному рівні: стає критично важливим для прийняття рішень про вихід на нові ринки, запуск нових продуктів та формування довгострокових конкурентних переваг.

Окремо слід відзначити, що розвиток клієнтської аналітики відбувається на тлі зростання вимог до етичності та прозорості використання даних. Питання захисту персональних даних та дотримання стандартів GDPR набувають стратегічного значення [3]. У багатьох випадках здатність компанії правильно працювати з клієнтською інформацією визначає її довіру на ринку.

Разом із тим, попри зростаюче усвідомлення важливості цього напрямку, більшість компаній перебувають лише на етапі інтеграції сучасних підходів у свої ІС. Традиційні інструменти управління спираються переважно на агреговані історичні дані, що обмежує точність прогнозування та ускладнює прийняття стратегічних рішень. Тому актуальною є потреба у подальшому розвитку комбінованого методу аналізу поведінки клієнтів та створенні аналітичного модуля, здатних вбудовуватися в ІС підприємств.

Отже, сучасний стан проблеми можна схарактеризувати як перехідний: від розуміння важливості клієнтської аналітики до її практичного впровадження як основи планування діяльності. Цей перехід відкриває перспективи для наукових досліджень, спрямованих на інтеграцію різних методів аналізу та їх застосування у практиці електронної комерції та суміжних галузей.

1.2 Аналіз існуючих методів аналізу поведінки клієнтів

Сучасна практика бізнес-аналітики пропонує широкий спектр методів

для дослідження поведінки клієнтів. Вони відрізняються підходами, математичним апаратом та рівнем деталізації результатів, проте всі мають спільну мету – підтримати підприємство у прийнятті управлінських рішень. Зокрема, ці методи спрямовані на вирішення таких ключових завдань, як прогнозування доходів, виявлення ризиків відтоку, оптимізація маркетингових інвестицій та підвищення ефективності взаємодії з клієнтами.

У науковій літературі та бізнес-практиці сформувалися кілька основних напрямів, які стали базовими в сучасних системах підтримки прийняття рішень. До них належать когортний аналіз, моделі прогнозування відтоку, підпискові та транзакційні моделі, оцінка CLV, а також методи атрибуції маркетингових каналів на основі імовірнісних моделей.

Популярність цих методів зумовлена їх широким використанням у бізнес-аналітиці та відносною доступністю реалізації в ІС підприємств. Вони стали поширеними завдяки простоті впровадження, наявності готових інструментів та накопиченому практичному досвіду застосування. Крім того, ці підходи дозволяють аналізувати клієнтську поведінку у часовому розрізі та формувати показники, що можуть бути використані в процесах планування діяльності. Зараз вони становлять основу більшості корпоративних аналітичних рішень, хоча їх застосування не позбавлене суттєвих обмежень, які потребують подальшого дослідження.

У подальших підрозділах розглянуто особливості кожного з цих методів, їхні переваги та обмеження, а також можливості застосування в рамках комбінованого підходу.

1.2.1 Когортний аналіз

Когортний аналіз [4] є одним із базових методів дослідження поведінки клієнтів, який дає змогу відстежувати зміни у динаміці з урахуванням часу та

спільних характеристик певних груп користувачів. Його сутність полягає у формуванні когорт – груп клієнтів, об'єднаних за спільною ознакою: часом першої покупки чи реєстрації, каналом залучення, реакцією на маркетингову акцію тощо. Далі аналізується, як змінюється їхня активність протягом вибраного періоду, що дозволяє виявляти закономірності утримання, відтоку та повторних транзакцій.

На відміну від суто агрегованих показників, когортний аналіз дає змогу уникнути викривлення результатів через змішування різних груп клієнтів. Наприклад, загальний рівень продажів може здаватися стабільним, однак детальний розгляд показує, що нові когорти втрачають лояльність швидше за попередні. Подібна аналітична інформація є важливою для планування, оскільки дає змогу вчасно виявляти проблеми та коригувати маркетингові та продуктові стратегії.

У практиці електронної комерції когортний аналіз особливо корисний для:

- оцінки рівня утримання клієнтів, тобто частки користувачів, які залишаються активними упродовж певного періоду;
- визначення масштабів відтоку і часу, коли клієнти найчастіше залишають сервіс;
- дослідження повторних покупок та змін середнього чека залежно від когорти;
- вимірювання ефективності рекламних кампаній або промоакцій, оскільки різні когорти можуть демонструвати різну чутливість до маркетингових стимулів;
- виявлення впливу зовнішніх факторів (сезонності, оновлення продуктів, конкуренції) на поведінку різних груп клієнтів [5].

Методика когортного аналізу зазвичай ґрунтується на побудові retention-таблиць і графіків, які відображають відсоток клієнтів, що зберігають активність у часі.

Приклад когортного аналізу подано у вигляді retention-таблиці (рис. 1.1).

У ній відображено, як змінюється частка активних користувачів, що повертається у наступні тижні після першої взаємодії з сервісом [6].


Rate (%) 	Week 0	Week 1	Week 2	Week 3	Week 4	Week 5
02.24 - 03.02 76 Users	100%	24%	18%	14%	8%	8%
03.02 - 03.09 88 Users	100%	28%	29%	24%	24%	
03.09 - 03.16 60 Users	100%	25%	15%	21%		
03.16 - 03.23 63 Users	100%	32%	20%			
03.23 - 03.30 58 Users	100%	26%				
03.30 - 04.06 63 Users	100%					
Average new users per week 68 Users	100%	27%	22%	21%	19%	10%

Рисунок 1.1 – Приклад когортного аналізу

У першому стовпці представлено когорти – групи клієнтів, які вперше приєдналися у певний тиждень. У стовпці Week 0 для всіх когорт прийнято частку утримання 100% користувачів, а далі відображається, який відсоток із них залишався активним у наступні тижні.

Наприклад, для першої когорти у Week 1 активними залишилося 24% клієнтів, у Week 2 – 18%, а у Week 4 – лише 8%. Середні значення внизу показують загальну тенденцію: вже за тиждень залишається близько 27% користувачів, а на п'ятому тижні – лише 10%.

Таким чином, застосування когортного аналізу у стратегічному плануванні має низку переваг. По-перше, він дає змогу точніше прогнозувати майбутні доходи, оскільки враховує не лише факт покупки, а й динаміку взаємодії клієнта з компанією. По-друге, на основі результатів аналізу можна сегментувати клієнтську базу та розробляти цілеспрямовані заходи щодо

утримання. По-третє, він допомагає оцінювати, які канали залучення приводять клієнтів із найвищою довгостроковою цінністю.

1.2.2 Методи прогнозування відтоку клієнтів

Прогнозування відтоку клієнтів [7] є одним із ключових завдань у сфері бізнес-аналітики та управління відносинами з клієнтами. Воно дозволяє своєчасно визначати тих користувачів, які з високою ймовірністю відмовляться від послуг компанії або перестануть здійснювати покупки. Актуальність цього напрямку пояснюється тим, що збереження наявного клієнта обходиться значно дешевше, ніж залучення нового, а стійкість доходів підприємства значною мірою залежить від рівня утримання.

Завдання прогнозування відтоку найчастіше формується як бінарна класифікація: клієнт або належить до групи «відтік», або до групи «утримання». При цьому важливим є не лише визначення ймовірності відтоку, але й інтерпретація факторів, що впливають на поведінку клієнтів. Це дозволяє компаніям не просто передбачати ризик, а й розробляти ефективні стратегії утримання.

До класичних підходів прогнозування відтоку відносять:

- логістичну регресію [8], яка дозволяє оцінювати вплив окремих характеристик на ймовірність відтоку та відрізняється простотою інтерпретації;
- аналіз виживання (Survival Analysis) [9], що використовується для прогнозування часу до відтоку клієнта і є особливо корисною в підпискових моделях;
- структурне моделювання рівнянь (structural equation modeling) [9], яке дає можливість аналізувати зв'язки між такими змінними, як, наприклад, задоволеність, лояльність та відтік.

Перевагою цих методів є прозорість та зрозумілість результатів, проте в умовах великих масивів даних і складних нелінійних залежностей їхня прогностична сила обмежена.

Розвиток машинного навчання значно розширив можливості у сфері прогнозування відтоку. Серед найбільш поширених підходів є:

- дерева рішень, random forest, gradient boosting, XGBoost [10] – ансамблеві алгоритми, здатні враховувати складні залежності та взаємодії між ознаками;
- метод опорних векторів [11] – ефективний для роботи з багатовимірними даними;
- k-найближчих сусідів [12] – базується на пошуку схожих клієнтів, що дає зрозумілий механізм прогнозування;
- штучні нейронні мережі [12] – застосовуються для роботи з великими та різномірними наборами даних, виявляючи приховані закономірності.

Однією з головних проблем у цій сфері є дисбаланс класів: кількість клієнтів, які залишаються, значно перевищує кількість тих, що відмовляються від послуг. Це ускладнює навчання моделей, які часто «схиляються» до прогнозування класу більшості. Для подолання цього обмеження застосовуються різні методи балансування даних, які синтетично створюють приклади для класу «відтік».

Дослідження у різних галузях [12] показують високу ефективність сучасних підходів до прогнозування відтоку. У телекомунікаційному секторі, де ця проблема вивчена найглибше, моделі на основі методу випадкових лісів із налаштованими параметрами демонструють точність понад 95% та високі показники відтворюваності результатів. Хоча приклади стосуються телеком-індустрії, подібні підходи успішно застосовуються і в електронній комерції. Тут ключовими факторами ризику відтоку є частота транзакцій, середній чек, час від останньої покупки, рівень взаємодії з маркетинговими каналами та задоволеність сервісом.

У контексті розробки інтегрованого аналітичного модуля ІС прогноз

відтоку відіграє центральну роль, оскільки результати такого аналізу безпосередньо використовуються для розрахунку CLV, оцінки ефективності маркетингових заходів та побудови фінансових прогнозів.

1.2.3 Методи, що використовують підпискові моделі

Використання підпискових моделей [13] є одним із найдинамічніших напрямів розвитку сучасних бізнес-стратегій. Вони дедалі активніше застосовуються не лише у сфері роботи з кінцевими споживачами, а й у сегменті взаємодії між підприємствами. Їх сутність полягає у встановленні тривалих відносин між підприємством та клієнтом, де останній здійснює регулярні платежі за доступ до продукту чи послуги. Такий підхід принципово відрізняється від транзакційних моделей, орієнтованих на разовий продаж, оскільки акцент переноситься на довгострокову взаємодію та утримання споживача.

Ключова особливість підпискових моделей полягає в формуванні передбачуваного та стабільного грошового потоку. Завдяки цьому підприємства отримують можливість точніше прогнозувати доходи, оптимізувати запаси та планувати інвестиції. З іншого боку, клієнти виграють завдяки зручності, гнучкості та персоналізованим пропозиціям, що відповідають їхнім потребам [14]. Постійне оновлення та індивідуалізація пропозицій формують основу для утримання користувачів і зниження рівня відтоку.

З методологічного погляду підпискові моделі дають змогу реалізувати низку інструментів для аналізу та прогнозування поведінки клієнтів:

- аналіз тривалості життєвого циклу клієнта (customer lifecycle, CLC), який дозволяє оцінити ймовірність збереження підписки у часі;
- прогнозування кількості транзакцій і частоти використання сервісу,

що є базою для подальшого розрахунку CLV;

- моделі поведінкової сегментації, які виділяють когорти клієнтів за рівнем лояльності, платоспроможності чи чутливості до ціни;
- динамічне ціноутворення у форматі pay-per-use або багаторівневих тарифів («good–better–best»), що адаптуються до індивідуальних сценаріїв споживання [15].

Застосування підпискових моделей набуває великого значення у сфері електронної комерції. Згідно з дослідженням [16], традиційні бізнес-моделі часто не здатні забезпечити гнучкість у мінливому цифровому середовищі, тоді як підпискові підходи дозволяють швидко реагувати на зміну попиту, надавати персоналізовані пропозиції та забезпечувати конкурентні переваги. Крім того, вони інтегруються з сучасними цифровими інструментами опрацювання великих даних, штучного інтелекту та Інтернету речей, що підсилює можливості прогнозування та персоналізації.

Важливим аспектом розвитку підпискових моделей є їхня різноманітність. Підписки можуть орієнтуватися як на досягнення певного результату, так і різнитися за ступенем інтеграції ресурсів: від односпрямованих (де більшість процесів автоматизовано) до взаємних, що передбачають активну співпрацю між постачальником і клієнтом.

Водночас підпискові моделі несуть і виклики, наприклад «втома від надмірної кількості підписок», складність утримання клієнтів у довгостроковій перспективі та необхідність постійних інновацій у продукті та сервісі. Для подолання цих бар'єрів дослідники [14] пропонують використовувати комбінацію інструментів: регулярне оновлення контенту чи функціоналу, застосування фріміум-моделі підписки для залучення нових клієнтів, а також розвиток спеціальних систем для супроводу клієнта на всіх етапах взаємодії.

Отже, методи, що базуються на підпискових моделях, не лише дозволяють прогнозувати поведінку клієнтів, але й формують основу для довгострокового стратегічного планування. Вони інтегрують маркетингові,

фінансові та аналітичні інструменти, створюючи підґрунтя для розрахунку CLV та визначення ефективності бізнес-моделі загалом.

1.2.4 Метод розрахунку життєвої цінності клієнта

Метод CLV [17] є одним із ключових інструментів у сучасному маркетинговому та аналітичному менеджменті. Його сутність полягає в оцінці сукупного доходу, який клієнт може принести підприємству за весь період співпраці. Такий підхід дозволяє компаніям не лише вимірювати економічну цінність клієнтської бази, а й приймати обґрунтовані рішення щодо сегментації, утримання та інвестування у розвиток відносин із клієнтами.

У класичному вигляді CLV обчислюється за формулою:

$$CLV = AOV * ALT * AGM * P, \quad (1.1)$$

де AOV – сума покупки, грн;

ALT – цикл життя покупця, міс.;

AGM – маржа, %.

P – кількість покупок, здійснених клієнтом.

Цикл життя покупця обчислюється за формулою:

$$ALT = \frac{1}{Ch}, \quad (1.2)$$

де Ch – коефіцієнт відтоку клієнтів, %.

Коефіцієнт відтоку клієнтів визначається за формулою:

$$Ch = \frac{(CB - CE)}{CB} * 100\%, \quad (1.3)$$

де СВ – кількість клієнтів на початок періоду, осіб;

СЕ – кількість клієнтів на кінець періоду, осіб.

Наукові роботи [18, 19] присвячені аналізу та порівнянню результатів традиційного підходу, що поєднує аналіз давності, частоти та обсяг витрат клієнта (recency, frequency, monetary, RFM), та сучасних алгоритмів машинного навчання. Хоча RFM залишається простим і зручним методом сегментації, його прогностична точність є обмеженою. Автори доводять, що застосування методів машинного навчання дозволяє отримати більш надійні прогнози CLV завдяки врахуванню ширшого спектра змінних: частоти відвідувань, поведінки в онлайн-каналах, реакції на маркетингові стимули. Такий підхід допомагає не лише оцінювати поточну цінність клієнта, але й формувати персоналізовані стратегії взаємодії.

Сучасна література [20] демонструє, що розвиток методів CLV рухається у напрямку від статичних фінансових моделей до динамічних, поведінково-орієнтованих алгоритмів. Залучення інструментів штучного інтелекту дозволяє не лише підвищити точність прогнозів, але й забезпечити стратегічну цінність для бізнесу зокрема у плануванні доходів, оптимізації маркетингових бюджетів і розробці програм лояльності.

1.2.5 Методи оцінювання ефективності маркетингових каналів

Оцінювання ефективності маркетингових каналів [21] – це система методів, яка дає змогу кількісно визначати внесок окремих точок контакту з клієнтом (реклама, пошук, e-mail, соціальні мережі тощо) у досягнення бізнес-цілей та приймати обґрунтовані рішення щодо розподілу бюджету. На базовому рівні використовують різні ключові показники ефективності, які пов'язують витрати, дохід і CLV клієнтської бази. До основних показників належать: рентабельність інвестицій, що показує, який дохід генерується на

одиницю вкладених коштів; окупність рекламних витрат, яка відображає дохід, отриманий завдяки цим витратам; вартість залучення клієнта, тобто сукупні витрати на залучення одного нового клієнта; CLV – очікуваний сумарний прибуток від клієнта за весь період співпраці.

Практичне правило прийняття інвестиційних рішень ґрунтується на співвідношенні CLV та вартості його залучення: інвестиції в канал вважаються доцільними, коли CLV перевищує вартість.

Важливим аспектом для аналізу виступає омніканальність. Сучасний клієнт взаємодіє з брендом через низку онлайн та офлайн каналів, і завдання бізнесу полягає в інтеграції цих дотиків у єдину послідовність для забезпечення послідовного та зручного користувацького досвіду. Дослідження [22] підкреслює, що успіх омніканальної стратегії залежить від гармонізації даних, формування єдиного профілю клієнта та використання аналітики для персоналізації взаємодій. Водночас у багатоканальному середовищі маркетингові впливи мають як прямий, так і синергетичний ефект на прийняття рішення, що потребує алгоритмічних моделей атрибуції для оптимізації бюджету.

У прикладних дослідженнях виділяють кілька основних груп методів атрибуції, кожна з яких має свої переваги та обмеження.

Перша група – правилі моделі (rule-based) [23]. До них належать підходи, які розподіляють вагу конверсії між каналами за наперед визначеним правилом: перший клік, останній клік, рівномірний розподіл. Такі моделі прості у реалізації та зрозумілі для бізнесу, проте вони не враховують порядок подій, взаємодію каналів і відмінність між конвертуючими та неконвертуючими шляхами.

Далі – марковські моделі [23]. Цей підхід розглядає взаємодію користувача з каналами як послідовність станів у ланцюгу Маркова. Кожен канал виступає станом, а конверсія або відмова – поглинаючим станом. На основі історичних даних формується матриця переходів, яка відображає ймовірності переходів між каналами. Внесок окремого каналу визначається за

підходом його виключення з моделі: канал видаляється, після чого повторно обчислюється ймовірність конверсії; різниця між базовим і модифікованим сценарієм інтерпретується як внесок цього каналу.

Наступна група – сучасні data-driven підходи [23]. До них відносять методи машинного навчання та причинно-наслідкові моделі, які враховують часові лаги, інтенсивність взаємодій та приховані фактори. Вони здатні будувати складні нелінійні залежності, однак потребують великих обсягів якісних даних і значних обчислювальних ресурсів.

1.2.6 Порівняння методів аналізу поведінки клієнтів

Методи, розглянуті у попередніх підрозділах, широко використовуються в бізнес-аналітиці та підтримці управлінських рішень. Проте їх окреме застосування має низку обмежень: вони орієнтовані лише на окремі аспекти поведінки клієнтів і не забезпечують повного охоплення завдань планування діяльності підприємства.

Для систематизації переваг і недоліків доцільно порівняти методи за ключовими критеріями (табл. 1.1).

Порівняння показує, що кожен із методів є корисним у власній сфері застосування: когортний аналіз описує історичну динаміку, прогнозування відтоку дає змогу діяти превентивно, підпискові моделі ефективні для сервісів із регулярними транзакціями, CLV використовується для стратегічних рішень, а атрибуційні методи допомагають оптимізувати маркетингові інвестиції.

Водночас жоден із підходів не забезпечує комплексного охоплення всієї клієнтської поведінки. Це обумовлює необхідність комбінування різних методів в єдиному аналітичному модулі ІС.

Таблиця 1.1 – Порівняння методів аналізу поведінки клієнтів

Метод	Цілі використання	Тип даних	Інтерпретованість	Обмеження
Когортний аналіз	Виявлення динаміки утримання, аналіз відтоку та повторних транзакцій	Транзакційні, часові	Висока	Не дає прогнозу, відображає лише ретроспективні тенденції
Прогнозування відтоку	Визначення клієнтів із високим ризиком відтоку, формування стратегій утримання	Поведінкові, транзакційні	Середня	Вимоги до якості, складність у поясненні результатів
Підпискові моделі	Прогноз тривалості CLC, кількості транзакцій, майбутніх доходів	Дані підписок, історія користування	Середня	Не придатні для разових покупок, чутливі до відтоку
CLV-моделі	Розрахунок довгострокової цінності, стратегічне планування та сегментація	Фінансові, транзакційні	Висока або середня	Базові моделі спрощені, сучасні потребують великих масивів даних
Оцінка ефективності маркетингових каналів	Атрибуція конверсій, визначення внеску каналів, оптимізація бюджету	Дані клієнтських шляхів, омніканальні	Висока або середня	Складність реалізації, залежність від якості даних, відсутність інтегрованого прогнозу CLV

1.3 Проблеми використання методів аналізу поведінки клієнтів при розробці інформаційної системи для планування діяльності підприємства

Розробка ІС для планування діяльності підприємства передбачає широке застосування методів аналізу поведінки клієнтів. Такі методи допомагають краще зрозуміти клієнтські потреби, прогнозувати відтік, оцінювати CLV та ефективність маркетингових каналів. Проте на практиці їх упровадження в межах ІС пов'язане з низкою проблем.

Однією з ключових складностей є інтеграція даних із різних джерел у межах єдиної ІС підприємства. У практиці електронної комерції та сервісних компаній дані про клієнтів надходять із платформ аналітики, маркетингових сервісів, соціальних мереж та систем управління продажами. Вони мають різний формат і частоту оновлення, містять дублікати або неповні записи. Для аналізу поведінки клієнтів потрібен єдиний профіль, який об'єднує транзакційні, поведінкові та маркетингові характеристики. Проте створення такого інтегрованого профілю є складним завданням. Додатково проблеми виникають у зв'язку з омніканальністю, коли клієнт одночасно взаємодіє з компанією через декілька каналів (онлайн і офлайн), а система не завжди здатна правильно поєднати ці події. Відсутність якісної інтеграції та синхронізації призводить до викривлення результатів аналізу та обмежує можливості системи.

Сучасні методи аналізу потребують значних обчислювальних ресурсів. Для підприємства це означає витрати на серверну інфраструктуру, бази даних і програмні засоби. Зі зростанням клієнтської бази та кількості каналів взаємодії навантаження на систему зростає. Проблемою стає і масштабування ІС: необхідно забезпечувати стабільну роботу алгоритмів у режимі реального часу, швидку обробку великих масивів даних і відсутність затримок у прийнятті рішень.

Методи аналізу поведінки клієнтів зазвичай зосереджені на окремих

завданнях. Наприклад, когортний аналіз дає змогу досліджувати динаміку, але не прогнозує майбутні результати; моделі відтоку допомагають зменшити втрати клієнтів, але не дають інтегральної оцінки цінності; атрибуційні моделі показують ефективність каналів, але не прогнозують довгострокову поведінку. Це ускладнює інтеграцію різних підходів в єдине аналітичне середовище ІС. Крім того, виникає питання балансу між точністю прогнозів і зрозумілістю результатів: складні моделі забезпечують високу точність, але їхні результати важко пояснити керівництву. Така методологічна неоднорідність знижує практичну цінність аналітики.

Крім технічних і методологічних обмежень існують організаційні та етичні аспекти. Розробка і підтримка складних моделей аналізу поведінки клієнтів вимагає висококваліфікованих спеціалістів та значних фінансових витрат, що може бути непосильним для малого чи середнього бізнесу. Етичні проблеми пов'язані з використанням персональних даних клієнтів: їх потрібно обробляти відповідно до вимог законодавства, що накладає додаткові обмеження на розробку ІС. Також актуальною є проблема прийняття рішень: складні результати аналітики не завжди зрозумілі менеджерам, які відповідають за планування, і це створює бар'єр у впровадженні таких систем.

Тому жоден із оглянутих методів не забезпечує комплексного охоплення всіх аспектів поведінки клієнтів, а ізольоване використання окремих моделей призводить до неповної картини. Це формує необхідність у створенні інтегрованого підходу, який би об'єднав переваги різних методів і усунув слабкі сторони.

1.4 Постановка задачі дослідження

Аналіз існуючих підходів показав, що застосування окремих методів дозволяє вирішувати лише часткові завдання аналізу поведінки клієнтів і не

формує цілісної картини. Для забезпечення комплексного бачення клієнтської бази, підвищення обґрунтованості управлінських рішень та для задоволення інформаційних потреб підприємств постає потреба у створенні інтегрованого підходу, який поєднує різні методи в межах єдиної ІС.

Об'єктом дослідження є процес аналізу поведінки клієнтів у сфері електронної комерції.

Предметом дослідження є методи аналізу поведінки клієнтів, що використовуються для підтримки процесів планування діяльності в ІС підприємства.

Метою дослідження є розробка комбінованого методу аналізу поведінки клієнтів на основі використання існуючих методів з подальшою реалізацією у вигляді аналітичного модуля ІС, що призначена для підтримки управлінських рішень щодо планування діяльності підприємства.

Для досягнення мети необхідно вирішити такі завдання:

- дослідити існуючі методи аналізу поведінки клієнтів та визначити можливості їх взаємодії;
- розробити комбінований метод аналізу поведінки клієнтів на основі використання існуючих методів;
- реалізувати запропонований комбінований метод у вигляді аналітичного модуля та інтегрувати його в ІС планування діяльності підприємства;
- провести експериментальну перевірку працездатності розробленого методу.

2 ДОСЛІДЖЕННЯ МЕТОДІВ АНАЛІЗУ ПОВЕДІНКИ КЛІЄНТІВ

2.1 Вибір та обґрунтування груп методів

У сучасних дослідженнях аналізу поведінки клієнтів спостерігається активне застосування методів, заснованих на поєднанні апарату математичної статистики, методів машинного навчання та економічних моделей. Це зумовлено зростанням обсягів даних, доступних підприємствам для аналізу, а також потребою у більш точному прогнозуванні та персоналізації бізнес-рішень.

У роботі [24] відзначається, що алгоритми ШІ кардинально впливають як на діяльність компаній, так і на поведінку клієнтів. Автори підкреслюють, що машинне навчання може зменшувати інформаційну асиметрію, покращувати прогнозування та сприяти персоналізації пропозицій, але водночас породжує ризики упередженості та посилення ринкової концентрації. Це підкреслює, що для практичного застосування слід обирати методи, які не лише здатні працювати з великими даними, але й забезпечують баланс між ефективністю та зрозумілістю результатів.

Окрема група досліджень присвячена застосуванню предиктивної аналітики. У статті [25] показано, що використання ШІ у сегментації та динамічному ціноутворенні дає можливість компаніям ефективніше планувати ресурси та формувати персоналізовані пропозиції для клієнтів. Автори наголошують, що традиційні методи сегментації поступаються за точністю та швидкістю сучасним алгоритмам, здатним працювати з великими масивами даних.

Узагальнюючи огляд літератури, можна виділити кілька основних груп методів аналізу поведінки клієнтів:

- класичні методи: когортний аналіз та базові моделі CLV, які відзначаються простотою та інтерпретованістю, дають змогу відстежувати динаміку утримання клієнтів та оцінювати їхню цінність;

- прогнозні методи: моделі прогнозування відтоку клієнтів та розширені алгоритми розрахунку CLV з використанням машинного навчання, які забезпечують більш точні та оперативні оцінки;
- методи персоналізації: підпискові моделі та інструменти динамічного ціноутворення, які безпосередньо впливають на дохід і задоволеність клієнтів, але водночас потребують високого рівня цифрової зрілості компанії;
- імовірнісні та мережеві методи: методи атрибуції каналів на основі Марковських ланцюгів, що дозволяють оцінити ефективність маркетингових інвестицій, але мають обмежене застосування для індивідуального аналізу клієнтів.

Попри різноманіття підходів, не всі методи однаково придатні для практичної реалізації у межах дослідження.

Методи, що використовують підпискові моделі, безперечно є актуальними для сучасної економіки, особливо у сфері електронної комерції. Підпискові моделі забезпечують стабільні грошові потоки, а алгоритми динамічного ціноутворення дають змогу максимізувати дохід шляхом адаптації цін до попиту в реальному часі. Проте їхня універсальність обмежена: підписки підходять лише для бізнесів зі специфічною моделлю монетизації, а динамічне ціноутворення потребує великих обсягів даних і може призводити до зниження довіри клієнтів у разі відсутності прозорості. Тому ефект від застосування цих методів як основи ІС планування діяльності для широкого класу підприємств є сумнівним.

Суттєвою перевагою методів оцінювання маркетингових каналів є можливість точніше визначати їхній вплив на залучення клієнтів, що допомагає оптимізувати рекламний бюджет. Проте цей напрям не дає комплексного уявлення про CLV та індивідуальну поведінку клієнтів. Крім того, методи атрибуції вимагають якісних даних про весь ланцюг взаємодій, які не завжди доступні підприємствам. Тому для розв'язання поставленої задачі вони також не є оптимальними.

Отже, для подальшого розгляду було обрано дві ключові групи методів,

що відповідають вимогам дослідження та мають потенціал до інтеграції:

- класичні методи: когортний аналіз та базовий показник CLV як прозорий «опорний шар» для моніторингу, сегментації та бізнес-інтерпретації;
- прогнозні методи: методи прогнозування відтоку та розширені алгоритми обчислення CLV, які підвищують точність і дають змогу врахувати ризики втрати клієнтів.

Поєднання цих двох груп відкриває можливість створення нового підходу, який має переваги порівняно з існуючими методами завдяки інтеграції ретроспективного (когорти), стратегічного (CLV) та прогнозного (відтік) аналізу. Така комбінація буде детально розглянута у наступних підрозділах як основа для розробки комбінованого методу.

2.2 Класичні методи аналізу поведінки клієнтів

Когортний аналіз є одним із базових інструментів для вивчення поведінки клієнтів, який широко застосовується як у наукових дослідженнях, так і на практиці. Суть методу полягає в групуванні клієнтів у когорти за певною ознакою (наприклад, час залучення, канал комунікації чи соціально-демографічні характеристики) та подальшому відстеженні їхньої поведінки в динаміці. Це дає змогу оцінювати відмінності між групами та робити висновки щодо результативності маркетингових стратегій, рівня утримання чи потенційної цінності клієнтів.

У прикладних дослідженнях когортний підхід довів свою ефективність для прогнозування CLV. У дослідженні [26] показано, що використання когортного підходу разом із класичними моделями CLV дозволяє значно зменшити похибку оцінки поза вибіркою порівняно з агрегованими моделями. Автори відзначили, що моделі, побудовані окремо для різних когорт, краще відображають реальні закономірності споживчої поведінки та забезпечують

більш точні прогнози.

Додаткове підтвердження практичної цінності когортного аналізу наведено в дослідженні [27], яке присвячене сприйняттю цифрових платформ поколіннями X, Y та Z. Автори підкреслюють, що різні когортні групи відрізняються за своїми цінностями, рівнем цифрової грамотності, активністю у соціальних мережах і навіть ставленням до реклами. Наприклад, покоління X надає перевагу більш традиційним каналам комунікації та цінує інформативність, тоді як покоління Z є «мобільними нативами» та активно взаємодіє з мікровідеоплатформами та соціальними інфлюенсерами. Ця робота знову доводить, що без когортного підходу компанії ризикують втратити релевантність своїх маркетингових стратегій.

Практичне значення когортного аналізу полягає в тому, що він надає підприємствам інструмент для оцінки результативності маркетингових заходів та прогнозування майбутніх результатів на основі минулих тенденцій. Аналіз різних когорт клієнтів дає змогу визначити, які канали залучення дають стабільніших користувачів із високим рівнем утримання, а які, навпаки, приносять клієнтів із коротким життєвим циклом. Окрім цього, когортний аналіз часто застосовується як інтерпретаційний шар для складніших прогнозних моделей. У випадках, коли результати моделей машинного навчання є важкими для пояснення, побудова retention-кривих по когортах дає зрозумілий для менеджменту інструмент валідації. Це допомагає поєднувати прозорість класичних методів із точністю сучасних алгоритмів.

Водночас когортний підхід має суттєві обмеження (табл. 2.1). Він є переважно описовим методом і не забезпечує повноцінного прогнозу майбутньої поведінки клієнтів. Його результативність значною мірою залежить від коректності обраного критерію сегментації, адже неправильно сформовані когорти можуть призвести до викривлених висновків. Крім того, метод не враховує індивідуальних особливостей клієнтів, оскільки результати формуються для груп у цілому.

Таблиця 2.1 – Переваги та обмеження когортного аналізу

Критерій	Переваги	Обмеження
Простота застосування	Легко реалізується, не потребує складних алгоритмів	Дає обмежену кількість параметрів
Інтерпретація	Дає наочні retention-криві, зрозумілі для менеджерів	Не відображає індивідуальних відмінностей
Прогнозованість	Дає змогу оцінювати тенденції утримання клієнтів у часі	Не забезпечує точного прогнозу без інтеграції з іншими методами
Практична цінність	Використовується для оцінки каналів, поколінь, груп клієнтів	Ризик викривлень у разі некоректної сегментації або низької якості даних

Когортний аналіз є фундаментальним класичним методом, який забезпечує базове розуміння структури клієнтської бази та дає змогу пояснювати результати складніших моделей. Його головна цінність полягає у прозорості та простоті, однак обмеження цього підходу підкреслюють необхідність його інтеграції з іншими методами, такими як CLV та прогнозування відтоку.

Класичні CLV-моделі активно застосовуються і в практиці бізнесу. Вони дозволяють:

- ідентифікувати найбільш прибуткових клієнтів;
- прогнозувати обсяг доходів від клієнтської бази;
- планувати маркетингові витрати та програми лояльності;
- визначати доцільність залучення клієнтів через різні канали.

Водночас ці методи мають і певні обмеження (табл. 2.2). По-перше, вони чутливі до зовнішніх змін середовища (структурні злами, як-от пандемія COVID-19 [26], можуть призводити до систематичних похибок у прогнозах).

По-друге, класичні моделі часто працюють з усередненими показниками та не завжди здатні врахувати індивідуальні відмінності в поведінці клієнтів. По-третє, для отримання надійних результатів вони потребують якісних історичних даних, що не завжди є доступним.

Таблиця 2.2 – Переваги та обмеження методів обчислення CLV

Критерій	Переваги	Обмеження
Простота застосування	Доступні для реалізації, зрозумілі бізнес-користувачам	Потребують якісних історичних даних
Інтерпретація	Дають чіткий показник CLV	Усереднюють результати, не враховують індивідуальну поведінку
Прогнозованість	Дозволяють оцінювати довгострокові доходи клієнтів	Чутливі до структурних змін середовища, що призводить до зниження точності в нестабільних умовах
Практична цінність	Використовуються для сегментації, бюджетування та планування маркетингу	Обмежена здатність адаптуватися до швидких змін у поведінці клієнтів

Класичний CLV разом із когортним аналізом формують базовий рівень аналітики в дослідженнях, що стосуються поведінки клієнтів. Вони забезпечують інтерпретованість та стратегічну корисність, проте їхні прогностичні можливості залишаються обмеженими. Це підкреслює потребу в поєднанні з сучаснішими підходами, які будуть розглянуті в наступному підрозділі.

2.3 Прогнозні методи

Відтік клієнтів є одним із найбільш критичних показників для будь-якого бізнесу, оскільки втрата користувачів безпосередньо впливає на дохід і стабільність компанії. Прогнозування відтоку розглядається як один із ключових напрямів сучасної аналітики поведінки клієнтів. На відміну від класичних методів, що здебільшого описують історичні дані, прогнозні підходи дозволяють передбачати ймовірність того, що клієнт залишить компанію, і своєчасно застосовувати заходи для його утримання.

У роботі [28] увагу приділено не лише алгоритмічним аспектам прогнозування відтоку, а й практичним механізмам підвищення рівня утримання клієнтів. Автори підкреслюють, що успішність таких моделей визначається не тільки точністю прогнозів, а й тим, наскільки ефективно бізнес інтегрує їх результати у свою операційну діяльність. Важливим акцентом статті є роль мультиджерельних даних. Автори наголошують, що найбільш інформативними є моделі, які поєднують транзакційні дані з поведінковими показниками (частота візитів, взаємодія із застосунками) та даними цифрової взаємодії (дані email-розсилок, push-повідомлень і активності в соціальних мережах). Комплексність ознак дає змогу формувати не тільки більш точний прогноз, а й профіль клієнта, на основі якого можна будувати стратегічні рішення.

Окремо відзначено питання етики та прозорості моделей. Автори звертають увагу, що алгоритми прогнозування відтоку (особливо «чорні скриньки» на кшталт штучних нейронних мереж) часто є складними для пояснення менеджерам. Це створює ризик, коли рішення впроваджуються без достатнього розуміння причин, чому саме клієнта віднесено до «ризикової» групи. Дослідження підкреслює необхідність балансу між точністю та інтерпретованістю, що особливо важливо для невеликих компаній, які не мають великих аналітичних команд.

У роботі [29] акцент зроблено на малих та середніх підприємствах, які зазвичай мають обмежені ресурси й стикаються з проблемою недостатності даних. Авторка доводить, що навіть в таких умовах можна досягати високої точності прогнозування відтоку, використовуючи поєднання логістичної регресії з алгоритмами підкріплювального навчання для побудови персоналізованих кампаній. Отримані результати підтверджують, що прогнозні моделі дозволяють не лише виявляти клієнтів із високим ризиком відтоку, а й підвищувати їхню довгострокову цінність завдяки таргетованим заходам. Важливою частиною роботи стало використання аналізу виживання, який допоміг змодельовати тривалість CLC та скоригувати розрахунок CLV з урахуванням ризику відтоку. Це підкреслює тісний зв'язок між прогнозними моделями та оцінкою цінності клієнтів і демонструє перспективи інтегрованих підходів у бізнес-аналітиці (табл. 2.3).

Таблиця 2.3 – Прогнозні методи аналізу поведінки клієнтів

Метод	Переваги	Обмеження
Логістична регресія	Простота, інтерпретованість, швидкість	Менша точність на складних і нелінійних наборах даних
Random Forest, XGBoost	Висока точність, стійкість до шуму, робота з великими даними	Обмежена інтерпретованість, потреба у ресурсах
Штучні нейронні мережі	Добре працюють із часовими рядами та послідовними даними	«Чорна скринька», потребують великих обсягів даних
Аналіз виживання + CLV	Дає оцінку тривалості CLC, інтегрується з CLV	Потребує якісних даних і регулярного оновлення моделей

Отже, прогнозування відтоку клієнтів виступає одним із ключових

напрямів у групі прогнозних методів, оскільки дозволяє компаніям не лише завчасно ідентифікувати ризикові сегменти, а й формувати цілеспрямовані стратегії утримання. Особлива цінність цього підходу полягає у можливості поєднання його результатів із класичними методами CLV, що забезпечує коригування оцінки довгострокової цінності клієнтів з урахуванням імовірності їхнього відтоку. Таким чином формується більш реалістична та практично значуща модель, яка враховує як потенційні доходи, так і ризики втрати клієнтської бази.

2.4 Порівняльний аналіз методів

Вище було розглянуто дві основні групи методів аналізу поведінки клієнтів – класичні та прогнозні. Кожна з них має власні переваги та обмеження, які визначають можливості їхнього застосування в ІС для планування діяльності підприємства.

Класичні методи відзначаються простотою реалізації та високим рівнем інтерпретованості. Вони забезпечують зрозумілі для менеджменту результати, дають змогу оцінити динаміку утримання клієнтів і базовий показник CLV. Проте їхній основний недолік – обмежена прогнозованість, що знижує точність оцінок у мінливому бізнес-середовищі.

Прогнозні методи характеризуються вищою точністю та здатністю враховувати індивідуальну поведінку клієнтів. Вони дають змогу своєчасно виявляти ризикові сегменти та інтегрувати результати в розширені моделі CLV для отримання більш реалістичних прогнозів. Водночас ці методи є складнішими у впровадженні, вимагають якісних даних, а також потребують додаткових інструментів для забезпечення інтерпретованості.

Узагальнені результати аналізу наведено в табл. 2.4.

Таблиця 2.4 – Порівняння класичних і прогностичних методів аналізу поведінки клієнтів

Критерій	Класичні методи	Прогностичні методи
Простота застосування	Висока, легко реалізуються, зрозумілі для менеджерів	Нижча, потребують додаткових інструментів і технічних знань
Інтерпретація	Наочні графіки та зрозумілі показники	Обмежена у «чорних скриньках», вимагає пояснювальних моделей
Прогнозованість	Низька, описують минулі тенденції	Висока, дозволяють прогнозувати відтік і цінність клієнтів
Вимоги до даних	Середні (історичні транзакції, базова сегментація)	Високі (повні транзакційні, поведінкові, цифрові дані)
Практична цінність	Сегментація, ретроспективний аналіз	Утримання клієнтів, персоналізовані стратегії

Отже, класичні методи забезпечують прозорість і простоту інтерпретації, а прогностичні – точність і гнучкість. Водночас обидві групи не є взаємовиключними: їхнє поєднання дасть змогу сформувати більш повну картину поведінки клієнтів. Інтеграція когортного підходу, CLV та методу прогнозування відтоку створює основу для розробки комбінованого методу.

3 РОЗРОБКА КОМБІНОВАНОГО МЕТОДУ АНАЛІЗУ ПОВЕДІНКИ КЛІЄНТІВ

3.1 Вимоги до розроблюваного методу

Аналіз існуючих підходів до вивчення поведінки клієнтів показав, що кожна з груп методів має власні сильні сторони, проте має й обмеження, які ускладнюють їхнє застосування в ІС планування діяльності підприємства.

Класичні методи (когортний аналіз, базові моделі CLV) відзначаються простотою реалізації та високим рівнем інтерпретованості. Вони дозволяють оцінювати динаміку утримання клієнтів і розраховувати довгострокову цінність окремих сегментів. Проте, як показує дослідження [26], точність таких моделей знижується у випадках структурних змін попиту, а інше дослідження [27] підкреслює, що когортний підхід здебільшого є описовим і не забезпечує належної прогнозованості.

Прогнозні методи, зокрема моделі відтоку на основі машинного навчання, забезпечують високу точність і дозволяють завчасно ідентифікувати клієнтів із високим ризиком втрати. У роботі [28] показано, що використання багатоканальних даних дає змогу значно підвищити якість прогнозів. Проте автори наголошують на обмеженій інтерпретованості складних алгоритмів, що ускладнює їх упровадження в практичну діяльність компаній. Крім того, дослідники відзначають, що для малих і середніх підприємств характерна нестача якісних даних, через що навіть точні моделі потребують статистичних підходів, здатних працювати зі скороченими вибірками. До таких належить аналіз виживання, який забезпечує можливість врахування часової динаміки та неповноти спостережень і, таким чином, розширює застосовність прогнозних моделей у практиці бізнесу.

Отже, маємо ситуацію, коли класичні підходи є зрозумілими та зручними, але малопрогностичними, тоді як прогнозні – точні, проте складні та ресурсомісткі. Це створює розрив між потребами бізнесу та можливостями

аналітичних інструментів. У зв'язку з цим виникає практична проблема: підприємствам потрібен метод, який забезпечує одночасно інтерпретованість для управлінців і достатню точність для прийняття рішень у мінливих умовах.

На підставі цього можна сформулювати вимоги до розроблюваного методу. Він повинен:

- використовувати фактичні дані підприємства (транзакційні, поведінкові);
- забезпечувати аналіз як на рівні груп клієнтів (когорт), так і на рівні окремих записів;
- надавати можливість отримувати результати у зрозумілому для бізнесу вигляді (через retention-криві, середню тривалість взаємодії та прогнозовану цінність клієнта);
- бути відтворюваним на доступних наборах даних без необхідності використання складних інструментів машинного навчання;
- мати можливість інтеграції в аналітичні модулі ІС підприємства.

Метою розробки є створення комбінованого методу, заснованого на когортному аналізі, що доповнюється оцінкою CLV та коригується шляхом прогнозування відтоку. Використання комбінації цих методів має забезпечити більш реалістичну оцінку клієнтської бази, підвищити точність прогнозування та водночас залишатися зрозумілою для користувачів ІС для планування діяльності підприємства.

3.2 Основні етапи методу

Комбінований метод аналізу поведінки клієнтів поєднує можливості когортного аналізу, розрахунку CLV та прогнозування відтоку. Таке поєднання дасть змогу оцінити клієнтську базу не лише з погляду отриманих у минулому доходів, а й врахувати імовірності майбутньої активності клієнтів,

тобто зробити оцінку більш реалістичною. Мета розробки методу полягає у створенні інструменту, який дозволить підприємству краще планувати свою діяльність, прогнозувати доходи та розробляти дієві стратегії утримання клієнтів.

Метод можна представити у вигляді такої послідовності етапів.

1. Усі клієнти групуються за певною ознакою (наприклад, місяць першої покупки) у когорти. Для кожної сформованої когорти будується retention-крива, яка показує, як змінюється кількість активних клієнтів у часі. Це дасть змогу виявити закономірності утримання клієнтів і визначити характер поведінки різних груп.

2. Обчислюється базовий показник CLV. Він відображає очікуваний сумарний дохід від кожної когорти протягом обраного періоду прогнозування.

3. Виконується прогнозування ймовірності втрати клієнтів на основі поведінкових характеристик. Отримана функція виживання відображає, з якою ймовірністю клієнти певної когорти залишатимуться активними протягом певного часу.

4. Базовий показник CLV коригується з урахуванням прогнозованої ймовірності відтоку, завдяки чому отримується ризик-скоригований показник CLV.

Запропонований комбінований метод реалізується у вигляді модуля аналізу поведінки клієнтів ІС планування діяльності підприємства.

Для опису логіки роботи модуля застосовано методологію структурного моделювання Integration Definition for Function Modeling (IDEF0) [30]. Ця нотація використовується для формалізованого зображення функціональної структури системи, опису потоків даних, управлінських впливів і ресурсів, необхідних для виконання процесів.

Основним елементом методології IDEF0 є функціональний блок, який відображає окрему операцію або підпроцес.

Кожен блок має чотири типи стрілок:

- входи (input) – дані або об'єкти, що обробляються у процесі;

- керування (control) – правила, параметри, нормативи, які визначають, як виконується процес;
- механізми (mechanism) – ресурси, засоби та виконавці, що забезпечують виконання процесу;
- виходи (output) – результати роботи або продукти, які утворюються після виконання функції.

Згідно з методологією IDEF0, побудова моделі відбувається поетапно:

- контекстна діаграма (A0) – подає загальний огляд системи, її межі, основні вхідні та вихідні потоки;
- декомпозиція першого рівня (A1) – деталізує головний процес, розкриваючи його на підпроцеси, між якими встановлюються логічні зв'язки.

У межах роботи за допомогою методології IDEF0 розроблено контекстну діаграму модуля аналізу поведінки клієнтів ІС планування діяльності підприємства, який реалізує запропонований комбінований метод, та декомпозицію її першого рівня – рис. 3.1–3.2.



Рисунок 3.1 – Контекстна діаграма модуля аналізу поведінки клієнтів

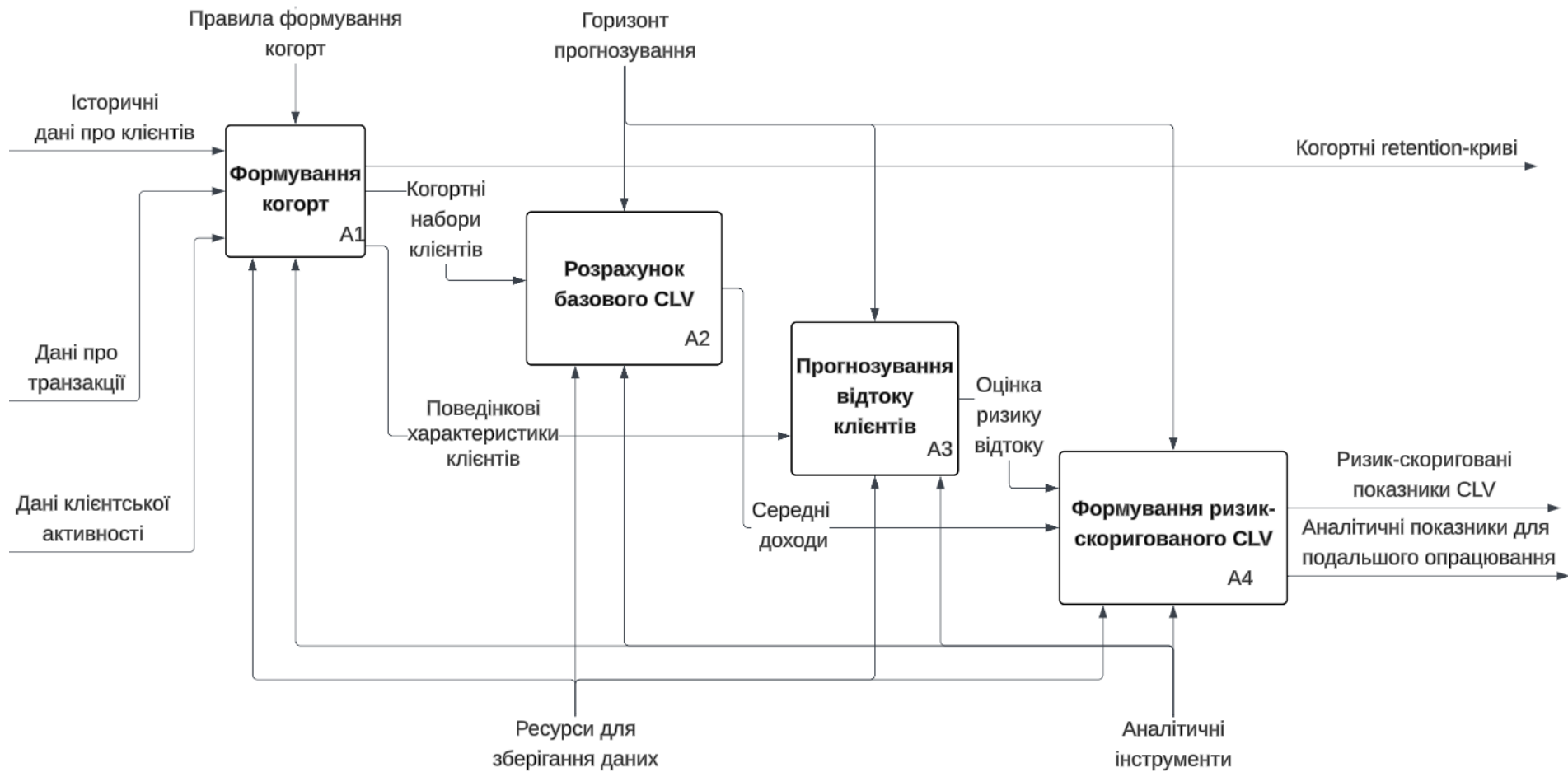


Рисунок 3.2 – Декомпозиція першого рівня

На рис. 3.2 наведено декомпозицію контекстної діаграми A0, що відображає основні етапи роботи аналітичного модуля. Кожен із підпроцесів A1–A4 виконує окрему функцію: формування когорт клієнтів, розрахунок базового показника CLV, прогнозування ймовірності відтоку та формування ризик-скоригованого показника CLV. Між підпроцесами передаються проміжні результати у вигляді когортних даних, поведінкових характеристик і фінансових показників. Спільні компоненти керування та механізми забезпечують функціонування всіх етапів процесу.

Для реалізації комбінованого методу використовуються дані клієнтської бази підприємства, що надходять до аналітичного модуля ІС. До складу вхідних потоків належать:

- історичні дані про клієнтів – ідентифікатори користувачів та інформація про дату першої покупки;
- транзакційні дані – дата та сума кожної покупки;
- дані клієнтської активності – поведінкові характеристики, сформовані на основі транзакційних записів (частота покупок, середній інтервал між транзакціями та інші ознаки взаємодії клієнта з підприємством).

Зазначені дані надходять з корпоративних джерел чи журналів подій, формуючи єдину аналітичну вибірку, що використовується на подальших етапах.

3.3 Формальний опис етапів комбінованого методу

Архітектура комбінованого методу являє собою, як зазначено вище, чітко визначену послідовність етапів (рис. 3.3), кожен з яких відіграє ключову роль у забезпеченні високої якості результатів.

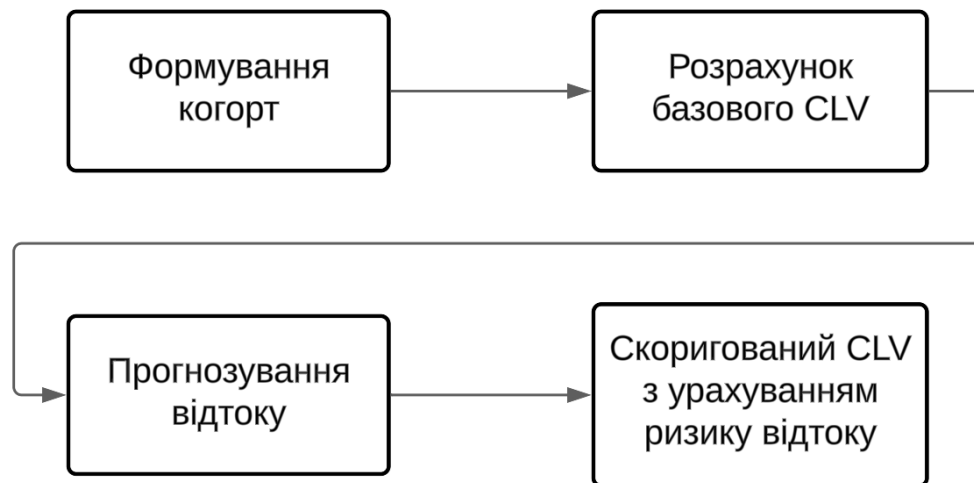


Рисунок 3.3 – Основні етапи комбінованого методу аналізу поведінки клієнтів

Розглянемо основні етапи запропонованого методу більш детально.

На першому етапі комбінованого методу здійснюється формування когорт клієнтів – груп користувачів, об’єднаних за спільною ознакою. Найчастіше когорта визначається за місяцем першої покупки або реєстрації, що дає змогу відстежити динаміку активності клієнтів упродовж часу.

Для кожного клієнта визначається дата першої транзакції за формулою:

$$first_transaction_u = \min(date_transaction_u), \quad (3.1)$$

де u – унікальний ідентифікатор клієнта (номер облікового запису або ID у базі даних клієнтів);

$date_transaction_u$ – дата здійснення кожної транзакції клієнта;

$first_transaction_u$ – дата першої покупки клієнта.

На практиці це виглядає таким чином. Якщо клієнт здійснив покупки 12.01.2025, 02.02.2025 та 15.03.2025, то $first_transaction_u = 12.01.2025$.

Відповідно, клієнт належить до когорти, що визначається місяцем першої транзакції, за формулою:

$$\text{cohort}(u) = \text{month}(\text{first_transaction}_u), \quad (3.2)$$

де $\text{cohort}(u)$ – функція, що визначає належність клієнта до певної когорти;

$\text{month}()$ – функція, що повертає місяць з дати;

$\text{first_transaction}_u$ – дата першої покупки клієнта.

Отже, якщо перша покупка відбулась 12.01.2025, тоді $\text{cohort}(u) = 01.2025$, тобто клієнт належить до когорти січня 2025 року.

Для подальших розрахунків позначимо через $c = \text{cohort}(u)$ когорту, до якої належить клієнт u .

Для кожної когорти c і періоду t обчислюються наступні показники:

– N_c – кількість клієнтів у когорті на момент її формування, тобто кількість користувачів, які вперше здійснили покупку або зареєструвалися в певному періоді;

– $A_c(t)$ – кількість клієнтів когорти, які залишаються активними у періоді t ;

– $R_c(t)$ – частка утриманих клієнтів, тобто коефіцієнт утримання, у когорті c на момент часу t .

Показник $R_c(t)$ визначається за формулою:

$$R_c(t) = \frac{A_c(t)}{N_c}, \quad t = 0, 1, 2, \dots, T, \quad (3.3)$$

де T – тривалість періоду спостереження, кількість місяців;

N_c – кількість клієнтів у когорті на момент її формування, осіб;

$A_c(t)$ – кількість клієнтів когорти, які залишаються активними у періоді t , осіб.

Показник $R_c(t)$ відображає частку утриманих клієнтів (коефіцієнт утримання) у когорті в момент часу t . Він може задаватися у вигляді частки від 0 до 1 або у відсотках (від 0% до 100%). Значення 1 означає, що всі клієнти когорти залишаються активними, тоді як 0 свідчить про повну втрату клієнтів.

Після визначення когорт дані агрегуються у вигляді когортної матриці, у якій рядки відповідають когортам, а стовпці – часовим періодам (віку когорти). Приклад когортної матриці наведено на рис. 3.4.

Місяць першої покупки (когорта)	Кількість користувачів	Місяць 1	Місяць 2	Місяць 3	Місяць 4	Місяць 5	Місяць 6
Січень 2025	1 245	100%	86%	79%	73%	69%	65%
Лютий 2025	1 184	100%	85%	77%	71%	67%	64%
Березень 2025	1 326	100%	84%	78%	72%	68%	64%
Квітень 2025	1 274	100%	83%	76%	71%	67%	63%
Травень 2025	1 352	100%	82%	75%	70%	65%	62%
Червень 2025	1 198	100%	83%	76%	71%	66%	63%
Липень 2025	1 286	100%	82%	75%	69%	65%	62%
Серпень 2025	1 314	100%	81%	74%	69%	64%	61%
Вересень 2025	1 177	100%	80%	73%	68%	63%	60%
Жовтень 2025	1 221	100%	79%	72%	67%	62%	59%
Листопад 2025	1 192	100%	78%	71%	66%	61%	58%
Грудень 2025	1 315	100%	77%	70%	65%	60%	57%
Усього	15 084	100%	82%	75%	69%	65%	62%

Рисунок 3.4 – Когортна матриця

Retention-крива – додатковий спосіб візуалізації когорт, який дає змогу оцінити, наскільки швидко відбувається відтік користувачів після першого контакту, а також порівняти різні когорти між собою та виявити закономірності.

Приклад графіку retention-кривих за даними трьох когорт з рис. 3.4 наведено на рис. 3.5.

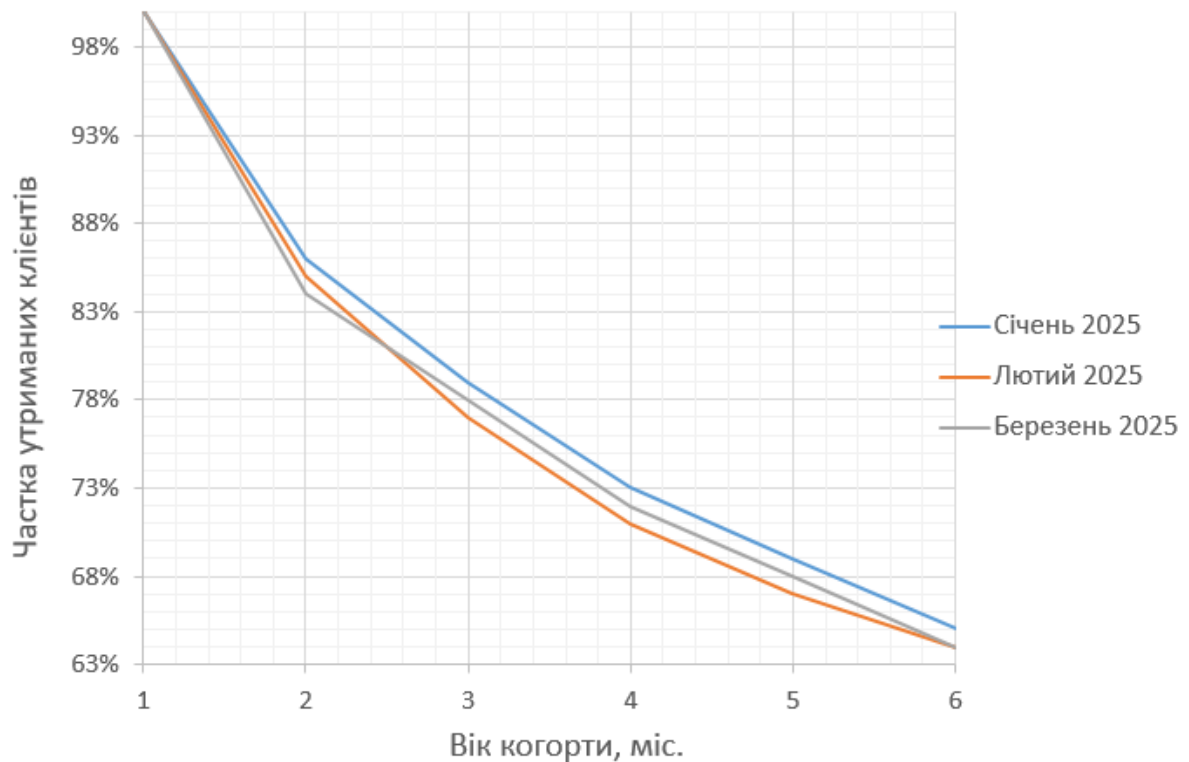


Рисунок 3.5 – Графік retention-кривих для когорт «Січень 2025», «Лютий 2025» та «Березень 2025»

На другому етапі комбінованого методу обчислюється базовий показник CLV. Цей показник використовується для оцінювання очікуваного прибутку, який підприємство може отримати від клієнта або групи клієнтів за певний період взаємодії.

У межах дослідження розрахунок проводиться на рівні когорт, сформованих за ознакою місяця першої покупки. Для кожної когорти c і періоду t визначаються:

- кількість клієнтів $A_c(t)$, які залишаються активними в кожному часовому періоді t ;
- частка утриманих клієнтів $R_c(t)$, що розрахована за формулою (3.3);
- середній дохід $m_c(t)$ від одного активного клієнта когорти c у період t , який обчислюється за формулою:

$$m_c(t) = \frac{Income_c^{act}(t)}{A_c(t)}, \quad (3.4)$$

де $Income_c^{act}(t)$ – сумарний дохід, отриманий від активних клієнтів когорти c у періоді t , грн;

$A_c(t)$ – кількість клієнтів когорти c , що залишаються активними у періоді t , осіб.

Базовий показник CLV когорти c визначається за формулою:

$$CLV_c^{base} = \sum_{t=1}^T \frac{R_c(t) * m_c(t)}{(1 + d)^t}, \quad (3.5)$$

де T – горизонт прогнозування, кількість періодів;

$R_c(t)$ – частка утриманих клієнтів у когорті c на момент часу t , безрозмірна величина;

$m_c(t)$ – середній дохід від одного активного клієнта когорти c у період t , грн/період;

d – ставка дисконтування, що враховує зміну вартості грошей у часі, вимірюється у частках від одиниці або у відсотках за період.

Для демонстрації розрахунків розглянемо когорту клієнтів, сформовану у січні 2025 року (рис. 3.4). Припустимо, що середній дохід від одного активного клієнта становить $m_c(t) = 100$ грн/місяць, а ставка дисконту дорівнює $d = 0,05$ (5% на місяць). Показники утримання для перших п'яти місяців становлять $R_c(t) = [0,86; 0,79; 0,73; 0,69; 0,65]$.

Тоді базовий показник CLV на одного клієнта згідно з формулою (3.5):

$$CLV_{cic}^{base} = \frac{0,86 * 100}{(1 + 0,05)^1} + \frac{0,79 * 100}{(1 + 0,05)^2} + \frac{0,73 * 100}{(1 + 0,05)^3} + \frac{0,69 * 100}{(1 + 0,05)^4} + \frac{0,65 * 100}{(1 + 0,05)^5} = 324,32. \quad (3.6)$$

Отже, очікувана CLV одного клієнта цієї когорти становить близько 324 грн, а для всієї когорти з 1245 клієнтів – приблизно 403,8 тис. грн.

Базовий показник CLV відображає лише історичну та прогнозовану динаміку доходів, не враховуючи ризиків втрати клієнтів. На наступному етапі метод доповнюється компонентом прогнозування ймовірності відтоку, що дає змогу скоригувати базову оцінку.

На третьому етапі здійснюється прогнозування ймовірності відтоку клієнтів. Для реалізації цього етапу комбінованого методу обрано метод аналізу виживання, що ґрунтується на використанні моделі пропорційних ризиків Кокса [31]. Вибір цього методу зумовлений специфікою поставленої задачі, а саме необхідністю враховувати не лише факт відтоку клієнта, а й час, через який він відбувається. Більшість традиційних методів прогнозування, таких як логістична регресія чи дерева рішень, дозволяють визначити лише ймовірність того, що клієнт припинить взаємодію з компанією, не враховуючи часову динаміку цього процесу.

Натомість аналіз виживання спеціально розроблений для задач оцінювання часу до настання події, де важливо моделювати не лише сам факт події, а тривалість до її виникнення. У контексті поведінки клієнтів це означає можливість описати, як змінюється ймовірність відтоку у часі.

Важливою перевагою survival-аналізу є здатність працювати з цензурованими спостереженнями, тобто з тими клієнтами, які на момент аналізу залишаються активними. У звичайних класифікаційних моделях такі випадки або відкидаються, або спотворюють результати, тоді як survival-підхід дає змогу коректно їх врахувати, зберігаючи повноту вибірки. Серед різних варіантів реалізації survival-аналізу обрано саме модель Кокса, яка поєднує обґрунтовану математичну основу з практичною інтерпретованістю. Вона дає змогу не лише оцінювати загальну функцію виживання клієнтів, а й визначати, як конкретні поведінкові фактори впливають на ризик відтоку.

Функція виживання $S_c(t)$ описуватиме ймовірність того, що клієнт із когорти c залишатиметься активним до моменту часу t . Це значення надалі

використовується для коригування базового показника CLV, даючи змогу врахувати ризик відтоку та зробити оцінку більш реалістичною.

Для оцінювання ймовірності відтоку використано математичну модель пропорційних ризиків Кокса, яка дозволяє врахувати вплив поведінкових характеристик клієнтів на ризик припинення взаємодії з підприємством.

Функція ризику для клієнта u має вигляд:

$$h_u(t) = h_0(t) * e^{\beta_1 * x_{u1} + \beta_2 * x_{u2} + \dots + \beta_k * x_{uk}}, \quad (3.7)$$

де $h_u(t)$ – інтенсивність відтоку клієнта u у момент часу t , частка/період;

$h_0(t)$ – базова інтенсивність ризику відтоку, спільна для всієї вибірки клієнтів;

β_i – коефіцієнти моделі, які показують вплив кожного поведінкового фактора на ризик відтоку, безрозмірна величина;

x_{ui} – поведінкові характеристики клієнта u (частота покупок, середній чек, тривалість взаємодії тощо), власні одиниці або нормовані в діапазоні $[0; 1]$;

k – кількість поведінкових факторів, включених у модель;

t – час або порядковий номер періоду спостереження (наприклад, місяць, тиждень, день після першої покупки).

Значення коефіцієнта β_i інтерпретується так:

- якщо $\beta_i > 0$, то фактор підвищує ризик відтоку;
- якщо $\beta_i < 0$, то фактор знижує ризик відтоку та свідчить про вищу лояльність клієнта.

Функція виживання $S_u(t)$ показує, з якою ймовірністю клієнт залишиться активним до моменту часу t , і має вигляд:

$$S_u(t) = [S_0(t)]^{e^{\beta_1 * x_{u1} + \beta_2 * x_{u2} + \dots + \beta_k * x_{uk}}}, \quad (3.8)$$

де $S_u(t)$ – функція виживання для клієнта u , що відображає ймовірність залишатися активним до моменту часу t , безрозмірна величина ($0 \leq S_u(t) \leq 1$);

$S_0(t)$ – базова функція виживання для «середнього» клієнта, яка оцінюється статистично за допомогою методів Каплана-Мейєра або Бреслоу [32], безрозмірна величина ($0 \leq S_0(t) \leq 1$);

β_i – коефіцієнти моделі, що показують вплив кожного поведінкового фактора на ризик відтоку, безрозмірна величина;

x_{ui} – поведінкові характеристики клієнта u (частота покупок, середній чек, тривалість взаємодії тощо), власні одиниці або нормовані в діапазоні $[0; 1]$;

k – кількість поведінкових факторів, включених у модель;

t – час або порядковий номер періоду спостереження (наприклад, місяць, тиждень, день після першої покупки).

Функція $S_u(t)$ є спадною в часі: чим більше значення t , тим менше значення ймовірності утримання клієнта. Це відображає природну тенденцію до зростання ризику відтоку з часом.

Експоненційний вираз $e^{\beta_1 * x_{u1} + \beta_2 * x_{u2} + \dots + \beta_k * x_{uk}}$ модифікує форму базової кривої $S_0(t)$ відповідно до індивідуальних характеристик клієнта:

– якщо $e^{\beta * x_u} > 1$, крива спадає швидше – клієнт має підвищений ризик відтоку;

– якщо $e^{\beta * x_u} < 1$, крива спадає повільніше – клієнт є більш лояльним.

Розберемо приклад обчислення індивідуальної ймовірності утримання клієнтів на основі моделі пропорційних ризиків Кокса – табл. 3.1.

Таблиця 3.1 – Приклад обчислення індивідуальної ймовірності утримання клієнтів

Клієнт	x_1 , частота покупок	x_2 , місяці неак- тивності	β_1	β_2	$e^{\beta_1 x_1 + \beta_2 x_2}$	$S_0(3)$	$S_u(3)$	Ймовірність утримання
1	5	1	-0,4	0,6	0,247	0,7	0,916	91,6%
2	4	2	-0,4	0,6	0,670	0,7	0,787	78,73%

Продовження таблиці 3.1

Клієнт	x_1 , частота покупок	x_2 , місяці неак- тивності	β_1	β_2	$e^{\beta_1 x_1 + \beta_2 x_2}$	$S_0(3)$	$S_u(3)$	Ймовірність утримання
3	4	1	-0,4	0,6	0,368	0,7	0,877	87,70%
4	3	3	-0,4	0,6	1,822	0,7	0,522	52,21%
5	4	5	-0,4	0,6	4,055	0,7	0,235	23,54%

Для демонстрації використано два поведінкові фактори: x_1 – частота покупок, яка знижує ризик відтоку, та x_2 – тривалість неактивності, яка, навпаки, його підвищує. Коефіцієнти моделі прийнято рівними: $\beta_1 = -0,4$ та $\beta_2 = 0,6$. Базова функція виживання для «середнього» клієнта через три місяці становить $S_0(3) = 0,7$. Для кожного клієнта розраховано значення експоненти $e^{\beta_1 x_1 + \beta_2 x_2}$, яка визначає індивідуальний рівень ризику, а також відповідну функцію виживання $S_u(3)$.

Як видно з результатів, наведених у табл. 3.1, ймовірність утримання зменшується зі зниженням частоти покупок і зі збільшенням періоду неактивності. Наприклад, клієнт №1, який здійснює 5 покупок на місяць і був неактивним лише 1 місяць, має прогноз ймовірності утримання 91,6%, а клієнт 5, який купує 4 рази, але був неактивним протягом 5 місяців, має значно нижчу ймовірність утримання – 23,54%.

Для всієї когорти клієнтів функція виживання обчислюється як середнє значення:

$$S_c(t) = \frac{1}{N_c} \sum_{u \in c} S_u(t), \quad (3.9)$$

де $S_c(t)$ – середня функція виживання для всієї когорти клієнтів c у момент часу t , безрозмірна величина ($0 \leq S_c(t) \leq 1$);

N_c – кількість клієнтів у когорті c , осіб;

$S_u(t)$ – індивідуальна функція виживання для кожного клієнта u у когорті c , безрозмірна величина ($0 \leq S_u(t) \leq 1$);

t – період спостереження.

Формула (3.9) відображає перехід від індивідуального прогнозу $S_u(t)$ до агрегованого показника для всієї групи – $S_c(t)$. Вона визначає, яка частка клієнтів певної когорти, у середньому, залишиться активною до моменту часу t . Оскільки $S_u(t)$ є ймовірністю, то й $S_c(t)$ також є безрозмірною величиною в діапазоні $[0; 1]$.

Використаємо дані з попередньої таблиці для п'яти клієнтів (табл. 3.1). Індивідуальні ймовірності утримання за три місяці становлять: $S_u(3) = [0,916; 0,787; 0,877; 0,522; 0,235]$. Тоді середня функція виживання для когорти розраховується як:

$$S_c(3) = \frac{1}{5} (0,916 + 0,787 + 0,877 + 0,522 + 0,235) = 0,667. \quad (3.10)$$

Отже, середня ймовірність того, що клієнти цієї когорти залишаться активними через три місяці, становить приблизно 66,7%.

На четвертому етапі розраховується скоригований показник CLV. На відміну від базового значення CLV, який враховує лише історичні дані про утримання клієнтів $R_c(t)$, скоригований показник CLV відображатиме прогнозовану поведінку клієнтів з урахуванням ризику відтоку.

Формула ризик-скоригованого показника CLV матиме вигляд:

$$CLV_c^{adj} = \sum_{t=1}^T \frac{S_c(t) * m_c(t)}{(1 + d)^t}, \quad (3.11)$$

де $S_c(t)$ – прогнозована ймовірність того, що клієнти когорти c залишатимуться активними у періоді t , безрозмірна величина ($0 \leq S_c(t) \leq 1$);

$m_c(t)$ – середній дохід від одного активного клієнта когорти c у період t , грн/період;

d – ставка дисконтування, що враховує зміну вартості грошей у часі, вимірюється у частках від одиниці або у відсотках за період;

T – горизонт прогнозування, кількість періодів.

Для демонстрації застосування ризик-скоригованої моделі розглянемо когорту клієнтів, для якої середня функція виживання через три місяці становить $S_c(3) = 0,67$, як було отримано у попередньому прикладі. Припустимо, що середній місячний дохід від одного активного клієнта дорівнює $m_c(t) = 100$ грн, ставка дисконту $d = 0,05$ (5% на місяць), а горизонт прогнозування $T = 5$ місяців.

Для демонстрації використано умовний ряд значень $S_c(t)$, що плавно зменшується в часі та відображає типову динаміку скорочення активності клієнтів: $S_c(t) = [0,87; 0,79; 0,67; 0,62; 0,60]$. Для кожного періоду t обчислюється очікуваний дохід за формулою (3.11). Результати обчислень наведено в табл. 3.2.

Таблиця 3.2 – Результат обчислення ризик-скоригованого показника CLV

t (місяць)	$S_c(t)$	$(1 + d)^t$	$\frac{S_c(t) * m_c(t)}{(1+d)^t}$, грн
1	0,87	1,05	82,86
2	0,79	1,10	71,66
3	0,67	1,16	57,88
4	0,62	1,22	51,01
5	0,6	1,28	47,01
Сума			310,41

У результаті отримано $CLV_c^{adj} \approx 310$ грн, тоді як базова оцінка без урахування відтоку становить $CLV_c^{base} \approx 324$ грн. Зменшення значення

пояснюється тим, що частка клієнтів має високий ризик відтоку, що зменшує очікувані майбутні прибутки. Такий приклад демонструє, як врахування функції виживання дає можливість отримати більш реалістичну та економічно обґрунтовану довгострокову цінність клієнтів.

Отже, подвійний розрахунок CLV має різне призначення на кожному етапі комбінованого методу. Спочатку обчислюється базовий показник, який відображає історичну дохідність клієнтської бази без урахування ризиків.

Далі проводиться коригування показника CLV з інтеграцією функції виживання. Вона використовується як ваговий коефіцієнт, що коригує майбутні грошові потоки відповідно до прогнозованої ймовірності утримання клієнтів. Такий підхід дає змогу отримати більш реалістичну оцінку довгострокової цінності, що враховує не лише обсяг доходу від клієнта, а й ризик його втрати в часі.

Такий підхід забезпечує зв'язок між історичною фінансовою оцінкою та прогнозною моделлю поведінки клієнтів. Це дає змогу не лише порівняти отримані результати, а й оцінити, як зміни у поведінці клієнтів впливають на фактичну прибутковість підприємства, та визначити більш точні стратегії їх утримання.

3.4 Інтеграція методу в ІС та очікувані результати

Розроблений комбінований метод аналізу поведінки клієнтів інтегрується в ІС планування діяльності підприємства як окремий аналітичний модуль. Він забезпечує обробку даних про клієнтів, розрахунок показників CLV та оцінювання ризику відтоку. Основною метою інтеграції є підвищення аналітичних можливостей ІС, формування комплексного розуміння клієнтської поведінки та отримання достовірних прогнозів для підтримки управлінських рішень.

Як показано на рис. 3.2, модуль отримує дані з кількох джерел: бази даних клієнтів, транзакційного журналу та системи моніторингу активності. На виході формується набір аналітичних результатів, що включає когортні retention-криві та ризик-скориговані показники CLV. Ці результати можуть використовуватися для побудови візуалізацій, створення аналітичних звітів або формування узагальнених показників у межах ІС.

Очікуваний результат інтеграції полягає у розширенні функціональних можливостей ІС: система зможе не лише фіксувати поточні показники діяльності, а й прогнозувати поведінку клієнтів у майбутніх періодах. Застосування комбінованого методу дасть змогу формувати більш точні прогнози доходів, своєчасно ідентифікувати ризикові сегменти клієнтської бази та визначати найбільш прибуткові когорти для подальшої роботи.

Інтеграція розглядається на концептуальному рівні та демонструє логічну взаємодію розробленого методу з іншими підсистемами ІС. У наступному розділі наведено приклад практичної реалізації комбінованого методу та проведено аналіз отриманих результатів.

4 ЕКСПЕРИМЕНТАЛЬНА ПЕРЕВІРКА ПРАЦЕЗДАТНОСТІ РОЗРОБЛЕНОГО КОМБІНОВАНОГО МЕТОДУ АНАЛІЗУ ПОВЕДІНКИ КЛІЄНТІВ

4.1 Приклад застосування розробленого комбінованого методу

Для демонстрації практичного застосування розробленого комбінованого методу аналізу поведінки клієнтів використано транзакційні дані умовної компанії. У наборі даних містяться:

- унікальний ідентифікатор клієнта;
- дата здійснення транзакції;
- сума покупки.

Такий формат даних є типовим для CRM- та e-commerce-систем і забезпечує можливість проведення повноцінного когортного та поведінкового аналізу, необхідного для реалізації комбінованого методу.

Відповідно до методики когортного аналізу, описаної в розділі 3, для кожного клієнта було визначено дату першої покупки за формулою (3.1). Серед сформованих когорт для подальшого аналізу обрано когорту лютого 2024 року. Такий вибір обґрунтований такими чинниками:

- когорта має оптимальний розмір, що дозволяє отримувати стабільні статистичні оцінки;
- період спостереження охоплює майже повний рік, що дає достатній горизонт для розрахунку показників $R(t)$, $m(t)$, CLV_{base} та побудови survival-моделі;
- поведінка клієнтів не містить виражених сезонних коливань, що підвищує репрезентативність отриманих результатів.

На наступному етапі комбінованого методу здійснюється аналіз динаміки утримання клієнтів. За формулою (3.3) розраховано частку утриманих клієнтів когорти. Як було зазначено вище, показник $R(t)$ виражається у частках і змінюється в діапазоні від 0 до 1 (або від 0% до 100%).

Результати обчислень наведено в табл. 4.1.

Таблиця 4.1 – Активні клієнти та частка утримання $R(t)$ когорти

Місяць віку когорти	Кількість активних клієнтів	Частка утриманих клієнтів
2024-02	163	1,00
2024-03	54	0,33
2024-04	75	0,46
2024-05	59	0,36
2024-06	68	0,42
2024-07	63	0,39
2024-08	56	0,34
2024-09	61	0,37
2024-10	56	0,34
2024-11	67	0,41
2024-12	73	0,45

Для порівняння динаміки утримання між різними групами клієнтів було побудовано retention-криві для трьох різних когорт: січня, лютого та березня 2024 року (рис. 4.1).

Згідно з рис. 4.1, для всіх когорт характерним є різкий спад активності після першого місяця. Це є типовою поведінковою моделлю: більшість клієнтів демонструють найвищу активність у момент першої взаємодії з компанією, після чого інтенсивність покупок поступово знижується. Починаючи з другого місяця значення утримання стабілізуються в інтервалі 0,30–0,45, що свідчить про формування групи постійних клієнтів, які продовжують взаємодію з компанією протягом кількох місяців поспіль.

Порівняння трьох когорт показує, що характер динаміки утримання є подібним: після початкового спаду криві мають близькі траєкторії, без суттєвих розбіжностей між когортами. Це дає підстави стверджувати, що

сезонні чинники або зовнішні умови не мали значного впливу на поведінку клієнтів протягом аналізованого періоду.

Отже, побудовані retention-криві підтверджують стабільність поведінкової моделі клієнтів та можуть бути використані як вхідні дані для подальших розрахунків середнього доходу $m(t)$, базового показника CLV та побудови survival-моделі.

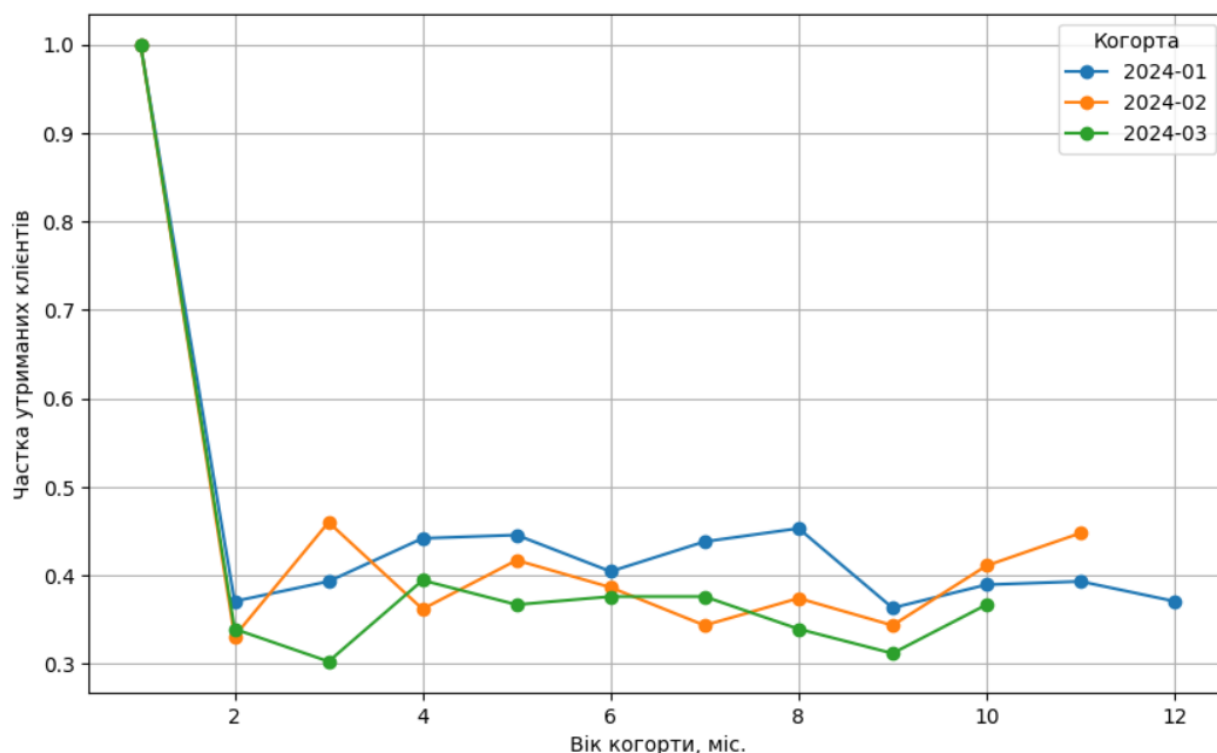


Рисунок 4.1 – Графік retention-кривих для когорт «2024-01», «2024-02» та «2024-03»

На другому етапі визначається показник середнього доходу на одного активного клієнта за формулою (3.4) та обчислюється базовий показник CLV за формулою (3.5). У межах дослідження ставку дисконту прийнято $d = 10\%$. Результати розрахунків наведено в табл. 4.2.

Таблиця 4.2 – Результати розрахунків базового показника CLV

Місяць віку когорти	t	Кількість активних клієнтів	Частка R(t)	Дохід m(t)	Множник $\frac{1}{(1+d)^t}$	Базовий показник CLV
2024-02	1	163	1,00	658,82	0,91	598,92
2024-03	2	54	0,33	662,57	0,83	181,41
2024-04	3	75	0,46	693,04	0,75	239,58
2024-05	4	59	0,36	589,46	0,68	145,73
2024-06	5	68	0,42	685,53	0,62	177,58
2024-07	6	63	0,39	687,23	0,56	149,93
2024-08	7	56	0,34	661,74	0,51	116,67
2024-09	8	61	0,37	692,73	0,47	120,94
2024-10	9	56	0,34	640,84	0,42	93,37
2024-11	10	67	0,41	641,32	0,39	101,63
2024-12	11	73	0,45	712,31	0,35	111,81

Сумарний базовий показник CLV когорти лютого 2024 року:

$$CLV_{2024-02}^{base} = \sum_{t=1}^T CLV_T = 2037,58 \text{ грн.} \quad (4.1)$$

На третьому етапі будується survival-модель, яка використовує наступні поведінкові змінні:

- frequency – частота транзакцій клієнта;
- avg_purchase_norm – нормований середній чек;
- avg_interval – середній інтервал між покупками.

Значення коефіцієнтів моделі прийняті такі:

- $\beta_{frequency} = -0,05$;
- $\beta_{avg_purchase_norm} = -0,12$;

– $\beta_{\text{avg_interval}} = 0,04$.

Значення коефіцієнтів обрано відповідно до очікуваного впливу поведінкових змінних на ризик відтоку:

– знак мінуса перед частотою покупок ($\beta_{\text{frequency}} = -0,05$) означає, що активні клієнти мають нижчу ймовірність припинення взаємодії;

– вищий середній чек також зменшує ризик втрати клієнта ($\beta_{\text{avg_purchase_norm}} = -0,12$);

– позитивний коефіцієнт при середньому інтервалі між покупками ($\beta_{\text{avg_interval}} = 0,04$) означає, що великі паузи між транзакціями є індикатором зростання ризику відтоку.

Індивідуальний показник інтенсивності ризику відтоку обчислюється за формулою:

$$\alpha_u = e^{\beta_{\text{frequency}} * x_1 + \beta_{\text{avg_purchase_norm}} * x_2 + \beta_{\text{avg_interval}} * x_3}. \quad (4.2)$$

Survival-функція клієнта обчислюється за формулою:

$$S_u(t) = S_0^{\alpha_u}, \quad S_0 = 0,80. \quad (4.3)$$

Середнє значення survival-функції для когорти становить $S_c = 0,44$. Це означає, що клієнт когорти буде залишатися активним у найближчому періоді з імовірністю 44%.

На останньому етапі обчислюється ризик-скоригований показник CLV за формулою (3.11). У табл. 4.3 наведено порівняння значень базового та ризик-скоригованого показників CLV.

Таблиця 4.3 – Порівняння отриманих результатів обчислень показників CLV

Місяць віку когорти	Базовий показник CLV	Ризик-скоригований показник CLV
2024-02	598,93	261,34
2024-03	181,41	239,73
2024-04	239,58	226,58
2024-05	145,73	174,73
2024-06	177,58	185,28
2024-07	149,93	167,76
2024-08	116,67	147,12
2024-09	120,94	141,93
2024-10	93,37	117,33
2024-11	101,63	109,03
2024-12	111,81	108,68

Графічне порівняння отриманих показників CLV наведено на рис. 4.2.

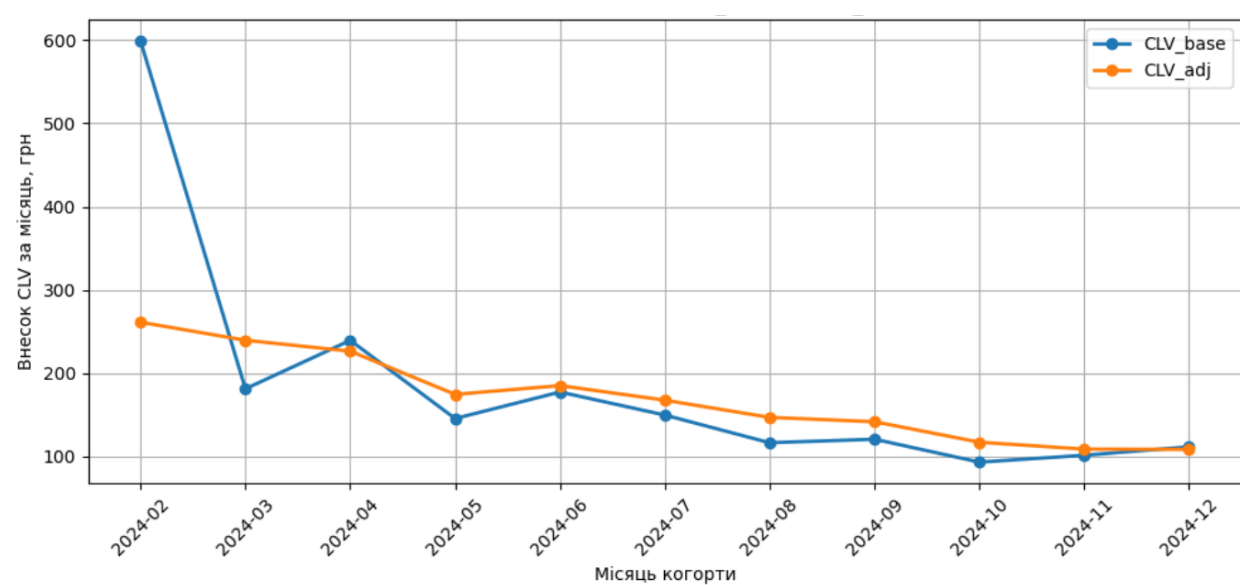


Рисунок 4.2 – Порівняння показників CLV

Сумарний ризик-скоригований показник CLV когорти лютого 2024 року:

$$CLV_{2024-02}^{adj} = \sum_{t=1}^T CLV_T = 1880,9 \text{ грн.} \quad (4.4)$$

4.2 Аналіз отриманих результатів

Використання survival-аналізу дозволило доповнити когортний підхід оцінкою ризику тривалого відтоку. Отримане значення функції виживання – 0,44 свідчить про те, що майже половина клієнтів має потенціал залишатися активними, навіть якщо вони не здійснили покупку в певному місяці.

Поведінкові фактори підтвердили очікувану логіку впливу на ризик відтоку:

- клієнти з високою частотою покупок демонструють нижчий ризик відтоку;
- вищий середній чек корелює зі стабільнішою активністю та лояльністю;
- тривалі інтервали між транзакціями є надійним індикатором ризику, що знижує прогнозовану цінність клієнта.

Результати обчислень продемонстрували, що скоригований за ризиком показник CLV є нижчим за базовий на близько 7–8%. Це очікуваний результат, оскільки базовий CLV не враховує індивідуальної ймовірності відтоку клієнтів, тоді як survival-корекція усуває цю неповноту та робить прогноз більш реалістичним.

В окремих періодах скоригований показник CLV перевищує базовий. Це може бути наслідком того, що аналіз виживання оцінює потенційну активність клієнта, тоді як когортний підхід базується лише на факті здійсненої покупки.

Survival-модель «пом'якшує» падіння активності у періоди, де $R(t)$ є низьким, і дає плавнішу динаміку.

Сумарне значення CLV зменшується, що цілком обґрунтовано з економічної точки зору: ризик-скоригований показник CLV відображає дійсні довгострокові перспективи клієнтів, враховуючи неоднорідність їхньої поведінки та ймовірність відтоку.

З метою комплексної оцінки запропонованого комбінованого методу було здійснено розрахунок сумарних показників CLV для всіх клієнтських когорт 2024 року.

Результати порівняння наведено в табл. 4.4.

Таблиця 4.4 – Результат порівняння сумарних показників CLV

Когорта клієнтів	Сумарний базовий показник CLV	Сумарний ризик-скоригований показник CLV
2024-01	2085,27 грн	2044,65 грн
2024-02	2037,58 грн	1880,9 грн
2024-03	1839,48 грн	1737,56 грн
2024-04	1729,99 грн	1561,57 грн
2024-05	1704,78 грн	1400,93 грн
2024-06	1736,88 грн	1841,59 грн
2024-07	1945,56 грн	1771,08 грн
2024-08	868,52 грн	1022,68 грн
2024-09	1374,21 грн	1183,99 грн
2024-10	539,99 грн	472,74 грн
2024-11	1171,69 грн	833,62 грн

Результати демонструють, що співвідношення між базовим та ризик-скоригованим показниками CLV залежить від поведінкових особливостей кожної когорти.

У когортах, де клієнти демонструють нерегулярні інтервали між

покупками, нестабільну активність або підвищений ризик відтоку, скоригований показник є нижчим за базовий. Це свідчить про те, що survival-модель коригує завищені очікування, притаманні класичному когортному підходу, та зменшує прогнозовану довгострокову цінність таких когорт відповідно до їх реальної поведінкової динаміки.

Водночас у деяких когорт ризик-скоригований показник CLV перевищує базовий. Це характерно для клієнтів зі стабільною активністю, короткими інтервалами між транзакціями або високим середнім чеком. У таких випадках survival-модель підтверджує підвищений потенціал клієнтів і демонструє, що класичний підхід може недооцінювати їхню цінність.

Таким чином, порівняння двох моделей дає змогу:

- оцінити поведінкові особливості когорт і ступінь стабільності їхньої поведінки;
- виявити найбільш перспективні сегменти клієнтів;
- коригувати управлінські рішення з урахуванням реальних поведінкових ризиків.

У подальших дослідженнях доцільно розширити набір поведінкових характеристик клієнтів та перевірити вплив інших survival-моделей на точність прогнозування довгострокової цінності. Перспективним напрямом є також інтеграція отриманих показників у рекомендаційні системи чи модулі персоналізації для формування індивідуальних пропозицій і цілеспрямованих стратегій утримання клієнтів. Це дозволить поглибити аналіз, підвищити адаптивність бізнес-процесів та забезпечити стійкість прогнозів у різних умовах і для вибірок різного розміру.

ВИСНОВКИ

У ході виконання магістерської кваліфікаційної роботи було здійснено комплексне дослідження методів аналізу поведінки клієнтів та визначено їх придатність до використання в ІС для планування діяльності підприємства. Встановлено, що сучасні підприємства, зокрема, у сфері електронної комерції, стикаються з низкою особливостей щодо поведінки клієнтів, які ускладнюють процес прогнозування та потребують поєднання різних аналітичних підходів.

У процесі дослідження було здійснено ґрунтовний огляд наявних підходів до аналізу клієнтської поведінки та визначено їх функціональні можливості. На основі порівняння класичних і прогнозних методів встановлено, що жоден із них окремо не забезпечує повного уявлення про довгострокову взаємодію клієнтів із підприємством. Тому було проведено аналіз когортного підходу, моделей оцінювання CLV та методів прогнозування відтоку з використанням аналізу виживання. У результаті обґрунтовано доцільність поєднання цих підходів у межах комбінованого методу.

Було розроблено комбінований метод аналізу поведінки клієнтів, заснований на застосуванні когортного аналізу, базового показника CLV та survival-моделі. Детально описано етапи роботи методу, а також побудовано його функціональну структуру у нотації IDEF0, що підтверджує можливість інтеграції методу в аналітичні модулі ІС.

Перевірено працездатність запропонованого комбінованого методу з використанням транзакційних даних умовної компанії. Було сформовано когорти клієнтів, визначено динаміку їхнього утримання, розраховано базові значення CLV та побудовано survival-модель з урахуванням поведінкових характеристик. Порівняння базового та скоригованого показників CLV продемонструвало, що їх співвідношення залежить від стабільності та особливостей поведінкової динаміки клієнтів у межах кожної когорти. Такий

підхід дав змогу отримати більш узгоджену оцінку довгострокової взаємодії клієнтів із підприємством і виявити групи з підвищеним або зниженим потенціалом. Таким чином, результати експериментальних досліджень підтвердили працездатність та ефективність розробленого комбінованого методу.

Отримані в роботі результати можуть бути використані під час розробки ІС для планування діяльності підприємства. Запропонований комбінований метод дає змогу формувати прогностичні показники, досліджувати особливості поведінки клієнтів, враховувати імовірність їхнього відтоку. Це створює підґрунтя для побудови гнучких моделей взаємодії з клієнтами, уточнення управлінських рішень і формування стратегій, які враховують реальні тенденції розвитку клієнтської бази.

Подальші дослідження можуть бути спрямовані на розширення набору поведінкових параметрів, удосконалення прогностичних моделей відтоку та інтеграцію отриманих результатів у системи персоналізації або рекомендаційні модулі. Це дасть змогу підвищити адаптивність аналітичних рішень та проводити більш глибоке моделювання поведінки клієнтів у різних галузях.

Результати дослідження пройшли апробацію на міжнародних наукових конференціях «Science and Information Technologies in the Modern World» (Афіни, 2025) [33] та «The Future of Science, Technology and Economy» (Софія, 2025) [34].

Роботу виконано відповідно до вимог щодо розробки та оформлення кваліфікаційних робіт [35], а також із дотриманням положень державних стандартів ДСТУ 3008:2015 [36] і ДСТУ 8302:2015 [37].

ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ

1. Urbany J., Dapena-Baron M. The pursuit of customer centricity // AMS review. 2025. URL: <https://doi.org/10.1007/s13162-024-00288-4> (date of access: 24.11.2025).
2. Момотков І. С. Нейромаркетинг в основі маркетингових рішень: нове бачення поведінки клієнтів // Український економічний часопис. 2025. No 9. С. 84–89. URL: <https://doi.org/10.32782/2786-8273/2025-9-14> (дата звернення: 24.11.2025).
3. Potapenko A., Shekhovtsova V. Protection of personal data of consumers in the analysis of demand for the company's products. Modern pedagogical technologies and innovative methods: materials of the 3d International scientific and practical conf., Seville, 25-28 Feb. 2025. Seville, 2025. P. 53–56.
4. Як розрахувати когортний аналіз // Data-life-ua. URL: <https://data-life-ua.com/analyst/yak-rozrakhuvaty-kohortnyy-analiz/> (дата звернення: 25.11.2025).
5. Somavarapu S. Cohort analysis for user behavior trends in marketplace platforms // International journal of research in modern engineering and emerging technology. 2025. No 1. P. 149–173.
6. Cohort analysis: 4 use cases for product managers // Smartlook blog. URL: <https://www.smartlook.com/blog/cohort-analysis/> (date of access: 25.11.2025).
7. Ahn J. A survey on churn analysis in various business domains // IEEE access. 2020. No 8. P. 220816–220839. URL: <https://doi.org/10.1109/access.2020.3042657> (date of access: 25.11.2025).
8. Земко П. А. Застосування логістичної регресії для прогнозування відтоку клієнтів телекомунікаційної компанії. Радіoeлектроніка та молодь у XXI столітті: матеріали 25-го Міжнар. молодіжн. форуму, м. Харків, 20-22 квіт. 2021 р. Харків, 2021. Т. 7. С. 53–54.
9. Geiler L., Affeldt S., Nadif M. A survey on machine learning methods for

churn prediction // International journal of data science and analytics. 2022. No 14. P. 217–242. URL: <https://doi.org/10.1007/s41060-022-00312-5> (date of access: 25.11.2025).

10. Ashtari H. XGBoost vs random forest vs gradient boosting // Spiceworks inc. URL: <https://www.spiceworks.com/tech/artificial-intelligence/articles/xgboost-vs-random-forest-vs-gradient-boosting/> (date of access: 26.11.2025).

11. Бовчалюк С., Гайдай Я. Аналіз методу опорних векторів у порівнянні з традиційними методами передбачення ринкових рухів // Збірник наукових праць. 2024. No 77. С. 89–92. URL: <https://doi.org/10.26906/sunz.2024.3.089> (дата звернення: 26.11.2025).

12. Edwine N., Wang W. Detecting the risk of customer churn in telecom sector: a comparative study // Mathematical problems in engineering. 2022. No 1. P. 1–16. URL: <https://doi.org/10.1155/2022/8534739> (date of access: 26.11.2025).

13. Dusmatov B. The subscription economy: transforming business models // Journal of analysis and inventions. 2024. No 2. P. 71–76.

14. Uwah V. Impact of subscription revenue model on 21st century businesses // Journal of science education and humanities. 2023. No 7. P. 132–142.

15. Kowalkowski C., Ulaga W. Subscription offers in business-to-business markets: conceptualization, taxonomy and framework for growth // Industrial marketing management. 2024. No 117. P. 440–456. URL: <https://doi.org/10.1016/j.indmarman.2024.01.014> (date of access: 26.11.2025).

16. Тимошенко К. Роль стратегічного управління в адаптації бізнес-моделей електронної комерції до динамічних змін ринку // Економіка та суспільство. 2025. No 74. С. 637–642. URL: <https://doi.org/10.32782/2524-0072/2025-74-94> (дата звернення: 26.11.2025).

17. Dogan O. Business analytics in customer lifetime value: an overview analysis // WIREs data mining and knowledge discovery. 2024. No 15. P. 1–18. URL: <https://doi.org/10.1002/widm.1571> (date of access: 27.11.2025).

18. Qismat T., Feng Y. Comparison of classical RFM models and machine learning models in CLV prediction // Master's thesis. Oslo. 2023.

19. Bauer J., Jannach D. Improved customer lifetime value prediction with sequence-to-sequence learning and feature-based models // *ACM transactions on knowledge discovery from data*. 2021. No 5. P. 1–37. URL: <https://doi.org/10.1145/3441444> (date of access: 27.11.2025).

20. Akter J., Roy A. Artificial intelligence-driven customer lifetime value (CLV) forecasting: integrating RFM analysis with machine learning for strategic customer retention // *Journal of computer science and technology studies*. 2025. No 1. P. 249–257. URL: <https://doi.org/10.32996/jcsts.2025.7.1.18> (date of access: 28.11.2025).

21. Герчаківський О. Аналіз впливу маркетингових комунікацій на споживачів в ритейлі: переваги та недоліки інтегрованої маркетингової стратегії // *Академічні візії*. 2023. No 17. С. 1–9. URL: <https://doi.org/10.5281/zenodo.7759321> (дата звернення: 28.11.2025).

22. Darvidou K. Omnichannel marketing in the digital age: creating consistent, personalized and connected customer experiences // *Technium business and management*. 2023. No 10. P. 34–54 (date of access: 29.11.2025).

23. Sinha R., Arbour D. Bayesian modeling of marketing attribution // *JSM proceedings*. 2022. URL: <https://doi.org/10.48550/arXiv.2205.15965> (date of access: 29.11.2025).

24. Abrardi L., Cambini C., Rondi L. Artificial intelligence, firms and consumer behavior: a survey // *Journal of economic surveys*. 2021. No 4. P. 969–991. URL: <https://doi.org/10.1111/joes.12455> (date of access: 01.12.2025).

25. Siti Z., Hazik M. Customer predictive analytics using artificial intelligence // *The Singapore economic review*. 2025. No 4. P. 1009–1020. URL: <https://doi.org/10.1142/s0217590820480021> (date of access: 01.12.2025).

26. Tudoran A., Hjerrild C., Thomasen S. Understanding consumer behavior during and after a pandemic: implications for customer lifetime value prediction models // *Journal of business research*. 2024. No 174. P. 114527–114540. URL: <https://doi.org/10.1016/j.jbusres.2024.114527> (date of access: 02.12.2025).

27. Gurunathan A., Lakshmi K. Exploring the perceptions of generations x, y

and z about online platforms and digital marketing activities: a focus-group discussion based study // International journal of professional business review. 2023. No 5. P. 1–34. URL: <https://doi.org/10.26668/businessreview/2023.v8i5.2122> (date of access: 03.12.2025).

28. Adekunle B., Balogun E., Ogunsola K. Improving customer retention through machine learning: a predictive approach to churn prevention and engagement strategies // International journal of scientific research in computer science, engineering and information technology. 2023. No 4. P. 507–523. URL: <https://ijsrcseit.com/CSEIT2311254> (date of access: 03.12.2025).

29. Kalishina D. Algorithmic customer churn prediction and targeted intervention: optimizing customer lifetime value in data-sparse SME environments // World journal of advanced research and reviews. 2025. No 1. P. 593–603. URL: <https://doi.org/10.30574/wjarr.2025.26.1.1045> (date of access: 03.12.2025).

30. IDEF // Maxym Zosym. URL: <https://www.maxzosim.com/idef/> (date of access: 04.12.2025).

31. Cox proportional hazards model // GeeksforGeeks. URL: <https://www.geeksforgeeks.org/data-science/cox-proportional-hazards-model/> (date of access: 05.12.2025).

32. Understanding customer attrition through survival analysis: Kaplan Meier & Cox proportional // Medium. URL: <https://medium.com/@agudaaidris5/understanding-customer-attrition-through-survival-analysis-kaplan-meier-cox-regression-on-in-45e3d871a82d> (date of access: 06.12.2025).

33. Потапенко А. О., Петров К. Е. Дослідження та використання методів аналізу поведінки клієнтів для планування діяльності підприємства. Наука та інформаційні технології у сучасному світі: матеріали Міжнар. наук.-практ. конф., м. Афіни, 21-23 трав. 2025 р. Афіни, 2025. С. 295–297.

34. Потапенко А. О., Петров К. Е. Комбінований метод аналізу поведінки клієнтів у системах планування діяльності підприємства. Майбутнє науки, технологій та економіки: матеріали Міжнар. наук.-практ. конф., м. Софія, 29-31 жовт. 2025 р. Софія, 2025. С. 206–210.

35. Методичні вказівки щодо розробки та оформлення кваліфікаційної роботи другого (магістерського) рівня вищої освіти за освітньо-професійною програмою «Інформаційні управляючі системи та технології» спеціальності 122 Комп'ютерні науки / Упоряд.: К.Е. Петров, В.М. Левикін, С.Ф. Чалий, М.В. Євланов, В.І. Саєнко, Д.К. Міхнов, А.В. Міхнова, О.В. Чала. ХНУРЕ: Харків, 2024. 24 с.

36. ДСТУ 3008:2015. Інформація та документація. Звіти у сфері науки і техніки. Структура та правила оформлювання. Чинний від 2017-07-01. – Київ: ДП «УкрНДНЦ», 2016. – 31 с.

37. ДСТУ 8302:2015. Інформація та документація. Бібліографічне посилання. Загальні положення та правила складання. Чинний від 2016-07-01. – Вид. офіц. Київ : УкрНДНЦ, 2016. 16 с.