

МІЖДИСЦИПЛІНАРНІ МЕТРИКИ В УПРАВЛІННІ ЛОЯЛЬНІСТЮ В КНИЖКОВОМУ РІТЕЙЛІ

І.К. Бабіч¹, Д.Л. Орловський²

¹ аспірант кафедри ПІТУ, НТУ «ХПІ», Харків, Україна

² доцент кафедри ПІТУ, канд. техн. наук, НТУ «ХПІ», Харків, Україна

Ihor.Babich@cs.khpi.edu.ua

Підвищення ефективності програм лояльності у сучасних умовах потребує значної уваги. Розробка методів та технологій, спрямованих на формалізацію та автоматизацію процесів оцінки, прогнозування і підтримці клієнтської лояльності, може спиратися на використанні міждисциплінарних показників. Класичні метрики, як RFM або CLV, можуть бути розширені такими показниками: коефіцієнт концентрації, ентропійні та дивергенційні міри та показники виживаності.

Індекс (коефіцієнт) Джині (Gini coefficient) та індекс Герфіндаля–Гіршмана (Herfindahl-Hirschman Index, ННІ) походять з економіки і вимірюють нерівномірність розподілу ресурсів. *Коефіцієнт Джині* оцінює нерівність у розподілі значень (наприклад, доходів) від 0 (повна рівність) до 1 (максимальна нерівність). *ННІ* визначає концентрацію ринку, обчислюючись як сума квадратів часток учасників ринку; значення близьке до 1 (або 10000, якщо виражати в умовних пунктах) вказує на майже монополію, тоді як низькі значення сигналізують про високу конкуренцію. Такі метрики можна рахувати і для всього бізнесу, аналізуючи концентрацію виручки по клієнтах (рис. 1)

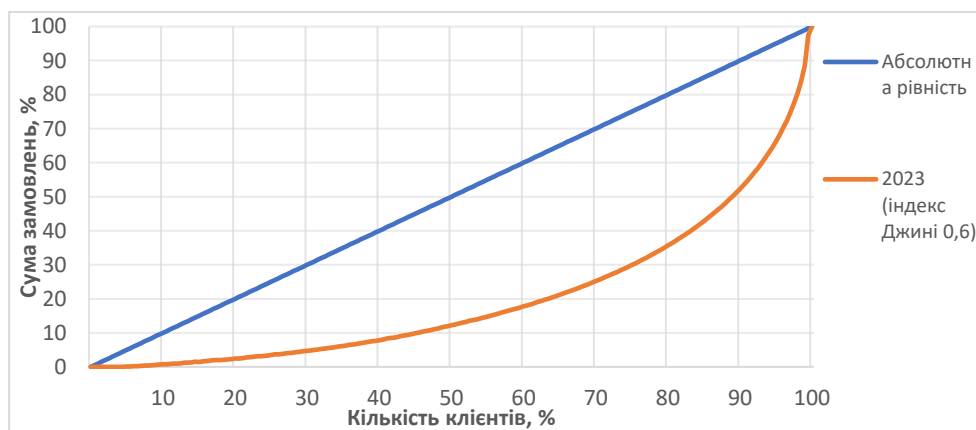


Рис. 1 – Крива Лоренца за сумою замовлень (2023)

Висока концентрація виручки 0,6 – позитивна ознака ризикова з точки зору стійкості бізнесу. Втрата декількох топ-клієнтів може суттєво вплинути на прибуток. Необхідно диверсифікувати клієнтську базу: активізувати середній сегмент через персональні пропозиції, тематичні підбірки. Показники концентрації дають змогу кількісно підтвердити правило Парето і виявити рівень залежності бізнесу від постійних клієнтів.

Поняття ентропії прийшло з теорії інформації: *ентропія Шеннона* вимірює середній рівень невизначеності в розподілі ймовірностей. Це міра різноманітності результатів – якщо всі випадки рівноймовірні, ентропія максимальна, а якщо переважає один результат, ентропія низька. *Індекс Сімпсона* – міра, запозичена з екології, що оцінює біорізноманіття середовища. Він відображає ймовірність того, що дві випадково вибрані

особини належать до одного виду. Таким чином, обидві метрики кількісно характеризують, наскільки рівномірно і широко розподіляються елементи серед різних категорій. У сфері роздрібних продажів книг можна розглядати покупки клієнта як набір категорій (жанрів або товарних розділів) зі своїми частками у витратах. Висока ентропія Шеннона означатиме, що клієнт купує дуже різножанрові книжки приблизно в рівних пропорціях. Натомість низька ентропія (близька до нуля) вказує, що одна категорія домінує (скажімо, 90% витрат лише на підручники). Для бізнесу, що має широкий асортимент, це має значення: чим ширший спектр потреб клієнта він закриває, тим сильніше клієнт прив'язаний до компанії. [1] Метрики різноманіття дозволяють сегментувати клієнтів за широтою їхніх інтересів: «вузькопрофільні» клієнти та «різнопланові».

Дивергенція Кульбака–Лейблера (KL) – це концепція з математичної статистики та інформаційної теорії, що вимірює відмінність між двома розподілами ймовірностей. KL-дивергенція несиметрична, тому для симетричного порівняння розподілів використовують похідні величини, зокрема *дивергенцію Джессена–Шеннона (JS)*. Дивергенція JS завжди визначена і обмежена від 0 до 1, і дорівнює нулю лише при повному збігу розподілів. Метрики KL та JS дозволяють сегментувати клієнтів за схожістю профілів покупок. Застосовується кластеризація клієнтів на основі JS-відстані між їхніми «жанровими профілями». Такі сегменти будуть дуже інтерпретованими: наприклад, один кластер може відповідати «любителям художньої літератури», інший – «поціновувачам наукпопу й бізнес-книг», третій – «читачам дитячої літератури» тощо. Для маркетолога метрики дивергенції дозволяють також виявити аномалії – клієнтів, які не вписуються ні в один великий сегмент. Знаючи динаміку KL-дивергенції вподобань клієнта, можна передбачати, чи не настає переломний момент. Якщо профіль покупок клієнта почав сильно відхилятися від його ж історичного профілю (він раптом різко перемкнувся на інший жанр) – можливо, попередні рекомендації для нього вже не актуальні і потрібен перегляд стратегії утримання.

Survival-аналіз виник у біостатистиці для аналізу часу до настання події (типово – смерті або відмови механізму). *Оцінка Каплана–Мейєра* – це непараметричний метод, що дозволяє розрахувати емпіричну функцію виживання – ймовірність «дожити» до певного часу, не зазнавши події. В задачах лояльності survival-аналіз може застосовуватися для моделювання часу до відтоку клієнта або його життєвого циклу. Якщо трактувати момент «смерті» як момент, коли клієнт не здійснює більше покупок, можна побудувати криву Каплана–Мейєра для клієнтської бази. Така крива наочно показує динаміку утримання: наприклад, може виявитися, що найбільший відсів клієнтів стається в перші 3 місяці (крива різко падає), а далі плато – ті, хто залишився після 1 року, купують і надалі майже постійно. Аналіз Каплана–Мейєра дозволяє порівнювати групи клієнтів: наприклад, побудувати окремі криві виживання для клієнтів, залучених через різні канали. Можна виявити, що клієнти, які прийшли по рекламі в соцмережах, «живуть» значно менше (швидше кидають купувати), ніж ті, хто прийшов за рекомендацією знайомих – різниця кривих буде статистично значущою за лог-ранк тестом.

Розглянуті типи метрик демонструють, як міждисциплінарні підходи можуть збагатити аналіз даних у програмах лояльності книжкового бізнесу. Це може допомогти глибше зрозуміти поведінку читачів, тонше сегментувати аудиторію та проактивно управляти лояльністю на основі даних.

Список літератури:

1. Ruben Nicolas-Sans & Daniel González Ibáñez (2021) Customer basket heterogeneity: how to measure it and some possible business applications, *Economic Research-Ekonomska Istraživanja*, 34:1, 2572-2592, DOI: 10.1080/1331677X.2020.1835518