

АНОТАЦІЯ

Іванов Д.В. Інтелектуальний аналіз даних B2B-клієнтів для персоналізації контенту систем електронної комерції. – Кваліфікаційна наукова праця на правах рукопису.

Дисертація на здобуття наукового ступеня доктора філософії за спеціальністю 122 – Комп'ютерні науки. – Національний університет «Одеська політехніка», МОН України, Одеса, 2026.

У **вступі** обґрунтовано актуальність створення персоналізованого контенту систем електронної комерції для підвищення своєчасності формування предикативних пропозицій у відповідності до класів B2B-клієнтів за рахунок інтелектуального аналізу масивів транзакційно-повелінкових даних. Визначено **об'єкт дослідження** як процес персоналізації контенту в системах B2B електронної комерції з урахуванням транзакційно-поведінкових даних клієнтів. **Предметом дослідження** є моделі та методи інтелектуального аналізу транзакційно-поведінкових даних B2B-клієнтів систем електронної комерції. Сформульовано **мету дослідження**, яка полягає в вирішенні науково-прикладної задачі персоналізації контенту систем електронної комерції за рахунок підвищення ефективності формування предикативних пропозицій у відповідності до класів B2B-клієнтів шляхом розробки інтелектуального ціннісно-зваженого темпорального підходу до аналізу масивів транзакційно-поведінкових даних. Досягнення зазначеної мети дослідження пов'язано з вирішенням наступних задач: проаналізувати існуючі технології персоналізації контенту в системах електронної комерції та визначити шляхи підвищення їх ефективності для B2B-клієнтів в сучасних умовах; розробити ціннісно-зважений темпоральний підхід до інтелектуального аналізу масивів транзакційно-повелінкових даних B2B клієнтів; розробити метод класифікації B2B клієнтів для предикативної персоналізації контенту систем електронної комерції; провести моделювання стратегій персоналізації контенту за рахунок підвищення ефективності формування предикативних пропозицій порівняно з класичними підходами та інтегрувати

розроблений аналітичний модуль в систему рекомендації для реальної B2B-платформи.

У **першому розділі** проаналізовано стан існуючих технологій персоналізації контенту в системах електронної комерції та визначено шляхи підвищення їх ефективності для B2B-клієнтів в сучасних умовах. Обґрунтовано специфіку B2B-сегмента в умовах динамічного ринку. Показано, що понад 60 % українських B2B-компаній зазначають, що основною перешкодою для переходу в онлайн є неможливість забезпечити клієнта релевантним контентом, а саме індивідуальними цінами, технічною документацією в реальному часі та предикативними автоматизованими рекомендаціями щодо здійснення нових транзакцій для поповнення складів. Дослідження показують, що лише 12 % вітчизняних B2B-платформ використовують елементи інтелектуального аналізу даних для персоналізації пропозицій, тоді як у світі цей показник перевищує 50 %.

Результати проведеного аналізу B2B-сегмента електронної комерції за схемою «*Wholesale, Marketing, Customer, Collaborative*» свідчать, що на відміну від B2C, де домінують емоційні та разові покупки, B2B-транзакції характеризуються циклічністю, раціональністю та тривалим життєвим циклом клієнта. Це доводить, що персоналізація в B2B-системах має базуватися на глибокому інтелектуальному аналізі ціннісних та темпоральних параметрів та докорінно відрізнятися від B2C-систем, де контент генерується переважно шляхом статистичного аналізу частотних покупок та загальних поведінкових шаблонів масового споживача.

Проаналізовано традиційні методи взаємодії з клієнтами в системах електронній комерції на основі створення рекомендацій та зауважено, що функціонал модулів колаборативної та контентної фільтрації при врахуванні операційної складності B2B-процесів має певні недоліки. Так виявлено, що класичні рекомендаційні системи часто ігнорують часову динаміку та специфічну цінність товарів, що призводить до низької релевантності пропозицій у сферах із рідкісними, але високовартісними закупівлями. Крім того показано, що незважаючи на спроби створення складних гібридних рекомендаційних систем, як надбудови над B2B-системою електронної комерції шляхом комбінування

колаборативного та контентного підходів залишаються недоліки ігнорування економічної цінності та стабільності циклу транзакцій та не гарантується своєчасність рекомендацій для бізнес-клієнта. Таким чином, аналіз технологій персоналізації контенту за допомогою базових технологій рекомендаційних систем виявив системний дефіцит інструментів для врахування ціннісно-темпоральної специфіки B2B-транзакцій. Це обґрунтовує необхідність використання методів інтелектуального аналізу транзакційно-поведінкових даних B2B-клієнтів для побудови асоціацій та послідовностей.

Однак, проведений компаративний аналіз методів ІАД, спрямованих на побудову асоціацій та послідовностей у транзакційних масивах дозволив встановити наступне:

- класичні методи аналізу асоціативних правил (*Association Rule Mining ARM*), зокрема Apriori, FP-Growth та Eclat, орієнтовані на знаходження статичних зв'язків товарів в транзакції. При цьому враховується лише частота появи предметів у транзакціях, що повністю нівелює інформацію про його вартість чи обсяг закупівлі. Як наслідок, стратегічно важливі, але рідкісні B2B-товари відсікаються, тоді як дешеві супутні товари з високою частотою стають основою моделі. Це все демонструють повну «фінансову сліпоту», класичних методів *ARM*;

- класичні методи аналізу послідовних шаблонів (*Sequential Pattern Mining, SPM*), такі як GSP, PrefixSpan та SPADE, спрямовані на виявлення кореляцій між транзакціями на тлі хронології подій. При цьому вони частково вирішують проблему врахування фактору часу, проте характеризуються високим рівнем «ідентифікаційної дискретності», розглядаючи клієнта як абстрактну послідовність різних ідентифікаторів, що не дає можливість виявити індивідуальний ритм транзакцій та призводить до «темпоральної сліпоти» *SPM*-методів.

Таким чином, виявлено, що існуючий інструментарій ІАД, а саме *ARM* та *SPM*, потребує адаптації шляхом інтеграції фінансових показників корисності та часових інтервалів між транзакціями.

Наприкінці розділу на основі аналізу можливостей розповсюдженого галузевого стандарту, а саме є моделі RFM (*Recency, Frequency, Monetary*) щодо

сегментації клієнтів в B2B-середовищі також доведено її суттєві обмеження. Так показник Recency (давність) фіксує лише статичну точку на осі часу, не враховуючи волатильність індивідуального закупівельного циклу, а показник Monetary (вартість) оцінює валовий дохід, ігноруючи специфічну корисність (Utility) та стабільність патернів споживання. Таким чином класична модель RFM-аналізу орієнтована на ретроспективний опис стану («що відбулося»), що не дозволяє будувати предикативні моделі CLV (Customer Lifetime Value) з високою точністю для B2B-клієнтів із нерівномірним ритмом активності.

Узагальнюючи результати аналізу стану та тенденцій галузі, встановлено критичну необхідність розробки нових підходів до персоналізації B2B-контенту. Обґрунтовано доцільність переходу від статичної сегментації до побудови ціннісно-зваженого темпорального профілю клієнта. Це дозволить подолати фінансову та темпоральну сліпоту існуючих моделей та забезпечити перехід від констатації минулих транзакцій до предикативного управління комунікаційними тригерами на основі прогнозованого закупівельного циклу, фінансової ємності та стабільності поведінки кожного конкретного контрагента.

У **другому розділі** розроблено ціннісно-зважений темпоральний підхід до інтелектуального аналізу масивів транзакційно-повелінкових даних B2B клієнтів. Результатом проведення системного дослідження існуючих методів інтелектуального аналізу асоціативних правил та послідовних шаблонів в накопичених масивах B2B транзакцій є встановлення *фундаментальних методичних обмежень*, які унеможливають їх безпосереднє застосування для створення персоналізованого контенту у системах електронної комерції. Обґрунтовано існування п'яти пов'язаних між собою критичних розривів (обмежень) між математичними припущеннями методів пошуку асоціацій та послідовностей та природою B2B-предметної галузі, а саме це *«фінансова сліпота»* бінарних моделей, наявність *«комбінаторного вибуху»* при зваженому аналізі, *«темпоральна сліпота»* статичних моделей, наявність *ідентифікаційної дискретності* та *відсутність диференційованої стратегії взаємодії з клієнтами*. Встановлено, що зв'язок цих обмежень має ієрархічну структуру. Тобто спроба

подолати фінансову сліпоту класичних моделей ІАД шляхом додавання вартісних показників активує проблему комбінаторного вибуху. Своєю чергою, перехід до аналізу темпоральних інтервалів транзакційних даних вимагає не лише нової архітектури обробки, а й усунення ідентифікаційної дискретності для B2B клієнтів, оскільки розрахунок циклічності можливий лише на цілих, а не фрагментованих часових рядах. Таким чином, фінальне обмеження, відсутність диференційованої стратегії взаємодії з клієнтами, тобто стратегічна монотонність є результатом дефіциту даних на всіх попередніх етапах аналізу. Тобто в B2B-системі електронної комерції не можливо диференціювати (визначати класи) клієнтів без ідентифікації їхніх індивідуальних фінансово-темпоральних патернів.

Таким чином, подолання зазначених обмежень є науковою основою для розробки ціннісно-зваженого темпорального аналізу, математичний апарат якого умовно поділений на дві складові частини: модель ціннісно-зваженої транзакції та метод розростання ціннісних патернів (*Utility Pattern Growth, UP-Growth*), а також модель темпоральної послідовності та метод темпорального аналізу ціннісних патернів (*Utility-Temporal Pattern Growth, UT- Growth*).

Розроблено модель ціннісно-зваженої B2B-транзакції шляхом додавання до класичного визначення транзакції формалізацію індикатора сумарної корисності (*Transaction Utility*) – $U(T)$ визначеної транзакції, як суми добутків кількості придбаних одиниць товарних складових на їх індивідуальну ціну. Для подолання ефекту комбінаторного вибуху, який виникає за використання індикатора $U(T)$ формалізовано визначення транзакційно-зваженої корисності (*Transaction-Weighted Utilization*) – $TWU(X)$, як сумарну вартість усіх транзакцій бази, які містять визначену транзакцію. З використанням моделі ціннісно-зваженої B2B-транзакції розроблено метод розростання ціннісних патернів *UP-Growth*. Показано, що його використання, дозволяє усунути «фінансову сліпоту» класичного підходу до пошуку асоціативних патернів в транзакційних даних. Так на *першому проході* при побудові *UP*-дерева в вузлах накопичується не кількість входжень товарної позиції, а їх грошова вартість. Використання механізму *TWU* *від час другого проході* відновлює властивість антимонотонності і забезпечує

обчислювальну стабільність навіть при аналізі довгих B2B-накладних. Результатом роботи методу є впорядкований перелік ціннісних патернів – комбінацій товарів, які формують основну частку виторгу підприємства.

Розроблено модель темпоральної B2B-послідовності шляхом додавання до класичного визначення послідовності часу виникнення транзакцій, що дозволило отримати вектор міжзакупівельних інтервалів (*Inter-Purchase Time, IPT*) для довільного B2B-клієнта. Це спільно з розрахованим безрозмірним показником – коефіцієнтом варіації (*Coefficient of Variation, CV*) дає можливість визначити стабільність циклу транзакцій, долаючи таким чином «темпоральну сліпоту» та ідентифікаційну дискретність класичного методу пошуку шаблонів в послідовностях клієнтів. З використанням моделі темпоральної B2B-послідовності розроблено метод темпорального аналізу ціннісних патернів *UT- Growth*, які попередньо отримано за методом *UP-Growth*. Показано, що використання методу *UT- Growth* дозволяє усунути «темпоральну сліпоту» класичного підходу до пошуку послідовних патернів в транзакційних даних через побудову *IPT*-векторів та їх подальшого аналізу за критерієм *CV* для довільного B2B-клієнта. Наведено інтерпретацію критерію *CV* в залежності від встановленого порогу стабільності. Значення порогу не є фіксованим і може налаштовуватися залежно від специфіки галузі та прийнятності підприємства до похибки прогнозу.

Як загальний підсумок розділу показано що використання запропанованого двоетапного конвеєру методів *UP-Growth* та *UT- Growth* дозволяє отримати перелік персоналізованих тригерів, завдяки якому для кожного контрагента та кожної фінансово значущої та циклічно стабільної транзакції визначається момент отримання предикативної пропозиції. Як підтвердження наведено детальний розгляд двох типових сценаріїв для B2B систем електронної комерції – «регулярна сезонна закупівля» та «нерегулярна закупівля». Саме можливість їх реалізації перетворює B2B-платформу з реактивного каталогу на інтелектуальну систему рекомендацій, здатну передбачати потреби клієнтів, щодо поповнення складських запасів і взаємодіяти з ними у найбільш сприятливий момент.

У третьому розділі розроблено метод класифікації B2B клієнтів для предикативної персоналізації контенту систем електронної комерції. Доведено існування трьох критичних обмежень використання класичного *RFM*-аналізу для класифікації B2B-клієнтів. Показано, використання параметру давності *R* дозволяє фіксувати лише час, що минув з останньої транзакції. При цьому не будується вектор межакупівельних патернів *IPT*, що робить показник *R* «темпорально сліпим» до індивідуального ритму B2B-бізнесу. Використання параметру частоти *F* дозволяє оцінювати загальну кількість транзакцій за період. При цьому ігнорується стабільність транзакцій, індикатором якої є коефіцієнт варіації *CV*, розрахований на основі *IPT*. Параметр грошового обсягу транзакції *M* розраховується як загальна сума витрат клієнта. Але в такому вигляді він не дає інформацію про структуру транзакцій. При цьому не враховується транзакційно-зважена корисність (*TWU*). Це не дозволяє виділити патерни товарів-драйверів, які забезпечують стабільність обороту.

Для подолання зазначених недоліків стандартного *RFM*-аналізу формалізовано ціннісно-зважений темпоральний профіль клієнта (*Utility-Weighted Temporal Customer Profile, UW-TCP*) у вигляді кортежу значень наступних ознак: середній цикл закупівлі (значення часу між повторними транзакціями), середня стабільність циклу транзакцій (коефіцієнт варіації) та середня грошова вартість транзакції

З метою формування предикативних пропозицій для B2B клієнтів розроблено метод класифікації, який є третьою складовою ціннісно-зваженого темпорального аналізу транзакційно-поведінкових даних B2B клієнтів системи електронної комерції. Він функціонує виключно як надбудова над запропонованими методами *UP-Growth* та *UT-Growth*. Використання методу класифікації B2B клієнтів базується ознаках ціннісно-зваженого темпорального профілю B2B клієнта, які є агрегатами вже наявних проміжних результатів, що робить метод класифікації B2B-клієнтів *третьім етапом* конвеєра формування предикативних пропозицій для B2B-клієнтів. Визначення номерів класів B2B-клієнтів інтегрується у систему формування комунікаційних тригерів. Це

забезпечує змістовний зв'язок між класом клієнта та стратегією взаємодії з ним через математично обґрунтоване призначення ширини вікна триггеру відповідно до цінності клієнта а не через суб'єктивне визначення порогів сегментації як в стандартному *RFM*-аналізі.

У четвертому розділі виконано експериментальну перевірку ефективності запропонованого ціннісно-зваженого темпорального підходу для формування предикативних пропозицій у B2B-платформі електронної комерції. Розроблені моделі та методи розглянуто як складові рекомендаційного модуля, інтегрованого в контур роботи реальної B2B-платформи «База взуття». Формалізовано взаємодію платформи, аналітичного конвеєра та інтерфейсу менеджера, у межах якої історія замовлень, дані про клієнтів, товарні позиції, артикули, ціни та суми транзакцій використовуються для формування активних персоналізованих рекомендацій.

Експериментальне дослідження проведено на масиві транзакційно-поведінкових даних за 2022–2025 роки, який охоплював 29,35 тис. замовлень, 2,53 тис. клієнтів, 56,20 тис. унікальних артикулів і загальний обсяг замовлень 416,28 млн грн. Після попередньої обробки, формування транзакційного подання та фільтрації надто довгих замовлень для подальшого аналізу використано 28,53 тис. замовлень, що дозволило зменшити комбінаторну складність без втрати основної частини експериментального масиву.

Для оцінювання ефективності проведено порівняльне експериментальне дослідження, у межах якого послідовно зіставлено чотири сценарії: базовий частотний S0, ціннісно-зважений S1, ціннісно-темпоральний S2 та інтегрований S3, що додатково враховує клас B2B-клієнта. Результати показали послідовне зростання якості рекомендацій: конверсія збільшилася з 3,91 % у S0 до 5,39 % у S1, 12,69 % у S2 та 15,91 % у S3. Одночасно кількість сформованих рекомендацій скоротилася з 920 до 220, тобто на 76,09 %, що свідчить про зменшення надмірного комунікаційного навантаження на менеджерів і клієнтів.

Найкращий результат отримано для інтегрованого сценарію S3. Він сформував 220 рекомендацій, з яких 35 виявилися успішними, забезпечив середню виручку на одну рекомендацію 6,91 тис. грн та охоплену тестову виручку 1,52 млн

грн. Порівняно з базовим сценарієм S0, кількість рекомендацій зменшилася на 700, однак кількість успішних звернень майже не змінилася. При цьому значення метрики *ARPR* (ефективність однієї рекомендації) збільшилось із 1,21 тис. грн до 6,91 тис. грн, тобто приблизно у 5,7 раза. Це підтвердило, що повна модель відсікає значну частину нерелевантних рекомендацій та підвищує грошову цінність кожної.

Економічна оцінка також засвідчила перевагу інтегрованого підходу. За умови вартості одного комунікаційного звернення 150 грн операційні витрати зменшилися зі 138,00 тис. грн у базовому сценарії до 33,00 тис. грн у сценарії S3. Вартість однієї успішної рекомендації знизилася з 3,83 тис. грн до 0,94 тис. грн, а чистий фінансовий результат інтегрованої моделі становив 1,49 млн грн. Отримані результати доводять, що запропонований підхід забезпечує перехід від масового формування рекомендацій до точнішої, більш своєчасної та економічно обґрунтованої персоналізації контенту B2B-системи електронної комерції.

Архітектурно розроблений прикладний рекомендаційний модуль розширює функціональність B2B-платформи без зміни її основних транзакційних процесів. Програмну реалізацію виконано у вигляді аналітичного модуля, який отримує дані про клієнтів, замовлення та товарні позиції, застосовує розроблений триетапний конвеєр і формує активні предикативні рекомендації для менеджера. Для демонстрації практичного використання результатів реалізовано інтерфейсні екрани головної панелі менеджера, активних рекомендацій та персональної пропозиції B2B-клієнту.

Наукова новизна одержаних результатів полягає в наступному:

- удосконалено модель ціннісно-зваженої транзакції, яка на відміну представлення у вигляді переліка товарних пропозицій (SKU), додатково включає обсяг SKU та індивідуальну ціну, що дозволяє в подальшому використовувати їх добуток для розрахунку транзакціо-зваженої корисності (TWU) та ідентифікувати транзакції не лише за критерієм повторюваності, а і за економічної значущості для подолання обмеження «фінансової сліпоты»;

- отримав подальший розвиток метод розростання цінністних патернів *UP-Growth* за рахунок використання показника транзакціо-зваженої корисності

(TWU) для розрахунку значень вузлів дерева *UP-Growth* замість частотних лічильників класичного *FP-Growth*, що дозволило відновити властивість антимонотоності долаючи обмеження «комбінаторного вибуху», а також змінити пріоритетність генерації асоціативних правил, фокусуючи систему персоналізації на економічно значущих транзакціях;

– удосконалено модель темпоральної послідовності, яка на відміну представлення у вигляді кортежів ідентифікаційно дискретних транзакцій враховує наявність детермінованого ідентифікатора B2B-клієнта та додатково включає час виникнення транзакції, що дозволяє створити вектор міжтранзакційних інтервалів (*IPT*) та використовувати його статистичні характеристики для розрахунку ймовірної дати активації предикативної пропозиції для подолання обмежень «темпоральної сліпоты» та ідентифікаційної дискретності;

– отримав подальший розвиток метод темпорального аналізу ціннісно патернів *UT-Growth*, який спільно з моделлю темпоральної послідовності дозволяє розрахувати коефіцієнт варіації *CV* вектору міжтранзакційних інтервалів, що дає можливість визначити дату предикативної пропозиції B2B-клієнту формуючі предикативні тригери вигляду «клієнт – товарний набір – дата звернення» для підвищення рівня персоналізації контенту B2B-систем електронної комерції.

– вперше розроблено метод класифікації B2B клієнтів систем електронної комерції, який на відміну від існуючого RFM-аналізу базується на ціннісно-зваженому темпоральному профілі клієнта сформованому у результаті реалізації двоетапного конвеєру аналізу транзакційно-поведінкових даних B2B клієнта. Це дозволило реалізувати трьох етапний конвеєр формування індивідуальних предикативних пропозицій в залежності від визначеного класу клієнта та підвищити фінальну ефективність предикативної персоналізації контенту

Практичне значення результатів дослідження полягає у розробленні програмного аналітичного модуля для формування персоналізованих предикативних пропозицій у системах B2B електронної комерції. Основою модуля є триетапний конвеєр, який поєднує ціннісно-зважений аналіз товарних патернів, темпоральну перевірку стабільності закупівельних циклів та класифікацію B2B-

клієнтів за ціннісно-зваженим темпоральним профілем. Це дозволяє формувати пропозиції не лише на основі товарної релевантності, а й з урахуванням фінансової значущості патерну, прогнозованого моменту звернення та класу клієнта.

Практичну апробацію результатів виконано шляхом інтеграції розробленого конвеєра у вигляді Python-модуля в контур роботи B2B-платформи «База взуття». Реалізований модуль використовує дані про клієнтів, замовлення, товарні позиції, артикули, ціни та суми транзакцій, після чого формує активні рекомендації для менеджера. Для демонстрації прикладного використання результатів розроблено інтерфейсні екрани головної панелі менеджера, активних рекомендацій та персональної пропозиції B2B-клієнту.

Ефективність запропонованого підходу підтверджено експериментальним порівнянням базового частотного сценарію з повним інтегрованим триетапним конвеєром з використанням адаптованих бізнес-орієнтованих метрик HR, VC та ARPR. Отримані результати свідчать про підвищення точності та своєчасності надання індивідуальних предикативних пропозицій B2B-клієнтам, зменшення кількості нерелевантних рекомендацій і підвищення практичної цінності кожного комунікаційного звернення.

Науково-практичну цінність дисертаційного дослідження становлять триетапний конвеєр формування індивідуальних предикативних пропозицій, методика його інтеграції в B2B-платформу та система адаптованих метрик оцінювання ефективності рекомендацій. Запропонований ціннісно-зважений темпоральний підхід до інтелектуального аналізу транзакційно-поведінкових даних B2B-клієнтів впроваджено в діяльність ФОП (NEWOPT) та ТОВ «СТІЛЛІ», а також використано в навчальному процесі та при виконанні НДР Національного університету «Одеська політехніка» на кафедрі інформаційних систем.

Ключові слова: інтелектуальний аналіз даних; B2B-системи електронної комерції, методи класифікації, персоналізація контенту, рекомендаційні системи, транзакційно-поведінкові данні.

Список публікацій здобувача за темою дисертації

1. Арсірій О., Іванов Д. Моделі афінитивного аналізу транзакційних і поведінкових даних клієнтів систем B2B електронної комерції для створення персоналізованого контенту Праці Одеського політехнічного університету. 2025. 2(72) С 90-101(<https://doi.org/10.15276/opu.2.72.2025>)
<https://pratsi.op.edu.ua/index.php/pratsi/article/view/549/511> ISSN 2076-2429 (print)
2. Arsirii , O., & Ivanov , D. (2026). A Model and Method of Transactional-behavioral Data Mining for B2B Content Personalization. Electronics and Control Systems, 1(87), 14–23. <https://doi.org/10.18372/1990-5548.87.20881> ISSN 1990-5548
3. Arsirii, O., & Ivanov , D. (2026). Data Mining of B2b Transactions Using a High-utility Temporal Approach. Electronics and Control Systems, 2(88), 16–31. <https://doi.org/10.18372/1990-5548.88.20958> ISSN 1990-5548
4. Olena O. Arsirii, Dmitriy V. Ivanov, Oksana Yu. Babilunha Analysis of models, methods and technologies for creating personalized content in B2B e-commerce systems. (2025). Інформатика. <https://doi.org/10.15276/ict.02.2025.41> ISSN 2522-1523 (Online)
5. Арсірій О. О., Любомська О. М., Руденко О. В., Іванов Д. В. Гібридна рекомендаційна система для підтримки UI/UX дизайнерів Матеріали міжнародної конференції Інформатика. Культура. Техніка. 2024; Том 1 № 1: 29–35
<https://ict.od.ua/index.php/journal/en/article/view/93> <https://doi.org/10.15276/.01.2024.03> ISSN 2522-1523 (Online)
6. Сіроштан М.В., Іванов Д.В, Арсірій О.О. Дослідження ефективності методів машинного навчання на прикладі задачі класифікації предметів одягу. Project, Program, Portfolio Management. РЗМ-2024: Тези доповідей IX Міжнародної науково-практичної конференції : [у 2т.]. Відповідальний за випуск П.О. Тесленко Том 1. Одеса: ІШІР, 2024 С. 201-204 <https://rep.nuos.edu.ua/server/api/core/bitstreams/f2c763dc-2728-47ca-af64-31618bd2a659/content> ISSN 2522-9435
7. Іванов Д.В. Арсірій О.О. Класифікація B2B-клієнтів системи електронної комерції за ціннісно-зваженим темпоральним профілем. Матеріали XVI міжнародної конференції "Modern Information Technology 2026", ІКС, Національний університет "Одеська політехніка", Одеса, Україна, 14-15 травня, 2026. – С. 118-121. https://drive.google.com/drive/folders/1uXN7b84231YhSfT_tY6I9eMhPjDrIBKJ ISSN 2708-311X

ABSTRACT

Ivanov D.V. B2B Customer Data Mining for Content Personalization in E-commerce Systems. – Qualification scientific work in the form of manuscript.

Thesis for the PhD degree in specialty 122 Computer science. – Odessa Polytechnic National University, Ministry of Education and Science of Ukraine, Odesa, 2026.

The Introduction substantiates the relevance of creating personalized content in e-commerce systems to improve the timeliness of generating predictive offers in accordance with B2B customer classes through the data mining of transactional and behavioral datasets. **The object of the study** is defined as the process of content personalization in B2B e-commerce systems based on the transactional and behavioral data of customers. **The subject of the study** consists of models and methods for the data mining of transactional and behavioral data of B2B customers in e-commerce systems. **The purpose of the study** is formulated as solving the scientific and applied problem of content personalization in e-commerce systems by increasing the efficiency of generating predictive offers in accordance with B2B customer classes through the development of an intelligent utility-weighted temporal approach to the analysis of transactional and behavioral datasets. Achieving this research purpose involves **solving the following tasks**: to analyze existing content personalization technologies in e-commerce systems and determine ways to increase their efficiency for B2B customers under modern conditions; to develop a utility-weighted temporal approach to the data mining of transactional and behavioral datasets of B2B customers; to develop a B2B customer classification method for predictive content personalization in e-commerce systems; to model content personalization strategies by increasing the efficiency of predictive offer generation compared to classical approaches and to integrate the developed analytical module into the recommendation system for a real-world B2B platform

The first chapter analyzes the state of existing content personalization technologies in e-commerce systems and identifies ways to increase their efficiency for B2B customers under modern conditions. The specifics of the B2B segment in a dynamic market environment are substantiated. It is shown that more than 60 % of Ukrainian B2B companies state that the main obstacle to moving online is the inability to provide

customers with relevant content, namely personalized prices, real-time technical documentation, and predictive automated recommendations for executing new transactions to replenish warehouse inventories. Studies indicate that only 12 % of domestic B2B platforms utilize data mining elements to personalize offers, whereas globally this figure exceeds 50 %.

The results of the analysis of the B2B e-commerce segment according to the «Wholesale, Marketing, Customer, Collaborative» framework indicate that, unlike B2C, where emotional and one-time purchases dominate, B2B transactions are characterized by cyclicity, rationality, and a long customer lifetime value. This proves that personalization in B2B systems must be based on deep data mining of utility and temporal parameters, fundamentally differing from B2C systems where content is generated primarily through the statistical analysis of purchase frequencies and general behavioral patterns of the mass consumer.

Traditional methods of customer interaction in e-commerce systems based on recommendation generation are analyzed, noting that the functionality of collaborative and content-based filtering modules has certain disadvantages when considering the operational complexity of B2B processes. Thus, it is revealed that classical recommendation systems often ignore temporal dynamics and the specific value of goods, leading to low relevance of offers in sectors with rare but high-value purchases. Furthermore, it is shown that despite attempts to create complex hybrid recommendation systems as add-ons to B2B e-commerce systems by combining collaborative and content-based approaches, this approach also suffers from ignoring economic utility and transaction cycle stability, failing to guarantee the rationality of recommendations for a business customer.

Consequently, the analysis of content personalization technologies using basic recommendation system frameworks revealed a systemic deficit of tools for incorporating the utility-temporal specifics of B2B transactions. This justifies the necessity of employing data mining methods on B2B customer transactional and behavioral data to discover associations and sequences. However, a comparative analysis of data mining

methods aimed at identifying associations and sequences in transactional datasets allowed for the establishment of the following:

- it was found that classical association rule mining (ARM) methods, particularly the Apriori, FP-Growth, and Eclat algorithms, focus on finding static relationships between items in a transaction. In doing so, only the frequency of items appearing in transactions is considered, completely neutralizing information regarding their cost or purchase volume. As a result, strategically important but rare B2B products are discarded, while cheap high-frequency corporate accessories become the core of the model. All of this demonstrates the complete «financial blindness» of classical data mining methods;

- it was found that classical sequential pattern mining (SPM) methods, such as GSP, PrefixSpan, and SPADE, aim to discover correlations between transactions against a chronological backdrop of events. While they partially address the time factor, they are characterized by a high level of "identification discreteness," treating the customer as an abstract sequence of distinct identifiers. This makes it impossible to detect individual transaction rhythms, leading to the «temporal blindness» of SPM methods.

Thus, it is revealed that the existing data mining toolkit requires adaptation through the integration of financial utility metrics and time intervals between transactions. At the end of the chapter, based on an analysis of the capabilities of the widespread industry standard, namely the RFM (Recency, Frequency, Monetary) model for customer segmentation in the B2B environment, its significant limitations are proven. Specifically, the Recency metric captures only a static point on the timeline without considering the volatility of the individual purchasing cycle, while the Monetary metric evaluates gross revenue, ignoring specific utility and the stability of consumption patterns.

Thus, the classical RFM analysis model is oriented toward a retrospective description of state («what happened»), which prevents building predictive CLV (Customer Lifetime Value) models with high accuracy for B2B customers with irregular activity rhythms. Summarizing the results of the state and trend analysis of the industry, a critical need to develop new approaches to B2B content personalization is established. The expediency of shifting from static segmentation to constructing a utility-weighted temporal customer profile is substantiated. This will allow overcoming the financial and

temporal blindness of existing models, ensuring a transition from merely stating past transactions to the predictive management of communication triggers based on the forecasted purchasing cycle, financial capacity, and behavioral stability of each specific counterparty.

The second chapter develops a utility-weighted temporal approach to the data mining of transactional and behavioral datasets of B2B customers. The systemic investigation of existing association rule mining and sequential pattern mining methods in accumulated B2B transaction arrays resulted in establishing fundamental methodical limitations that prevent their direct application for creating personalized content in e-commerce systems.

The existence of five critical gaps (limitations) between the mathematical assumptions of association and sequence mining methods and the nature of the B2B domain is substantiated, namely: the «financial blindness» of binary models, the occurrence of a «combinatorial explosion» during weighted analysis, the «temporal blindness» of static models, the presence of identification discreteness, and the lack of a differentiated customer interaction strategy. It is established that these limitations are conceptually interconnected and their relationship possesses a hierarchical structure.

That is, an attempt to overcome the financial blindness of classical data mining models by adding cost metrics activates the combinatorial explosion problem. In turn, moving to the analysis of temporal intervals in transactional data requires not only a new processing architecture but also the elimination of identification discreteness for B2B customers, since calculating cyclicity is possible only on integral rather than fragmented time series. Thus, the final limitation – the lack of a differentiated customer interaction strategy, i.e., strategic monotony – is a direct result of data deficits at all previous stages of analysis. In other words, in a B2B e-commerce system, it is impossible to differentiate (determine classes of) customers without identifying their individual financial and temporal patterns.

Therefore, overcoming these limitations serves as the scientific foundation for developing a utility-weighted temporal analysis, the mathematical apparatus of which is conventionally divided into two components: a utility-weighted transaction model and a

utility pattern growth method (UP-Growth), alongside a temporal sequence model and a method for the temporal analysis of utility patterns. A utility-weighted B2B transaction model has been developed by adding the formalization of a total transaction utility indicator $U(T)$ to the classical transaction definition, calculated as the sum of the products of the quantities of purchased item components and their individual prices. To overcome the combinatorial explosion effect that arises when using this utility indicator, the definition of Transaction-Weighted Utilization $U(T)$ is formalized as the total value of all transactions in the database that contain the specified transaction. Using the utility-weighted B2B transaction model, the UP-Growth utility pattern mining method was developed. It is shown that its application eliminates the «financial blindness» of the classical approach to discovering associative patterns in transactional data.

Thus, during the first pass of UP-tree construction, nodes accumulate the monetary value of items rather than their support count. Utilizing the $U(T)$ mechanism during the second pass restores the downward closure (anti-monotonicity) property, ensuring computational stability even when analyzing long B2B invoices. The output of the method is an ordered list of utility patterns – combinations of goods that generate the main share of the enterprise's revenue. A temporal B2B sequence model has been developed by adding the transaction occurrence time to the classical sequence definition, which allowed obtaining an *Inter-Purchase Time (IPT)* vector for an arbitrary B2B customer. This, in conjunction with a calculated dimensionless indicator – the Coefficient of Variation CV – provides the capability to determine transaction cycle stability, thereby overcoming the «temporal blindness» and identification discreteness of classical sequential pattern mining.

Using the temporal B2B sequence model, a method for the temporal analysis of utility patterns obtained via the UP-Growth method was developed. It is shown that the implementation of this method eliminates the «temporal blindness» of the classical approach to sequential pattern mining by constructing *IPT* vectors and their subsequent evaluation against the CV criterion for any B2B customer. An interpretation of the CV criterion depending on the established stability threshold is provided. The value of the threshold is not fixed and can be adjusted depending on industry specifics and the

enterprise's tolerance for prediction errors. In most practical applications, the recommended value is in a range from 0,2 to 0,5. For utility-weighted sequential patterns that pass the verification check under the *CV* criterion, the date of the predictive offer to the B2B customer is calculated as the moment of communication trigger activation, factoring in the preventiveness coefficient, which determines the warning depth. As a general summary of the chapter, it is demonstrated that using the proposed two-stage pipeline consisting of the UP-Growth utility pattern mining method and its temporal analysis yields a list of personalized triggers.

Thanks to this, the moment for receiving a predictive offer is determined for each counterparty and each financially significant and cyclically stable transaction. As confirmation, a detailed examination of two typical scenarios for B2B e-commerce systems – «regular seasonal purchasing» and «irregular purchasing» – is provided. It is precisely the capability to implement these scenarios that transforms a B2B platform from a reactive catalog into an intelligent recommendation system capable of anticipating customer needs regarding warehouse inventory replenishment and interacting with them at the most favorable moment.

The third chapter develops a B2B customer classification method for predictive content personalization in e-commerce systems. The existence of three critical limitations of using classical RFM analysis for B2B customer classification is proven. It is shown that using the Recency parameter *R* only captures the time elapsed since the last transaction. In doing so, an *IPT* vector of inter-purchase patterns is not constructed, making the *R* metric «temporally blind» to the individual rhythm of B2B operations. Using the Frequency parameter *F* evaluates the total number of transactions over a period, yet ignores transaction stability, the indicator of which is the *Coefficient of Variation CV* calculated based on IPT. The Monetary volume parameter *M* is calculated as the total expenditure of the customer, but in this form, it provides no information on transaction structure. Furthermore, it does not account for Transaction-Weighted Utilization *TWU*, which prevents isolating the driver-product patterns that ensure turnover stability.

To overcome these drawbacks of standard RFM analysis, a utility-weighted temporal customer profile (UW-TCP) is formalized as a tuple of features.

To generate predictive offers for B2B customers, a classification method was developed, which constitutes the third component of the utility-weighted temporal analysis of B2B customer transactional and behavioral data in an e-commerce system. It functions exclusively as an add-on over the proposed UP-Growth utility pattern mining method and the method for their temporal analysis. The application of the B2B customer classification method is based on the features of the B2B customer's utility-weighted temporal profile, which are aggregates of already available intermediate results, making the classification method the third stage of the pipeline for generating predictive offers for B2B customers. The assignment of B2B customer class IDs is integrated into the communication trigger generation system. This provides a meaningful link between the customer class and the interaction strategy through a mathematically substantiated assignment of the trigger window width according to customer value, rather than through a subjective definition of segmentation thresholds as seen in standard RFM analysis.

The fourth chapter presents an experimental validation of the effectiveness of the proposed utility-weighted temporal approach for generating predictive offers in a B2B e-commerce platform. The developed models and methods are considered as components of a recommendation module integrated into the operational workflow of the real B2B platform “Baza Vzuttya”. The interaction between the platform, the analytical pipeline, and the manager interface is formalized, where order history, customer data, product items, SKUs, prices, and transaction amounts are used to generate active personalized recommendations.

The experimental study was conducted on a transactional-behavioral dataset for 2022–2025, which included 29.35 thousand orders, 2.53 thousand customers, 56.20 thousand unique SKUs, and a total order volume of UAH 416.28 million. After preprocessing, forming the transactional representation, and filtering excessively long orders, 28.53 thousand orders were used for further analysis, which made it possible to reduce combinatorial complexity without losing the main part of the experimental dataset.

To evaluate effectiveness, the reverse ablation study methodology was tested. Four scenarios were sequentially compared: the baseline frequency-based scenario S0, the utility-weighted scenario S1, the utility-temporal scenario S2, and the integrated scenario

S3, which additionally takes into account the B2B customer class. The results showed a consistent improvement in recommendation quality: conversion increased from 3.91 % in S0 to 5.39 % in S1, 12.69 % in S2, and 15.91 % in S3. At the same time, the number of generated recommendations decreased from 920 to 220, that is, by 76.09 %, indicating a reduction in excessive communication load on managers and customers.

The best result was obtained for the integrated scenario S3. It generated 220 recommendations, 35 of which were successful, ensured an average revenue per recommendation of UAH 6.91 thousand, and covered test revenue of UAH 1.52 million. Compared with the baseline scenario S0, the number of recommendations decreased by 700, while the number of successful interactions remained almost unchanged. At the same time, ARPR increased from UAH 1.21 thousand to UAH 6.91 thousand, that is, approximately 5.7 times. This confirmed that the full model filters out a significant share of irrelevant recommendations and increases the practical value of each interaction.

The economic evaluation also demonstrated the advantage of the integrated approach. Assuming the cost of one communication interaction is UAH 150, operational costs decreased from UAH 138.00 thousand in the baseline scenario to UAH 33.00 thousand in scenario S3. The cost of one successful recommendation decreased from UAH 3.83 thousand to UAH 0.94 thousand, while the net financial result of the integrated model amounted to UAH 1.49 million. The obtained results prove that the proposed approach enables a transition from mass recommendation generation to more accurate, timely, and economically justified personalization of B2B offers.

The place of the developed system in the architecture of the B2B platform was defined as an applied recommendation module that extends the platform's functionality without changing its core transactional processes. The software implementation was carried out as an analytical module that receives data on customers, orders, and product items, applies the developed three-stage pipeline, and generates active predictive recommendations for the manager. To demonstrate the practical use of the results, interface screens were implemented for the manager dashboard, active recommendations, and a personalized offer for a B2B customer.

The scientific novelty of the obtained results is as follows:

– Improved: The utility-weighted transaction model. Unlike traditional representations in the form of a flat stock-keeping unit (SKU) list, it additionally incorporates SKU volume and individual price. This enables the subsequent use of their product to calculate Transaction-Weighted Utilization (\$TWU\$) and identifies transactions not only by the repetition criterion but also by their economic significance, thereby overcoming the "financial blindness" limitation;

– Further developed: The utility pattern growth (UP-Growth) method. By utilizing the Transaction-Weighted Utilization (\$TWU\$) metric to compute the values of UP-tree nodes instead of the frequency counters used in classical FP-Growth, the downward closure (anti-monotonicity) property has been restored, overcoming the "combinatorial explosion" limitation. Furthermore, this shifts the priority of association rule generation, focusing the personalization system on economically significant transactions;

– Improved: The temporal sequence model. Unlike traditional representations in the form of tuples of identificationally discrete transactions, it accounts for a deterministic B2B customer identifier and additionally includes the transaction timestamp. This allows for the creation of an Inter-Purchase Time *IPT* vector and the utilization of its statistical characteristics to calculate the probable activation date of a predictive offer, overcoming the limitations of "temporal blindness" and identification discreteness;

– Further developed: The method for the temporal analysis of UP-Growth utility patterns. Working in conjunction with the temporal sequence model, it enables the calculation of the Coefficient of Variation *CV* of the Inter-Purchase Time vector. This provides the capability to determine the exact date of a predictive offer to a B2B customer, generating predictive triggers in the form of «customer – product set – interaction date» to increase the level of content personalization in B2B e-commerce systems;

– Developed for the first time: A B2B customer classification method for e-commerce systems. Unlike the existing RFM analysis, it is based on a utility-weighted temporal customer profile formed through a two-stage analysis pipeline of B2B customer transactional and behavioral data. This enabled the implementation of a three-stage

pipeline for generating individual predictive offers depending on the determined customer class, ultimately increasing the final accuracy of predictive content personalization.

The practical significance of the research results lies in the development of a software analytical module for generating personalized predictive offers in B2B e-commerce systems. The module is based on a three-stage pipeline that combines utility-weighted analysis of product patterns, temporal verification of purchasing cycle stability, and classification of B2B customers based on a utility-weighted temporal profile. This makes it possible to generate offers not only on the basis of product relevance, but also taking into account the financial significance of the pattern, the predicted moment of interaction, and the customer class.

The practical approbation of the results was carried out by integrating the developed pipeline as a Python module into the operational workflow of the B2B platform «Baza Vzuttya». The implemented module uses data on customers, orders, product items, SKUs, prices, and transaction amounts, after which it generates active recommendations for the manager. To demonstrate the applied use of the results, interface screens were developed for the manager dashboard, active recommendations, and a personalized offer for a B2B customer.

The effectiveness of the proposed approach was confirmed by an experimental comparison of the baseline frequency-based scenario and the full integrated model using adapted business-oriented metrics. The obtained results indicate an increase in the accuracy and timeliness of providing individual predictive offers to B2B customers, a reduction in the number of irrelevant recommendations, and an increase in the practical value of each communication interaction.

The scientific and practical value of the dissertation research consists of the three-stage pipeline for generating individual predictive offers, the methodology for its integration into a B2B platform, and the system of adapted metrics for evaluating recommendation effectiveness. The proposed utility-weighted temporal approach to the data mining of B2B customers' transactional and behavioral data has been implemented into the operations of Sole Proprietor (NEWOPT) and LLC "STILLI", and has also been utilized in the educational process and during the execution of R&D (Research and

Development) projects at the Department of Information Systems of Odesa National Polytechnic University.

Keywords: data mining; B2B e-commerce systems; B2B customer classification; content personalization; recommendation systems; transactional and behavioral data.

List of publications of the applicant on the topic of the dissertation

1. Arsirii, O., & Ivanov, D. (2025). Models of affinity analysis of B2B e-commerce customers' transactional and behavioral data for personalized content creation]. *Pratsi Odeskoho Politekhnichnoho Universitetu*, 2(72), 90–101. <https://doi.org/10.15276/opu.2.72.2025>
2. Arsirii, O., & Ivanov, D. (2026). A Model and Method of Transactional-behavioral Data Mining for B2B Content Personalization. *Electronics and Control Systems*, 1(87), 14–23. <https://doi.org/10.18372/1990-5548.87.20881>
3. Arsirii, O. O., & Ivanov, D. V. (2026) Data mining of B2B customer transactional data based on a utility-weighted temporal approach]. *Elektronika ta Systemy Upravlinnia*, 2(88), 16–31. <https://doi.org/10.18372/1990-5548.88.20958>
4. Arsirii, O. O., Ivanov, D. V., & Babilunha, O. Yu. (2025). Analysis of models, methods and technologies for creating personalized content in B2B e-commerce systems. *Informatyka. Kultura. Tekhnika*, 2, 273–278. <https://doi.org/10.15276/ict.02.2025.41>
5. Arsirii, O. O., Liubomska, O. M., Rudenko, O. V., & Ivanov, D. V. (2024). Hibrydna rekomendatsiina systema dlya pidtrymky UI/UX dyzaineriv [A hybrid recommendation system for supporting UI/UX designers]. *Materialy mizhnarodnoi konferentsii "Informatyka. Kultura. Tekhnika"*, 1(1), 29–35. <http://dspace.opu.ua/jspui/handle/123456789/14629>
6. Siroshtan, M. V., Ivanov, D. V., & Arsirii, O. O. (2024). Investigation of machine learning methods efficiency on the example of clothing items classification problem. *Project, Program, Portfolio Management (P3M-2024): Tezy dopovidei IX Mizhnarodnoi naukovo-praktichnoi konferentsii*, 1, 201–204. <https://rep.nuos.edu.ua/server/api/core/bitstreams/f2c763dc-2728-47ca-af64-31618bd2a659/content>
7. Ivanov, D. V., & Arsirii, O. O. (2026). B2B customer classification in an e-commerce system based on a utility-weighted temporal profile. *Materialy XVI mizhnarodnoi konferentsii "Suchasni Informatsiini Tekhnologii 2026" (Modern Information Technology 2026)*, Odesa, Ukraine, 118–121. https://drive.google.com/drive/folders/1uXN7b84231YhSfT_tY6I9eMhPjDrIBKJ