

МЕТОДИ DATA SCIENCE: АНАЛІЗ ПІДХОДІВ ДО СЕГМЕНТАЦІЇ КЛІЄНТІВ

© 2025 АНДРУСИК Є. В., КАГАНОВСЬКИЙ О. С.

УДК 33.330.4
JEL Classification: C1; D4; D9

Андрусик Є. В., Кагановський О. С. Методи Data Science: аналіз підходів до сегментації клієнтів

У статті розглянуто проблему підвищення ефективності маркетингових стратегій в умовах сучасного динамічного ринкового середовища, що характеризується високим рівнем конкуренції, інформаційним перевантаженням і швидкою зміною споживчих вподобань. Проведений аналіз показав, що традиційні методи сегментації, засновані на демографічних, географічних та соціально-економічних характеристиках, не дозволяють враховувати індивідуальні потреби, вподобання та поведінкові особливості споживачів, що знижує ефективність маркетингових кампаній. Зазначено, що в цих умовах на перший план виходять новітні альтернативні підходи, що ґрунтуються на використанні неструктурованих даних, динамічній сегментації та психографічних факторах. Реалізація альтернативних підходів неможлива без застосування таких методів Data Science, як *k-means*, ієрархічна кластеризація та DBSCAN, які дозволяють виявляти приховані закономірності в поведінці клієнтів і формувати більш точні сегменти. Визначено переваги та недоліки методів кластеризації під час реалізації альтернативних підходів до сегментації клієнтів. Зазначено, що вибір методу сегментації повинен ґрунтуватися на комплексному аналізі в контексті конкретного завдання (селективний підхід) або гібридизації, яка дозволяє підсилити переваги та нівелювати недоліки кожного з методів. Підкреслено, що селекція або гібридизація під час роботи з неструктурованими даними потребує перетворення даних (текстових, аудіо, медіа) на числові. У вирішенні цього завдання допоможе сполучення *K-means*, ієрархічної кластеризації або DBSCAN разом з іншими методами, такими як метод виділення ключових слів, тематичне моделювання, векторні представлення слів, методи обробки природної мови (NLP) або комп'ютерного зору. Для динамічної сегментації можуть бути використані *mini-batch K-means*, що являє собою модифікацію алгоритму *K-means*, яка використовується для кластеризації великих наборів даних. Це дозволяє значно прискорити процес кластеризації, особливо при роботі зі значними обсягами інформації. Наголошено, що психографічні фактори мають, напевно, найбільшу значущість, бо глибше відображають цінності і мотивацію споживачів. Ієрархічна кластеризація дозволяє візуалізувати групи різних психографічних факторів, але інтерпретація таких об'єднань вимагатиме додаткового аналізу за допомогою ШІ. Найбільш доцільним методом для сегментації клієнтів за психографічними факторами є DBSCAN. Цей метод дозволяє виявляти групи з подібними профілями, навіть якщо вони мають складну форму, а також кластери з нелінійними межами. Підкреслено, що подальші дослідження мають бути спрямовані на розробку гібридних моделей, які поєднують різні алгоритми кластеризації з альтернативними підходами для збільшення результативності маркетингової стратегії компанії.

Ключові слова: маркетингова діяльність, динамічна сегментація, психографічні фактори, Data Science, метод *k-means*, *mini-batch K-means*, ієрархічна кластеризація, DBSCAN.

Табл.: 2. Бібл.: 10.

Андрусик Євгеній Валерійович – аспірант кафедри економічної кібернетики і системного аналізу, Харківський національний економічний університет імені Семена Кузнеця (просп. Науки, 9а, Харків, 61166, Україна)

E-mail: andrusike@mail.com

ORCID: <https://orcid.org/0009-0001-3654-5067>

Кагановський Олександр Семенович – доктор філософії, аспірант кафедри менеджменту та бізнесу, Харківський національний економічний університет імені Семена Кузнеця (просп. Науки, 9а, Харків, 61166, Україна)

E-mail: kag.ole.68@gmail.com

ORCID: <https://orcid.org/0009-0008-1965-625X>

UDC 33.330.4
JEL Classification: C1; D4; D9

Andrusyk Y. V., Kaganovskyi O. S. Data Science Methods: Analyzing Approaches to Customer Segmentation

The article addresses the issue of enhancing the efficiency of marketing strategies in the context of the modern dynamic market environment, characterized by a high level of competition, information overload, and rapid changes in consumer preferences. The analysis conducted revealed that traditional segmentation methods based on demographic, geographic, and socioeconomic characteristics fail to account for individual needs, preferences, and behavioral traits of consumers, which diminishes the efficiency of marketing campaigns. It is noted that in these conditions, innovative alternative approaches based on the use of unstructured data, dynamic segmentation, and psychographic factors come to the forefront. The implementation of alternative approaches is impossible without the application of Data Science methods such as *k-means*, hierarchical clustering, and DBSCAN, which allow for the detection of hidden patterns in customer behavior and the formation of more accurate segments. The advantages and disadvantages of clustering methods during the implementation of alternative approaches to customer segmentation have been identified. It is specified that the choice of segmentation method should be based on a comprehensive analysis in the context of a specific task (selective approach) or hybridization, which allows for enhancing the advantages and mitigating the disadvantages of each method. It is underscored that selection or hybridization when working with unstructured data requires the transformation of data (text, audio, media) into numerical form. The combination of *K-means*, hierarchical clustering, or DBSCAN, along with other methods such as keyword extraction, topic modeling, word vector representations, natural language processing (NLP), or computer vision, will assist in solving this task. For dynamic segmenta-

tion, mini-batch K-means, which is a modification of the K-means algorithm used for clustering large datasets, can be employed. This significantly accelerates the clustering process, especially when dealing with substantial amounts of information. It is emphasized that psychographic factors likely hold the greatest significance, as they more deeply reflect the values and motivations of consumers. Hierarchical clustering allows for the visualization of groups of different psychographic factors, but the interpretation of such associations will require additional analysis using AI. The most appropriate method for segmenting clients based on psychographic factors is DBSCAN. This method allows for the identification of groups with similar profiles, even if they have complex shapes, as well as clusters with nonlinear boundaries. It is underlined that further research should be directed towards the development of hybrid models that combine various clustering algorithms with alternative approaches to enhance the effectiveness of the company's marketing strategy.

Keywords: marketing activities, dynamic segmentation, psychographic factors, Data Science, k-means method, mini-batch K-means, hierarchical clustering, DBSCAN.

Tabl.: 2. **Bibl.:** 10.

Andrusyk Yevhenii V. – Postgraduate Student of the Department of Economic Cybernetics and System Analysis, Simon Kuznets Kharkiv National University of Economics (9a Nauky Ave., Kharkiv, 61166, Ukraine)

E-mail: andrusike@mail.com

ORCID: <https://orcid.org/0009-0001-3654-5067>

Kaganovskyi Oleksandr S. – PhD, Postgraduate Student of the Department of Management and Business, Simon Kuznets Kharkiv National University of Economics (9a Nauky Ave., Kharkiv, 61166, Ukraine)

E-mail: kag.ole.68@gmail.com

ORCID: <https://orcid.org/0009-0008-1965-625X>

У сучасних умовах динамічного ринкового середовища, що характеризується високим рівнем конкуренції та інформаційним переважанням, ефективна сегментація клієнтів є одним із ключових факторів успіху будь-якої маркетингової стратегії. Традиційні методи сегментації, засновані на демографічних, географічних і соціально-економічних характеристиках, поступово втрачають свою актуальність, оскільки не враховують індивідуальні потреби, вподобання та поведінкові особливості споживачів. У цій ситуації на перший план виходять підходи, що ґрунтуються на застосуванні методів Data Science. Аналіз великих масивів різноманітних даних дозволяє виявляти приховані закономірності в поведінці клієнтів, проводити більш точну сегментацію та розробляти персоналізовані маркетингові пропозиції, що зрештою впливає на лояльність клієнта до компанії, обсяг продажів, частку ринку.

Підвищення динамічності ринків, скорочення тривалості дії трендів споживачьких вподобань призводить до стрімкого розвитку моделей сегментації клієнтів та формування маркетингових стратегій. Однією з цікавих робіт у цьому напрямку є праця Санкет Лодха, Сонал Дешмух, Судхір Чітніс, Арун Патіл [1], у якій досліджується проблематика сегментації клієнтів в електронній комерції з використанням методів машинного навчання. Автори наголошують на актуальності задачі управління відносинами з клієнтами (CRM) та її впливі на підвищення ефективності бізнес-процесів. В роботі проведено порівняльний аналіз різних методів кластеризації, включаючи k-means, ієрархічну кластеризацію, кластеризацію на основі щільності та методу поширення спорідненості (Affinity Propagation). Особлива увага приділяється

можливості використання гібридних моделей, які поєднують в собі переваги різних алгоритмів. Автори доводять, що гібридні моделі кластеризації можуть забезпечити більш високу точність сегментації порівняно з використанням окремих алгоритмів. Також підкреслюється важливість вибору релевантного методу кластеризації залежно від специфіки даних та цілей дослідження. За результатами дослідження автори роблять висновок, що ефективна сегментація клієнтів за допомогою машинного навчання є важливим інструментом для підвищення прибутковості компаній, покращення взаємовідносин з клієнтами та їх утримання.

Також варто взяти до уваги роботу Мудунурі Варма, Анапарті Ахіл, Кокку Анушка та Прітхві Мадукурі [2], у якій розглянуто використання великих даних для сегментації клієнтів у роздрібній торгівлі. Автори підкреслюють важливість аналізу великих даних для розуміння поведінки споживачів і підвищення ефективності бізнес-процесів. У роботі зроблено опис дизайну дослідження з дедуктивним підходом. Дані було зібрано за допомогою методу вторинного збору даних з різних джерел, включаючи наукові публікації, вебсайти та інші відкриті джерела інформації. Автори демонструють, що роздрібний ринок значно виріс завдяки цифровізації. Великі дані дозволяють роздрібним продавцям аналізувати ринкові тенденції та розуміти поведінку споживачів при покупках. Сегментація клієнтів на основі великих даних є важливим інструментом для збільшення обсягу продажів на роздрібному ринку. Дослідження підтверджує важливість використання великих даних для сегментації клієнтів у роздрібній торгівлі. Автори роблять висновок, що ефективна сег-

ментация клієнтів дозволяє роздрібним продавцям підвищити ефективність маркетингових кампаній і збільшити прибутковість.

Проблему сегментації цифрового ринку розглядає стаття Ашрафа Уддіна, Аламіна Талукдера, Редвана Ахмеда та ін. [9]. Автори досліджують, як машинне навчання може бути використано для визначення та аналізу різних груп цифрових користувачів на основі їхніх інтересів та поведінки в соціальних мережах. В їхній роботі представлено модель машинного навчання, яка використовує різні методи кластеризації, такі як k-means, AGNES та нечіткі C-середні, для аналізу даних соціальних мереж. Автори використовують відкритий набір даних, що містить інформацію про 30 000 учнів старших класів США, зібрану з провідної платформи соціальних мереж. Результати дослідження показують, що метод mini-batch k-means є найефективнішим методом кластеризації для цього набору даних, оскільки демонструє порівнянну продуктивність, тоді як AGNES та нечіткі C-середні є менш ефективними. Дослідження має практичне значення для маркетологів, оскільки дозволяє розробляти цільові рекламні кампанії та маркетингові стратегії для різних груп цифрових аборигенів. Крім того, результати дослідження можуть бути використані для розробки нових продуктів та послуг, які відповідають потребам та інтересам різних сегментів ринку [9].

Заслуговує на увагу дослідження Яна Блемкера та Кармен-Марії Альбрехт [10], що присвячене вивченню впливу індивідуальних психологічних характеристик споживачів на їхню поведінку в умовах багатоканального ринку. Автори використовують латентний класовий аналіз на даних 1512 клієнтів, досліджуючи психографічні змінні з п'яти областей: ставлення до ризику, когнітивні здібності, мотивація, особистість і стиль прийняття рішень.

Дослідження виявляє шість сегментів клієнтів, кожен з яких характеризується унікальним поєднанням психографічних змінних. Наприклад, одні сегменти демонструють високу схильність до ризику та інтуїтивний стиль прийняття рішень, тоді як інші віддають перевагу раціональному підходу та мають високу потребу в тактильному контакті з товаром. Результати дослідження мають практичне значення для компаній, що працюють в умовах багатоканального ринку. Розуміння психологічних факторів, що впливають на вибір каналу покупцем, дозволяє компаніям оптимізувати свої маркетингові стратегії, персоналізувати взаємодію з клієнтами та підвищити ефективність кожного каналу [10].

З українських науковців варто назвати Т. Затонацьку та Я. Фаренюк [3], які зазначають, що сучасні маркетингові стратегії все частіше спираються на можливості штучного інтелекту (ШІ) для досягнення кращих результатів. Алгоритми ШІ дозволяють глибоко проаналізувати потреби та бажання як існуючих, так і потенційних клієнтів. Це дає змогу зрозуміти не тільки хто вони, чого прагнуть та які рішення обирають, але й врахувати їхні емоції, вподобання та ставлення до продукту.

Маркетологи можуть використовувати ШІ для опитування клієнтів щодо задоволеності продуктом, а фахівці з обробки даних (DS) – для аналізу отриманих результатів. ШІ та DS відіграють ключову роль у формуванні маркетингової стратегії, допомагаючи вирішити три основні завдання: сегментація: розділення ринку на групи клієнтів з унікальними потребами; таргетинг: вибір сегментів, на яких фокусуються маркетингові зусилля; позиціонування: підкреслення найкращих характеристик продукту та його конкурентних переваг.

ШІ дозволяє рекомендувати різні цільові групи та застосовувати прогностичне моделювання для вибору релевантного сегмента. Генеративний ШІ є потужним інструментом для розробки слоганів та рекламних кампаній, які резонують з цільовою аудиторією.

У дослідженні І. Перезової, І. Землякова та В. Шайбана [7] зроблений висновок, що персоналізація стала основною стратегією компаній у боротьбі за клієнтів, дозволяючи за допомогою великих обсягів даних про поведінку користувачів створювати персоналізовані рекомендації, що суттєво підвищують показники конверсії. Також автори підкреслюють, що прогнозна сегментація є одним із основних підходів до коректної персоналізації.

Праця В. Стамата та А. Скорука [8] резюмує, що нині сегментація цільової аудиторії є однією з провідних технологій не тільки для створення маркетингової стратегії для досягнення цілей компанії, а й для розробки рекламної кампанії на всіх можливих каналах комунікації для збільшення лояльності споживача до бренду та створення нових послуг, зацікавлення споживачів. Ефективна рекламна стратегія має вирішальне значення – різноманітні дослідження доводять, що 80 % споживачів віддають перевагу компаніям, що пропонують персоналізований контент, який формується за допомогою цільового маркетингу [8].

Огляд сучасних напрацювань і публікацій дозволяє зробити висновок, що сегментація спо-

живачів має величезне значення, оскільки дозволяє бізнесу чітко розуміти, які саме продукти або послуги цікавлять різні групи клієнтів. Це знання допомагає визначити, що слід виключити з персональних пропозицій, а що, навпаки, додати, щоб задовольнити потреби кожного сегмента. Таким чином, сегментація сприяє підвищенню впізнаваності бренду, формуванню лояльності клієнтів, збільшенню доходу бізнесу.

Наголошуючи на безперечній ефективності розглянутих вище наукових розробок, слід зазначити, що недостатньо повно розглянуті альтернативні підходи до сегментації клієнтів, зокрема, динамічна сегментація та методи її реалізації.

Метою цього дослідження є визначення потенціалу застосування методів Data Science для альтернативних підходів до сегментації клієнтів (використання неструктурованих даних, створення динамічної сегментації і включення психографічних факторів), що сприятиме розробці більш точних моделей сегментації з урахуванням сучасних трендів розвитку цифрової економіки і змін тенденцій у споживачької поведінки. Для досягнення цієї мети вирішено такі завдання:

1. Визначено альтернативні підходи до сегментації.
2. Проведено порівняльний аналіз методів кластеризації, що використовуються для сегментації клієнтів.
3. Визначено потенційні переваги та недоліки застосування методів кластеризації в альтернативних підходах до сегментації клієнтів, перспективи їх розвитку.

Як було зазначено вище, в умовах зростаючої конкуренції та динамічних змін на ринку ефективна сегментація клієнтів є ключовим фактором успіху для будь-якої компанії. Традиційні методи сегментації поступово поступаються місцем більш складним підходам, які враховують поведінкові, психографічні та інші фактори.

Не менш важливим аспектом є також і етичні міркування, бо у сучасному світі конфіденційність даних – на першому місці для споживачів. Порушення обов'язків про зберігання та неоприлюднення даних може спричинити суттєве зменшення лояльності до компанії та значний відтік клієнтів.

На першому кроці дослідження у результаті аналізу літературних джерел визначено найбільш перспективні альтернативні підходи до сегментації клієнтів, їхні переваги, недоліки й етичні аспекти використання (табл. 1).

Слід зазначити, що альтернативні підходи до сегментації можуть бути ефективними в прогнозуванні зміни поведінки споживачів, а також розширенні уявлення про вподобання клієнтів.

На другому кроці дослідження проведений порівняльний аналіз поширених методів Data Science, що застосовуються в сегментації клієнтів. Як показав проведений аналіз літературних джерел, до них відносяться такі алгоритми кластеризації, як K-means, ієрархічна кластеризація і DBSCAN. На основі аналізу літературних джерел [4–6] було визначено переваги та недоліки цих методів (табл. 2).

Таблиця 1

Альтернативні підходи до сегментації клієнтів

Підхід	Опис	Переваги	Недоліки	Етичні міркування
Аналіз неструктурованих даних	Дані, які не мають загальної моделі або організації (текст, зображення, аудіо, відео)	Можуть надати глибше розуміння клієнтів, виявити приховані тенденції в поведінці	Складність аналізу, потреба в спеціалізованих інструментах та алгоритмах	Захист конфіденційності, забезпечення анонімності даних, прозорість використання
Динамічна сегментація	Сегментація в реальному часі на основі поточної поведінки клієнтів	Персоналізація взаємодії з клієнтами, підвищення ефективності маркетингових кампаній	Висока вартість впровадження, потреба в постійному моніторингу даних	Уникнення ризику створення "інформаційних бульбашок"
Психографічні фактори	Врахування особистісних характеристик, стилю життя, цінностей, інтересів клієнтів	Глибше розуміння мотивації клієнтів, створення більш цілеспрямованих повідомлень	Складність збору та аналізу даних, суб'єктивність інтерпретації	Повага до індивідуальності, уникнення стереотипів та упереджень

Джерело: розроблено авторами.

Порівняльний аналіз методів кластеризації, які використовуються в сегментації клієнтів

Метод сегментації	Переваги	Недоліки	Точність сегментації	Результат застосування
K-means	Простота реалізації, обчислювальна ефективність, здатність обробляти великі обсяги даних	Чутливість до початкової ініціалізації центрів, необхідність визначення кількості кластерів, схильність до локальних оптимумів	Варіюється залежно від структури даних та вибору кількості кластерів	Розбиття клієнтів на задану кількість груп за принципом мінімізації внутрішньокластерної відстані
Ієрархічна кластеризація	Здатність виявляти кластери різної форми та розміру, наочна візуалізація результатів у вигляді дендрограми	Висока обчислювальна складність для великих наборів даних, складність інтерпретації результатів при великій кількості кластерів	Залежить від вибору метрики відстані та методу зв'язування кластерів	Ієрархічна структура кластерів
DBSCAN	Ефективність для даних зі шумом та викидами, здатність виявляти кластери довільної форми, не вимагає задання кількості кластерів	Чутливість до вибору параметрів (радіус, мінімальна кількість точок), може бути неефективним для даних з різною щільністю	Висока для даних з чітко вираженими кластерами	Розбиття клієнтів на кластери на основі щільності розподілу даних

Джерело: розроблено авторами.

Проведений порівняльний аналіз поширених методів сегментації клієнтів демонструє, що кожен з них має свої переваги та недоліки. Вибір релевантного методу залежить від завдань дослідження та специфіки даних.

K-means відрізняється простотою та швидкістю, що робить його придатним для попереднього аналізу даних та виділення основних груп клієнтів. DBSCAN ефективний для даних зі шумом та викидами, що дозволяє виявляти нестандартні сегменти. Ієрархічна кластеризація має переваги у візуалізації кластерних структур за допомогою дендрограм та гнучкості у визначенні кількості кластерів, але основними недоліками цієї групи методів є обчислювальна складність, чутливість до шуму та викидів.

Слід зазначити, що важливо враховувати не тільки метрики якості класифікації, але й інтерпретованість результатів, щоб забезпечити їх практичну значущість для прийняття управлінських рішень. Загалом вибір методу сегментації повинен ґрунтуватися на комплексному аналізі в контексті конкретного завдання (селективний підхід) або гібридизації, яка дозволяє підсилити переваги та нівелювати недоліки кожного з методів.

На третьому кроці дослідження визначено потенційні переваги та недоліки застосування методів кластеризації в альтернативних підходах до сегментації клієнтів, перспективи їх розвитку.

Слід зазначити, що селекція або гібридизація під час роботи з неструктурованими даними потребує перетворення даних (текстових, аудіо, медіа) на числові. У вирішенні цього завдання допоможе сполучення K-means, ієрархічної кластеризації або DBSCAN разом з іншими методами, такими як метод виділення ключових слів, тематичне моделювання, векторні представлення слів, методи обробки природної мови (NLP) або комп'ютерного зору.

Для динамічної сегментації можуть бути використані mini-batch K-means, що являє собою модифікацію алгоритму K-means, яка використовується для кластеризації великих наборів даних. Замість того, щоб обробляти весь набір даних на кожній ітерації, як це робить стандартний K-means, mini-batch K-means використовує невеликі випадкові вибірки (міні-пакети) даних. Це дозволяє значно прискорити процес кластеризації, особливо при роботі зі значними обсягами інформації. Незважаючи на підвищену швидкість порівняно зі стандартним K-means, алгоритм mini-batch K-means зберігає певну обчислювальну складність, особливо при обробці великих наборів даних з високою розмірністю та великою кількістю кластерів. Така складність обумовлена ітеративним характером, необхідністю багаторазового обчислення відстаней та оновлення центрів, а також впливом випадковості, розмірності даних та пошуку оптимальної кількості кластерів.

Використання ієрархічної кластеризації для динамічної сегментації потребує постійної перебудови дендрограми при кожному оновленні даних. Такий підхід має певний ризик неефективності, що обумовлено прагненням ієрархічної кластеризації виявити статичну структуру, тоді як динамічна сегментація фокусується на змінах у часі. Найбільш доречним буде поєднання динамічної сегментації разом із DBSCAN. Цей метод дозволяє виявляти кластери довільної форми та обробляти дані з «шумом». Крім того, обробка нових даних та оновлення кластерів відбуватиметься без необхідності повної перебудови моделі.

Психографічні фактори мають, напевно, найбільшу значущість, бо глибше відображають цінності і мотивацію споживачів, що при правильній інтерпретації стане, безумовно, значною перевагою у конкурентоспроможності компанії. Ієрархічна кластеризація здатна на демонстрацію об'єднання в групи різних психографічних факторів, але інтерпретація таких об'єднань буде вимагати додаткового аналізу за допомогою ШІ або кваліфікованого працівника. Для сегментації клієнтів за психографічними факторами DBSCAN є найбільш доцільним методом. Цей метод дозволяє виявляти групи з подібними профілями, навіть якщо вони мають складну форму, а також кластери з нелінійними межами.

ВИСНОВКИ

Таким чином, проведений аналіз підтверджує важливість сегментації клієнтів в сучасних маркетингових стратегіях з використанням методів Data Science. В умовах динамічного ринкового середовища та зростаючої конкуренції ефективна сегментація дозволяє компаніям глибше зрозуміти потреби клієнтів, підвищити ефективність маркетингових кампаній та збільшити прибутковість.

Традиційні методи сегментації, засновані на демографічних, географічних та соціально-економічних характеристиках, поступово втрачають актуальність. На перший план виходять новітні підходи, що ґрунтуються на використанні неструктурованих даних, динамічній сегментації та психографічних факторах. Застосування методів Data Science, таких як k-means, ієрархічна кластеризація та DBSCAN, дозволяє виявляти приховані закономірності в поведінці клієнтів та формувати більш точні сегменти. Однак кожен метод кластеризації має свої переваги та недоліки, і вибір релевантного методу залежить від специфіки даних та завдань. Перспективним напрямком є розробка гібридних моделей, що поєднують різні алгоритми кластеризації з альтернативними підходами до сегментації.

Важливим аспектом є також і етичні міркування при використанні даних клієнтів. Компанії повинні забезпечити конфіденційність та анонімність даних, а також прозорість їх використання.

Подальші дослідження мають бути спрямовані на розробку гібридних моделей, що поєднують різні алгоритми кластеризації з альтернативними підходами для збільшення результативності маркетингової стратегії компанії. ■

БІБЛІОГРАФІЯ

1. Sanket Lodha, Dr. Sonal Deshmukh, Dr. Sudhir Chitnis, Dr. Arun Patil. Customer segmentation using machine learning. *Shodhak: A Journal of Historical Research*. 2023 Vol. 53. Issue 3, no. 5. P. 141–147. URL: https://www.researchgate.net/publication/376396210_CUSTOMER_SEGMENTS_USING_MACHINE_LEARNING
2. Mudunuri Varma, Anaparthi Akhil, Kokku Anoushka, Prithvi Maddukuri. Use of Big Data in the Process of Customer Segmentation in the Retail Sector. *Technoarete Transactions on Advances in Data Science and Analytics*. 2022 Vol. 1, Issue 2. P. 7–12. DOI: <https://doi.org/10.36647/TTADSA/01.02.A002>
3. Затоначька Т., Фаренюк Я. Використання Data Science технологій в Е-комерції: успішні кейси *Міжнародна науково-практична конференція Data Science та інформаційно-аналітичні системи: застосування в економіці та фінансах* : тези доп. 23 трав. 2024 р. Київ, 2024. С. 26–30. URL: <https://goo.su/5YyGA>
4. Qian Zhou, Bo Sun Adaptive K-means clustering based under-sampling methods to solve the class imbalance problem. *Data and Information Management*. 2024, Vol. 8, Issue 3. P. 1–12. DOI: <https://doi.org/10.1016/j.dim.2023.100064>
5. Zaitsev D., Filimonchuk T., Huk A., Maistrenko H. Огляд засобів ефективної сегментації зображень з використанням методів кластеризації даних. *Системи управління, навігації та зв'язку*. 2024. № 1 (75). С. 77–81. DOI: <https://doi.org/10.26906/SUNZ.2024.1.077>
6. Dingsheng Deng Application of DBSCAN Algorithm in Data Sampling. *J. Phys.: Conf. Ser.* 1617012088. URL: <https://stats.iop.org/article/10.1088/1742-6596/1617/1/012088>
7. Перезовова І. В., Земляков І. С., Шайбан В. М. Алгоритми персоналізації контенту у web-маркетингу як чинник підвищення конверсії інтернет-продажів. *Академічні візії*. 2023. № 24. URL: <https://www.academy-vision.org/index.php/av/article/view/1340>
8. Стамат В. М., Скорук А. Ю. Сегментація цільової аудиторії як важливий етап маркетингу на ринку готельно-ресторанного бізнесу. *Електронне наукове фахове видання з економічних наук «Modern Economics»*. 2022. № 35. С. 112–117. DOI: [https://doi.org/10.31521/modecon.V35\(2022\)-17](https://doi.org/10.31521/modecon.V35(2022)-17)

9. Ashraf Uddin, Alamin Talukder, Redwan Ahmed et al. Data-driven strategies for digital native market segmentation using clustering. *International Journal of Cognitive Computing in Engineering*. 2024. Vol. 5. P. 178–191.
DOI: <https://doi.org/10.1016/j.ijcce.2024.04.002>
10. Blömker J., Carmen-Maria A. Psychographic segmentation of multichannel customers: investigating the influence of individual differences on channel choice and switching behavior. *Journal of Retailing and Consumer Services*. 2024. Vol. 79. P. 1–17.
DOI: <https://doi.org/10.1016/j.jretconser.2024.103806>

REFERENCES

- Ashraf, Uddin et al. "Data-driven strategies for digital native market segmentation using clustering". *International Journal of Cognitive Computing in Engineering*, vol. 5 (2024): 178-191.
DOI: <https://doi.org/10.1016/j.ijcce.2024.04.002>
- Blomker, J., and Carmen-Maria, A. "Psychographic segmentation of multichannel customers: investigating the influence of individual differences on channel choice and switching behavior". *Journal of Retailing and Consumer Services*, vol. 79 (2024): 1-17.
DOI: <https://doi.org/10.1016/j.jretconser.2024.103806>
- Dingsheng, Deng. "Application of DBSCAN Algorithm in Data Sampling". *J. Phys.: Conf. Ser.* 1617 012088. <https://stats.iop.org/article/10.1088/1742-6596/1617/1/012088>
- Mudunuri, Varma et al. "Use of Big Data in the Process of Customer Segmentation in the Retail Sector". *Technoarete Transactions on Advances in Data Science and Analytics*, vol. 1, no. 2 (2022): 7-12.
DOI: <https://doi.org/10.36647/TTADSA/01.02.A002>
- Perevozova, I. V., Zemliakov, I. S., and Shaiban, V. M. "Alhorytmy personalizatsii kontentu u web-marketynhu yak chynnyk pidvyshchennia konversii internet-prodazhiv" [Content Personalization Algorithms in Web Marketing as a Factor in Increasing Online Sales Conversion]. *Akademichni vizii*. 2023. <https://www.academy-vision.org/index.php/av/article/view/1340>
- Qian, Zhou, and Bo, Sun. "Adaptive K-means clustering based under-sampling methods to solve the class imbalance problem". *Data and Information Management*, vol. 8, no. 3 (2024): 1-12.
DOI: <https://doi.org/10.1016/j.dim.2023.100064>
- Sanket, Lodha et al. "Customer segmentation using machine learning". *Shodhak: A Journal of Historical Research*. 2023. https://www.researchgate.net/publication/376396210_CUSTOMER_SEGMENTATION_USING_MACHINE_LEARNING
- Stamat, V. M., and Skoruk, A. Yu. "Sehmentatsiia tsilyovoi audytorii yak vazhlyvyi etap marketynhu na rynku hotelno-restorannoho biznesu" [Target Audience Segmentation as an Important Stage of Marketing in the Hotel and Restaurant Business Market]. *Modern Economics*, no. 35 (2022): 112-117.
DOI: [https://doi.org/10.31521/modecon.V35\(2022\)-17](https://doi.org/10.31521/modecon.V35(2022)-17)
- Zaitsev, D. et al. "Ohliad zasobiv efektyvnoi sehmentatsii zobrazhen z vykorystanniam metodiv klasteryzatsii danykh" [Overview of Effective Image Segmentation Tools Using Data Clustering Methods]. *Systemy upravlinnia, navihatsii ta zviazku*, no. 1(75) (2024): 77-81.
DOI: <https://doi.org/10.26906/SUNZ.2024.1.077>
- Zatonatska, T., and Farenjuk, Ya. "Vykorystannia Data Science tekhnolohii v E-komertsii: uspishni keisy" [Using Data Science Technologies in E-commerce: Successful Cases]. *Data Science ta informatsiino-analitychni systemy: zastosuvannia v ekonomitsi ta finansakh*. 2024. <https://goo.su/5YyGA>