

This is a self-archived version of an original article. This version may differ from the original in pagination and typographic details.

Author(s): Salminen, Joni; Mustak, Mekhail; Sufyan, Muhammad; Jansen, Bernard J.

Title: How can algorithms help in segmenting users and customers? : A systematic review and research agenda for algorithmic customer segmentation

Year: 2023

Version: Published version

Copyright: © The Author(s) 2023

Rights: CC BY 4.0

Rights url: <https://creativecommons.org/licenses/by/4.0/>

Please cite the original version:

Salminen, J., Mustak, M., Sufyan, M., & Jansen, B. J. (2023). How can algorithms help in segmenting users and customers? : A systematic review and research agenda for algorithmic customer segmentation. *Journal of Marketing Analytics*, 11(4), 677-692.

<https://doi.org/10.1057/s41270-023-00235-5>



How can algorithms help in segmenting users and customers? A systematic review and research agenda for algorithmic customer segmentation

Joni Salminen¹ · Mekhail Mustak² · Muhammad Sufyan³ · Bernard J. Jansen⁴

Revised: 11 April 2023 / Accepted: 21 June 2023
© The Author(s) 2023

Abstract

What algorithm to choose for customer segmentation? Should you use one algorithm or many? How many customer segments should you create? How to evaluate the results? In this research, we carry out a systematic literature review to address such central questions in customer segmentation research and practice. The results from extracting information from 172 relevant articles show that algorithmic customer segmentation is the predominant approach for customer segmentation. We found researchers employing 46 different algorithms and 14 different evaluation metrics. For the algorithms, K-means clustering is the most employed. For the metrics, separation-focused metrics are slightly more prevalent than statistics-focused metrics. However, extant studies rarely use domain experts in evaluating the outcomes. Out of the 169 studies that provided details about hyperparameters, more than four out of five used segment size as their only hyperparameter. Typically, studies generate four segments, although the maximum number rarely exceeds twenty, and in most cases, is less than ten. Based on these findings, we propose seven key goals and three practical implications to enhance customer segmentation research and application.

Keywords Customer segmentation · Machine learning · AI · Algorithms

Introduction

Business success depends on understanding customers and their needs. A key method to achieve this is *customer segmentation*, i.e., dividing individual customers into groups based on their similarities and differences (Cooil et al. 2008). As postulated by Punj and Stewart (1983: 135), “All segmentation research, regardless of the method used, is

designed to identify groups of entities (people, markets, organizations) that share certain common characteristics (attitudes, purchase propensities, media habits, etc.)”. Customer segmentation, in particular, allows businesses to create targeted marketing strategies, improve customer experience, and ultimately, increase revenue (Hosseini and Shabani 2015; Simões and Nogueira 2022; Spoor 2022). As the potential benefits are plentiful, most firms perform customer segmentation across industries, irrespective of their size (Zhou et al. 2021). With the rise of artificial intelligence (AI) and machine learning (ML) technologies (Mustak et al. 2021), firms are increasingly turning to more advanced AI and ML algorithms (we refer to these as AI/ML algorithms henceforth; for the reader interested in the conceptual distinction of these two terms, we recommend reading Kühl et al. 2022) to perform customer segmentation. To this end, customer segmentation undoubtedly represents a cornerstone application of AI for business purposes, as it represents an unsupervised learning problem where AI/ML algorithms are known to be applicable (Joung and Kim 2023; Ranjan and Srivastava 2022). This relationship is illustrated in Fig. 1.

However, at the same time, firms struggle to understand these novel AI/ML methods and their implementation in

✉ Joni Salminen
jonisalm@uwasa.fi
Mekhail Mustak
mekhail.mustak@hanken.fi
Muhammad Sufyan
muhammad.m.sufyan@jyu.fi
Bernard J. Jansen
bjansen@hbku.edu.qa

¹ University of Vaasa, Vaasa, Finland

² Hanken School of Economics, Helsinki, Finland

³ University of Jyväskylä, Jyväskylä, Finland

⁴ Qatar Computing Research Institute, Hamad Bin Khalifa University, Doha, Qatar





Fig. 1 The hierarchy of key concepts in this study. AI is a general concept referring to (pseudo-)intelligent algorithms performing tasks that require intelligence. Machine learning is an application of AI. Unsupervised learning is a type of machine learning, of which clustering is the most common approach. Clustering, when applied to a customer dataset, then becomes the customer segmentation task

everyday customer segmentation—for example, what algorithm to choose for their given data? Should they use one algorithm or many? How many customer segments should be created? How to evaluate the results? These motivational questions reveal a *practical* research gap for customer segmentation research. Simultaneously, extant academic research lacks a synthesis of how customer segmentation is carried out in research studies, in terms of methods, parameters, evaluation, and so on—in other words, there is a *theoretical* research gap. Taken together, these two gaps hinder the development of the body of knowledge around the practice and theory of customer segmentation, especially in the light of novel AI technologies (the most cited article assessing the use of clustering in marketing, by Punj and Stewart, is from 1983, so there is a need for an updated study). It becomes harder for firms and researchers to develop better approaches, evaluate them, and implement them if we do not know adequately well *how* customer segmentation has been done in the past.

We address this knowledge gap by exploring the various customer segmentation methods used in business practices and then delve into the various AI/ML algorithms used for customer segmentation. Therefore, this research emphasizes algorithmic approaches to study what we term algorithmic customer segmentation (ACS). The central premise of our treatise is that AI/ML methods have become increasingly commonplace in marketing, specifically in customer segmentation. However, the extant academic literature lacks a holistic view of *how customer ACS is done*. Offering such a view is valuable for scholars (in terms of identifying patterns and gaps) and practitioners, as lessons from previous work are likely to offer a useful starting point for robust customer segmentation while helping to identify novel angles for future work in this domain.

Overall, this study aims to increase knowledge of customer segmentation research, and, more explicitly, the role of algorithms and AI in customer segmentation, toward understanding the theory and practice of ACS. In addition, based on our analysis of the extant literature, we offer a detailed and comprehensive agenda for future academic inquiry. To this end, we formulate specific research questions (RQs) that are addressed by a systematic literature review (SLR):

- **RQ1:** *What algorithms are typically used for customer segmentation?* To understand ACS, we must familiarize ourselves with the different algorithms applied for the customer segmentation task in scholarly literature.
- **RQ2:** *Does customer segmentation typically use one or many algorithms? When many, what algorithms emerge in combination?* The interaction among different algorithms poses an impactful question, as the field of customer segmentation migrates toward interactive systems that enable stakeholders to interact with the customer segments directly—thus, one algorithm might not be sufficient for more advanced systems (Jung et al. 2018).
- **RQ3:** *How many customer segments are typically created?* There could be two, three, or ten segments—but it is unclear what the optimal number of customer segments should be. Furthermore, this number will likely vary by the dataset in question (Salminen et al. 2022). The current body of knowledge lacks systematic insights into how many customer segments are created in research, so we investigate this RQ.
- **RQ4:** *How is customer segmentation typically evaluated?* ‘Evaluation’ means determining the quality of an algorithmic process; so, how well the customer segmentation worked (Thirumuruganathan et al. 2023). This typically measures the segmentation algorithm’s ability to clearly distinguish one group of customers from another group. As evaluation is an essential part of the customer segmentation process, we investigate this RQ.
- **RQ5:** *What hyperparameters are used in algorithmic customer segmentation?* A ‘hyperparameter’ refers to a parameter external to the model (Jansen et al. 2021). The human (or an automated script) selects the “optimal” hyperparameters based on manual viewing or some technical measure. The most central hyperparameter for customer segmentation is the number of segments created (i.e., the segment size). However, there could be other relevant hyperparameters, which is why we are investigating this RQ.
- **RQ6:** *How frequently are subject matter experts used for evaluating the customer segments?* The quality of customer segmentation can be ascertained using both automatic and manual means. ‘Automatic’ implies that the decision-making relies on technical metrics. However, at least of equal importance is exposing the customer segments to human decision-makers, e.g., managers, software developers, designers, and others that rely on customer segmentation as inputs in their decision-making. To this end, it is worthwhile to investigate how these stakeholders are involved in the customer segmentation process.



Methodology

Overview

We adopt the SLR methodology to discover and analyze pertinent existing studies. The SLR process provides a precise and reliable appraisal of the topic under examination, acknowledges existing flaws, and is less biased than typical judgment-based evaluations done by professionals in a specific field. In our approach, we adhered to the review procedure described by Kitchenham et al. (2009) and conducted the review in three sequential stages: (a) literature search, (b) assessment of the evidence base, and (c) analysis and synthesis of the findings.

Literature search

In this study, we have included literature from various business domains, for instance, marketing and management. Further, as a high degree of relevant knowledge is available in the field of computer science, it is worthwhile to pursue knowledge developed there while concurrently examining its marketing implications (Mustak et al. 2021; Salminen et al. 2019). Predominantly, we focus on *customer* segmentation, not *market* segmentation. Although these two concepts appear similar, they are not: Market segmentation is the process of dividing a market into different subgroups of consumers with similar needs, wants, behavior, or other characteristics. In turn, customer segmentation is the process of dividing customers into groups based on their qualities, attributes, and behaviors. So, the former deals with the *overall* market and the latter with the *specific, current* customer base.

To identify relevant literature, we used four prominent academic databases: Web of Science (WoS), Emerald Insight, ACM Digital Library, and ABI/INFORM Collection (ProQuest). WoS is the most comprehensive generic database, encompassing more than 12,000 high-impact journals and research articles from more than 3,300 publishers. The Emerald Insight and ABI/INFORM Collection (ProQuest) are also similar—as generic databases, they offer a collection of many relevant journals and scientific articles. The ACM Digital Library is a specialized database focusing on technical disciplines and thus helps uncover the articles focused on technical aspects of customer segmentation. Combined, these four databases offer balanced coverage of the existing literature on customer segmentation from multiple scientific disciplines.

We conducted detailed searches in each of the four databases. We did not want to pre-limit the searches with highly specific keywords and narrow terms, which may exclude crucial articles, as those articles may address the very same

topic but use different terms. Rather, to identify a wide range of publications that may shed light on customer segmentation, we used only the following keywords: “customer segmentation*”, “user segmentation*”, and “audience segmentation*” (“*” denotes plural forms) and identified all the associated articles. However, in the case of WoS, after the initial search, which resulted in 574 entries, we kept articles from the fields of business and computer science (as categorized by WoS) and excluded articles from other fields that we deemed outside the scope of this study. These fields included, for example, materials science, physics, chemistry, or integrative complementary medicine. The specific search details, along with the assessment of the evidence base, are presented in Table 1.

Assessment of the evidence base

We carried out the article screening in two stages. In the first stage, we focused on “hygiene factors” (see Table 2), such as removing duplicates, articles with missing information, those written in a language other than English, and articles published before the year 2000, as we considered the past 20 years or so to contain most relevant work for our research purpose. Figure 2 illustrates the general interest in segmentation among studies in or related to marketing.

After the first round of screening, we checked the remaining 204 articles based on their contents, i.e., assessing their relevance to our research purpose. We found that 32 (15.7%) articles were irrelevant to our research purpose (e.g., they were literature reviews or did not use empirical data to conduct customer segmentation). This reduction left us with 172 (84.3%) articles, 134 (77.9%) of which contained algorithm-based approaches to customer segmentation (i.e., representing ACS), while 38 (22.1%) articles applied non-algorithm-based customer segmentation (see Fig. 3). In our analysis, for RQ1-2 and RQ4-RQ5, we focus on the 134 articles that were algorithm-based. For RQ3 and RQ6, we focus on the full 172 articles that passed the pre-screening as these RQs do not require the articles to be algorithm-based.

Article coding

After the screening, we extracted information from the articles to address our RQs using a data extraction sheet. The information fields were designed to correspond with the RQs (see Table 3). The coding was carried out by one researcher, with another researcher verifying the quality of the coding outcomes by randomly investigating a sample of 20 articles. This inspection was carried out successfully and revealed that the data was coded following the guidelines in Table 3. The results (see Section “RQ2: does customer segmentation typically use one or many algorithms?”) were obtained by carefully reviewing the full-text articles.



Table 1 Details of literature searches for this study

| Database | Keyword searches | Search query | Number of preliminary identified articles | Selection of specific fields | Number of identified articles |
|----------------------------------|---|---|---|---|-------------------------------|
| Web of Science | Searching the keywords in "title", and in either "abstracts" or "article keywords" | ((TI=("customer segmentation" OR "user segmentation" OR "audience segmentation")) AND (AB=("customer segmentation" OR "user segmentation" OR "audience segmentation")) OR AK=("customer segmentation" OR "user segmentation" OR "audience segmentation")) | 574 | Included (WoS Category): computer science; business economics; & operation research management science Excluded, for example (WoS category): materials science; physics; chemistry; integrative complementary medicine | 390 |
| Emerald Insight | Searching the keywords in "title", and in "abstracts." This database does not contain a function to search in "article keywords" | "customer segmentation" OR "user segmentation" OR "audience segmentation" | 15 | The database does not contain the function | 15 |
| ABI/INFORM Collection (ProQuest) | Searching the keywords in "title", and in "abstracts." This database does not contain a function to search in "article keywords" | "customer segmentation" OR "user segmentation" OR "audience segmentation" | 69 | The database does not contain the function | 69 |
| Total | | | | | 474 |

Results

RQ1: What algorithms are typically used for customer segmentation?

After a thorough overview of the ACS literature, *we identified 46 different algorithms used for customer segmentation*, whose usage frequencies are illustrated in Fig. 4.

As can be seen in Fig. 4, K-means clustering is the most frequently used algorithm, as it is used 27 times (20.1%) in our sample of reviewed literature. Other prominent algorithms include variants of K-means clustering with a frequency of 10 (7.5%), fuzzy algorithms, and latent class analysis models are all used 8 (6.0%) times, respectively. The recency-frequency-monetary gain (RFM) and its variants have been used 6 (4.5%) times, while Self-Organizing Maps (SOM) and Genetic Algorithms (GA) have been used three times (2.25%) each. Other algorithms, for example, the Louvain algorithm, Ward's algorithm, and hierarchical clustering algorithms have been used two to three times varying. Furthermore, some algorithms have only been used once, indicating that these approaches may not have been thoroughly explored in the context of customer segmentation. Examples include direct grouping iterative merge and consistency-based clustering algorithms, suggesting that there is scope for further nuanced research in these areas.

Overall, the results highlight the wide range of algorithms available for customer segmentation and the need for further exploration and comparison of these methods to determine the most effective approach for different business scenarios.

RQ2: Does customer segmentation typically use one or many algorithms?

After a thorough overview of the ACS literature, it appears that in most cases, researchers have utilized one algorithm for customer segmentation (i.e., in roughly 80% of them). However, there are also instances where multiple algorithms have been combined for more effective results. For example, K-means clustering, SOM, and RFM approaches have been applied in combination with other approaches (see Table 4).

In general, the employment of multiple algorithms in combination may aid in addressing the shortcomings specific to each algorithm and may also help in the creation of more robust and distinct customer segments. Furthermore, technical reasons hailing from the domain of applied ML can explain the use of multiple algorithms. It is expected that AI/ML research studies compare and evaluate multiple algorithms for one task; this involves experimenting with combinations of different algorithms and calculating the accuracy/performance among them. This is often done by conducting an ablation study (Symeonidis et al. 2018).

Table 2 Inclusion/exclusion steps applied in the first stage of screening, along with the number of articles remaining after each step

| Step | Measures taken | N | Remains |
|------|---|-----|---------|
| 0 | INITIATE: Start of screening | 479 | 479 |
| 1 | MISSING INFORMATION: Remove articles with missing information (publication year, source title) | 8 | 471 |
| 2 | PUBLICATION TYPE: Remove conference proceedings, book chapters, theses, and other non-journal articles | 232 | 239 |
| 3 | RECENCY: Remove articles published before 2000 | 5 | 234 |
| 4 | DUPLICATES: Remove duplicates | 26 | 208 |
| 5 | LANGUAGE: Remove non-English articles | 3 | 205 |
| 6 | DISCIPLINES: Remove articles from the geography discipline | 1 | 204 |

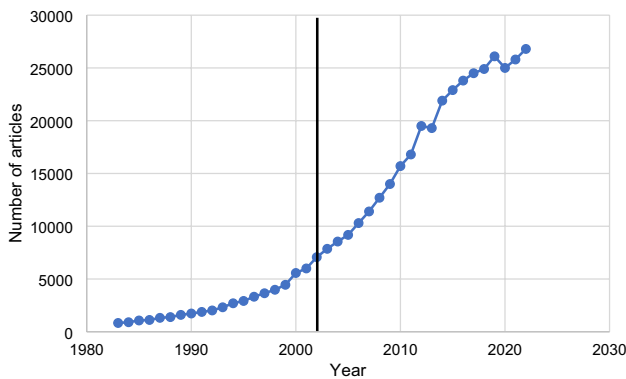


Fig. 2 Approximate search results when searching Google Scholar for [+ segmentation + marketing]. The results indicate a general increase in interest in segmentation among studies in or related to marketing. The vertical line (the year 2000) indicates our sample's cut-off year

An ablation study is a form of experimental design used to study the effect of removing a specific part or feature of a model on its overall performance, typically used in the field of ML and AI. It can involve removing individual components of a larger model or system, a subset of data

features, or hyperparameters to see how they affect the accuracy, cost, and other metrics of the ML model. Similarly, the researchers developing new customer segmentation algorithms are required to demonstrate the value of their approach; for this reason, algorithms are often compared against one another.

RQ3: How many customer segments are typically created?

Our analyses show that various numbers of customer segments have been suggested/created by researchers based on the application area and/or target market and objectives of the research (see Fig. 5). Most commonly, researchers have suggested/created *four* customer segments ($n = 32, 21.2\%$). On average, researchers created 5.7 segments ($SD = 3.9$). Interestingly, no study produced more than 20 customer segments (apart from the outlier study mentioned in Fig. 1 caption), while the lowest number of segments produced was one (in this particular study, the researchers applied a decision-rule algorithm to discover the most ideal customer type (Lee, J. H. & Park, 2005)). Interestingly, more than half ($n = 103, 68.2\%$) of the studies generated between two and

Fig. 3 Research process leading to article coding. Essentially, this took place in four stages: Screening, Relevance Assessment, Algorithm Assessment, and Data Extraction

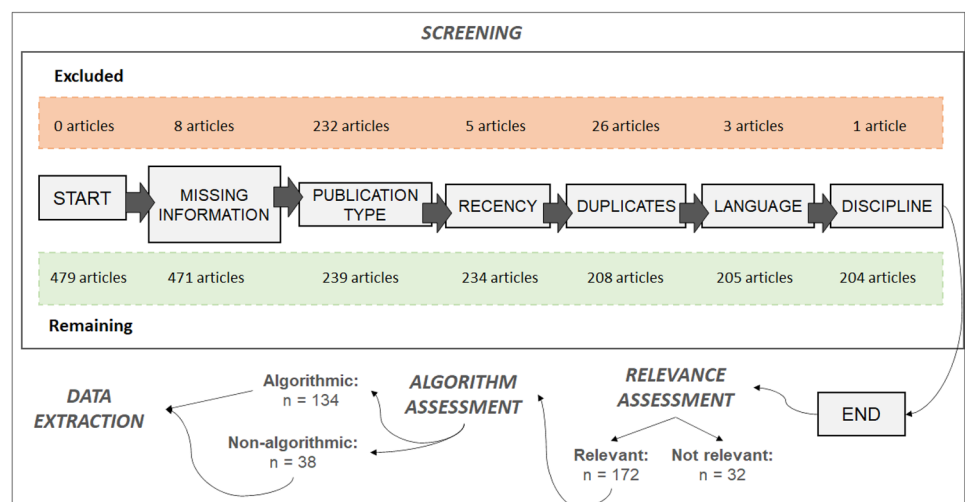
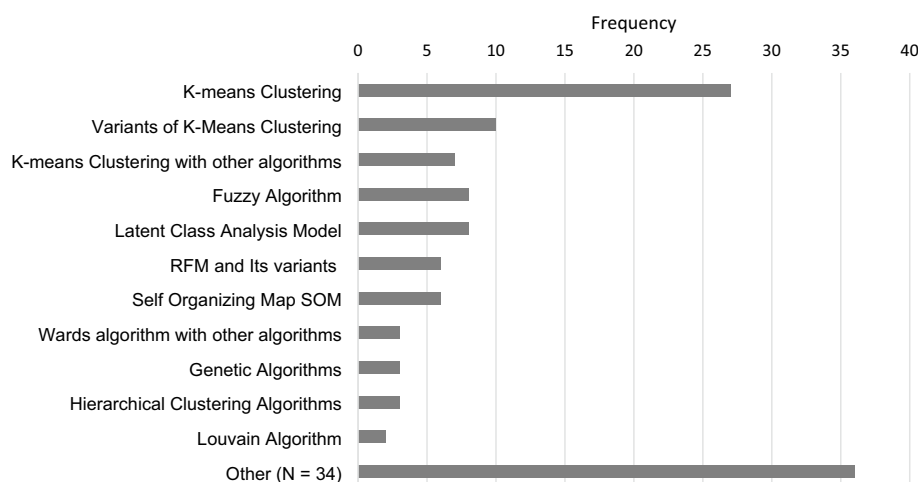


Table 3 Data extracted from the research articles

| ID | Information field | Possible values | RQ |
|----|---|-----------------|---------------------|
| 01 | ALGORITHM-BASED: Is the article algorithm-based? In other words, does it apply algorithms on a dataset about customers or people to create segments of the customers or people | Yes/No | (Article selection) |
| 02 | METHOD: Empirical (E), Conceptual (C), or Other (O) | E/C/O | (Article selection) |
| 03 | ALGORITHMS: Name of the algorithm or algorithms used for customer segmentation | (open) | RQ1, RQ2 |
| 04 | SEGMENT SIZE: Number of customer segments created | (integer) | RQ3 |
| 05 | EVALUATION: Approach or approaches used for evaluation | (open) | RQ4 |
| 06 | METRIC: Evaluation metric or method for the number of segments | (open) | RQ4 |
| 07 | HYPERPARAMETERS: Did the study involve other hyperparameters than the number of segments? (If “No”, what other hyperparameter(s) were used?) | Yes/No | RQ5 |
| 08 | EXPERTS: Were subject matter experts used for validating the segments? | Yes/No | RQ6 |

Fig. 4 Frequency chart for the algorithms used. The y-axis represents the name of the algorithms used for customer segmentation, and the x-axis shows the frequency, i.e., the number of times each algorithm has been used in the literature for customer segmentation. The ‘Other’ class contains all the algorithms that were used only once



five segments (see Fig. 5). Furthermore, 92.1% of the studies generated ten or fewer segments.

Typically, the number of segments is not determined beforehand but in a *data-driven way* (Hiziroglu 2013; Hong and Kim 2012), which means determining the segment size based on quantitative evaluation metrics (i.e., the number of segments is such that best fits the data based on an evaluation metric when experimentally varying the segment size). For example, the researchers may attempt multiple numbers for k where k indicates the number of segments, and then visualize at what number of k the obtained information noticeably decreases—the so-called elbow method where elbow indicates this decrease (Syakur et al. 2018). So, the number of clusters is neither random nor determined a priori, but all the reviewed articles used some quantitative metrics/criteria to validate or choose the number of segments.

RQ4: How is customer segmentation typically evaluated?

In ACS, the evaluation of the customer segments is crucial for the effectiveness of the segmentation process. As ACS tends to be an unsupervised ML task (where there is no single “correct” value for a segment, but instead, the algorithm aims to organize the data into groups, i.e., segments), accuracy is used more rarely in ACS than in other types of ML tasks, including supervised learning tasks such as prediction. Hence, there are several measures and criteria that researchers have used to evaluate the quality of customer segments. Overall, we identified 14 different metrics for evaluating customer segmentation outputs, of which six (42.9%) focused on statistical indicators and eight (57.1%) focused on distances and/or similarity calculation (see Table 5). We discuss the metrics in alphabetical order.

Accuracy (ACC) is used to evaluate the quality of segments, i.e., how well “unseen” or new members (customers) the segmentation algorithm can allocate to the correct segment (Wu, S. et al., 2021). ACC is typically calculated by dividing the number of correct values by the total number



Table 4 Examples of combining customer segmentation algorithms

| Main algorithms | Auxiliary algorithms |
|---------------------------------------|---|
| K-means clustering | RFM and LRFPM (Zhu et al. 2015; Nisan et al. 2022) K-medoids (Nie et al. 2020) Semi-Supervised Affinity Propagation (Deng and Gao 2020) Genetic Algorithm, Artificial Bee Colony, Particle Swarm Optimization (Kuo et al. 2016) Ward’s method (Güçdemir and Selim 2015; Repschlaeger et al. 2013) Expectation Maximization EM (Tsai et al. 2015) |
| Self-Organizing Maps (SOM) | Fuzzy C-Mean (Hiziroglu 2013) Ward’s method (Yao et al. 2014) K-means clustering (Hong and Kim 2012; Seret et al. 2015) Decision Trees (Lee and Park 2005) |
| Recency Frequency Monetary gain (RFM) | Customer Lifetime Value (CLV) modeling (Wong and Wei 2018) Logistic regression, Decision Tree (Coussement et al. 2014) |

of predicted values. The use of ACC requires that there are ‘ground truth’ values or labels against which the predictive ability of an algorithm is compared.

The *Adjusted Rand Index* (ARI) measures the similarity between two segmentations (Xu, X. et al., 2007). The ARI considers the random chance that objects between the segmentations might be similar. It compares object pairs in two sets of segments and computes the difference between the observed agreement and the expected agreement under random labeling. The value of -1 indicates perfect dissimilarity, and + 1 indicates perfect similarity. So, the higher the ARI, the more similar the two customer segmentations are.

The *Analysis of Variance* (ANOVA) test is a statistical method used to test the difference in means between two or more groups. In customer segmentation, ANOVA can test for significant numerical differences in customer characteristics between the segments (Ballestar et al. 2018; Kashwan and Velu 2013).

The *Average Clustering Error* (ACE) evaluates the average distance between data points within each segment (Manjunath and Kashef 2021). The ACE is calculated by taking the average of the sum of the distances from each point in the dataset to its closest segment center. It indicates how well the segmentation algorithm has been able to group

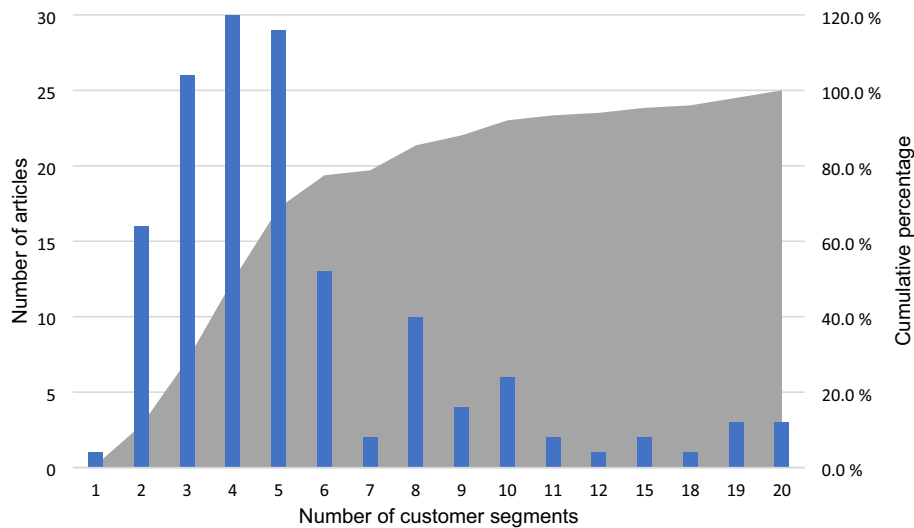


Fig. 5 Number of customer segments created in research articles. The numbers are based on 151 articles (87.7% of the total 172) that expressed the number of segments created. In the case the researchers presented multiple segmentations with different numbers (e.g., 7, 7, 8), as was sometimes the case when experimenting with multiple algorithms or datasets, we have taken the average of the reported

numbers and rounded it either up or down based on standard rounding rules (e.g., 7, 7, 8 would yield the average of 7.3 which rounds to the segment size of 7). Twenty-one articles (12.2%) did not report the segment size. Also, we omitted one outlier article from the analysis, as the researchers created 1209 and 8984 segments from two datasets, respectively (Böttcher et al. 2009)



points with similar characteristics. If the ACE is low, this indicates an effective segmentation algorithm.

The *Bayesian Information Criterion* (BIC) is a measure used for model selection and comparison, including in segmentation (Bhade et al. 2018). To calculate the BIC, the log-likelihood of the segmentation model is adjusted by penalizing complex models. The BIC score is obtained by subtracting the penalty term from the maximum log-likelihood, where a lower BIC value indicates a better-fitting and more parsimonious model.

The *Calinski-Harabasz Index* (CHI) evaluates the separation between the segments and the compactness within each segment (Kandeil et al. 2014). The CHI is calculated by comparing the within-segment dispersion to the between-segment dispersion. A higher value indicates that the segments are well-defined and that the dataset has been well-split into distinct segments.

The *Davies Bouldin Index* (DBI) assesses the similarity between the segments based on the distance between their centroids (i.e., midpoints). The lower the DBI score, the better the segmentation result, indicating that the segments are more compact and less scattered (Aryuni et al. 2018). The DBI penalizes algorithms that produce segments with a wide variety of sizes and shapes, with larger diameters relative to the separation between the segments.

The *Dunn Index* (DI) evaluates the separation between the segments and the compactness within each segment (Khajvand and Tarokh 2011). The calculation of the DI is done by taking the ratio of the minimum intra-segment distance to the maximum inter-segment distance. A higher value indicates that the segments are better separated.

The *Fukuyama and Sugeno* method (FS) is an approach to evaluate segmentation results based on fuzzy sets and fuzzy logic. The FS involves assigning membership values to data points indicating their degree of belongingness to each segment. The process incorporates expert knowledge through

the formulation of fuzzy rules, which guide the decision-making process. These membership values are then used to calculate a validity index, which measures the quality of the segmentation by considering their compactness and separation (Nemati et al. 2018).

The *Mann-Whitney rank test* (MW) is a non-parametric statistical test used to compare the difference in the median between two groups. In customer segmentation, the MW can be used to test whether there are significant differences in customer characteristics between the different segments (Jiang and Tuzhilin 2009).

The *Silhouette Index* (SI) is a metric used to measure how well-defined a segment is and how strongly a data point is assigned to its associated segment (Dzulhaq et al. 2019). The SI ranges from -1 to 1 , with a score of 1 indicating that the data point is perfectly matched to its own segment, a score of -1 indicating that the data point is more closely associated with another segment, and a score close to 0 indicating that the data point does not have a clear segment assignment. The SI is calculated by taking the average of the difference between the data point's own segment similarity and the lowest segment similarity with the other segments.

The *Total Clustering Effectiveness* (TCE) is a metric used by Lu and Wu (2009). The TCE combines an inter-cluster correlation indicator and an inner density indicator. The numerator represents the sum of densities for two segments, where a higher value indicates better performance. The denominator represents the correlation coefficient between the two segments, with a smaller value indicating better results. Incorporating both values, a higher TCE indicates better results.

The *Validity Index* (VI) is a measure that evaluates the quality of the segmentation result based on specific criteria (Pramono et al. 2019). These criteria can include factors such as intra-segment cohesion (compactness), inter-segment separation, or the overall structure of the segments. VIs provide a numerical score or value that indicates the goodness of fit of the segmentation solution, with higher values suggesting better quality or validity. Different VIs employ distinct formulas or methodologies to capture different aspects of segmentation performance.

Finally, the *Xie-Beni Index* (XBI) is a validity index used to evaluate the quality of segmentation results (Munusamy and Murugesan 2020). The XBI quantifies the trade-off between segment compactness and separation by calculating the ratio of the sum of squared distances between data points and segment centroids to the product of the segment compactness and the number of data points. A lower value indicates better segmentation with tighter and well-separated segments.

A few conclusions can be made. First, the metrics can be divided into statistics- and separation-focused metrics, with the latter being slightly more typical. Second,

Table 5 Segmentation evaluation metrics

| <u>Statistics-focused</u> (compare two or more groups using statistical testing or probabilities) | <u>Separation-focused</u> (calculate distances or similarities between items and centroids) |
|--|--|
| ACC | ACE |
| ANOVA | CHI |
| ARI | DBI |
| BIC | DI |
| MW | FS |
| TCE ¹ | SI |
| | VI ² |
| | XBI |

The acronyms can be found in the body text following the table

¹Deals with correlation

²Typically relies on separation measurement



statistics-focused metrics emphasize segment-to-segment differences, while separation-focused metrics emphasize low intra-segment distance (i.e., compactness) and high inter-segment distance (i.e., separation). Third, customer segmentation evaluation is centered on using metrics derived from clustering practices, rather than using metrics especially tailored to customer segmentation, business outcomes, or ecological validity.

RQ5: What hyperparameters are used for algorithmic customer segmentation?

In ML, a hyperparameter refers to a configuration setting (i.e., a numeric value) that is external to the model itself and is typically set before the learning process begins (Jansen et al. 2021). Unlike model parameters, which are learned from the training data, hyperparameters are predefined choices (range of values when experimenting with multiple hyperparameter values) that affect the model's performance and behavior. These parameters can include things like the learning rate, the number of hidden layers in a neural network, or the regularization parameter. Selecting appropriate hyperparameter values is crucial for achieving optimal model performance, and it often involves experimentation, trial and error, or using techniques like grid search or Bayesian optimization (Jansen et al. 2021).

In our review, out of the 169 studies that offered information about hyperparameters, more than four out of every five articles ($n = 138$, 81.7%) applied only segment size as a hyperparameter, while less than one out of five ($n = 31$, 18.3%) applied additional hyperparameters. In ACS, the algorithms may combine technical and business hyperparameters, with the technical parameters stemming from the inputs required by the algorithm (most commonly, the number of segments to create, i.e., the segment size and the distance measure—how the distance between the segments is calculated) to perform its computation and the latter arising from the particular business scenario the segmentation aims to address. For example, Munusamy and Murugesan (2020) performed customer segmentation based on the Fuzzy C-mean clustering algorithm and defined U matrix parameters to make their data compatible with the Fuzzy C-mean algorithm.

In contrast, Peker et al. (2017) formulated a hybrid approach for customer behavior prediction and used many AI/ML algorithms, including Neural Networks, SVM, Decision Tree, and Radial Basis Functions that make use of different hyperparameters, for instance, cost, gamma, number of hidden layers, weights, number of leaf nodes and number of trees, etc. In another contribution, the authors formulated a hybrid big data model for analyzing customer patterns in an integrated supply chain network (Wang et al. 2020). They applied Linked Based Bloom Filters (LBF)

that served as parameter functions directly linked with customer segmentation. Liu et al. (2009) formulated a hybrid approach for a product recommendation that directly relates to customer segmentation. Their proposed approach used learning rate, grid structure, and distance normalization as hyperparameters.

Although these parameters are not directly linked with customer segmentation, they contribute to the overall segmentation process in terms of providing a method for ascertaining a *technically optimal* number (and structure) of the segments. In turn, the business parameters aim to provide more information about the domain-specific business context. These may include the likes of Length of customer involvement (L) and Periodicity (P) that were applied by Nemati et al. (2018). Zhu et al. (2015) applied Profitability (*prof*), Accuracy (*acc*), and lead time as hyperparameters that contribute to customer segmentation in demand fulfillment of customers in case of supply shortage. Wu and Liu (2020) incorporated group preferences and linguistics parameters into their Type 2 fuzzy customer segmentation models and concluded that these parameters greatly affect the customer segmentation task.

Overall, the hyperparameters applied by most articles stem from the standard/default hyperparameters used by these algorithms in any dimensionality reduction context, of which customer segmentation is a special case.

RQ6: How frequently are subject matter experts used for evaluating the customer segments?

In this study, we encountered only seven cases (4.1%) where subject matter experts were used to evaluate the quality of customer segmentation and provide expert opinions. In other words, it is a rare, perhaps too rare, practice to invite stakeholders and subject matter experts to validate the results of the customer segmentation process in academic research articles. Out of the rare examples that do exist, Nemati et al. (2018) formulated a customer lifetime value (CLV) approach for prioritizing marketing strategies in the telecom industry. The experts were first asked through a questionnaire to provide different parameters for the said tasks. Once the segmentation was done, the experts were again consulted to evaluate and validate the results. Safari et al. (2016) formulated an RFM-based CLV determination approach that performs customer segmentation based on the RFM values. To do so, subject matter experts were asked through a questionnaire, and once the segmentation was carried out, a total number of 16 experts expressed their opinions about the accuracy of the segments.

Similarly, Sun et al. (2021) introduced a heuristic approach to customer segmentation. The experimental results show that the customer segmentation output by their proposed method was consistent with the customer



segmentation result given by experts. Manidatta et al. (2021) introduced an integrated approach for customer segmentation and evaluated their approach through experiments. They collected responses from nine subject matter experts from the Indian retail industry regarding their perception of the relative importance of four CLV criteria and evaluated the weights of each criterion using fuzzy AHP. Transaction data for 18 months was analyzed to segment 1,600 customers into eight segments using the fuzzy c-means clustering analysis technique. The segmentation results of their proposed integrated method were further validated by the nine experts from the Indian retail industry.

In another study, Li et al. (2011) formulated an agglomerative clustering-based approach for customer segmentation, and as a result of their proposed approach, the customer was segmented into four distinct groups/segments. Subject experts validated and evaluated the experimental results for customer segmentation by their approach (Li et al. 2011). Lee and Cho (2021) formulated a customer segmentation approach based on the Leuven algorithm. To verify and validate the segmentation results of their proposed approach, they consulted a subject matter expert, and the algorithm determined the modularity for ten segments, which was the same number of segments identified by the domain expert. Warner (2019) conducted a study on audience segmentation using a survey approach. Before conducting the study, a seven-member expert panel was asked to review the instrument and provide their expert opinion on the number of segments created. The domain expert team validated the audience segmentation results.

Thus, our analysis indicates that most often, the results of customer segmentation (including both algorithmic and non-algorithmic customer segmentation research) are not validated using external feedback, but the authors tend to rely on technical evaluation metrics to justify the quality of their work. This practice likely stems from the ML research tradition, in which metrics such as accuracy, precision, recall, etc., are used to evaluate the performance of an algorithm (Bell 2014; Kühl et al. 2022), rather than “subjective” human feedback. However, when they are used, most typically, multiple subject matter experts participate in evaluating the created segments.

Discussion

Study highlights

Here, we discuss the highlights of our findings.

We identified 46 different algorithms applied by researchers for customer segmentation. This finding highlights not only the methodological (algorithmic) plurality within customer segmentation studies but also the influence of the

current AI and ML technologies in this domain, as these algorithms overwhelmingly stem from the ML research tradition.

Most of the reviewed studies used one segmentation algorithm, making multi-algorithm customer segmentation a minority endeavor. The promise in multi-algorithms customer segmentation is that, in theory, it is better able to handle the plurality of segmentation criteria and be more responsive to organizational requirements for changing the segmentation parameters with updating business requirements.

In ACS, the number of segments is not pre-assigned, but it is inductively determined based on quantitative evaluation metrics. *On average, researchers create 5.7 customer segments per study (SD = 3.9, Mode = 4, Median = 5). So, even though creating four segments is the most common, the number of segments created varies substantially across the reviewed studies. No analyzed study created more than 20 segments (Min = 1, Max = 20).*

Few studies explicitly define the concept of customer segmentation. Instead, the concept is often treated implicitly, as “everyone knows what it is”. This conceptual vagueness can hinder the development of scientific advances in customer segmentation, as ‘customer segmentation’ might not be a similar task to other clustering tasks [the (dis)similarity of customer segmentation to other clustering tasks remains unaddressed in the literature].

Research outlines numerous ‘theoretical’ use cases and benefits for customer segmentation. These benefits, ranging from pricing to targeting and personalization of offerings and messaging, emphasize the central role of customer segmentation as a key business application of AI and ML technologies. Concurrently, few studies show empirical evidence of these benefits in organizational use cases or systems (i.e., ecological validity).

There is no one set of customer segmentation criteria, but the studies vastly vary in terms of the segmentation criteria applied. In fact, no two studies may have a single criterion (i.e., customer attribute) in common. This plurality of criteria partially explains the algorithmic or methodological plurality, as different criteria represent different data types that require distinct preprocessing and analysis approaches to apply the algorithms. Therefore, it is unlikely that we would end up in a situation where only one (or even a few) algorithms would cover all use cases for customer segmentation.

Researchers developing novel customer segmentation algorithms tend to see customer segmentation as a computational task. This viewpoint is visible in how algorithms are used, studies are structured, and outputs are evaluated. Machine learning studies follow a particular paradigm of benchmark comparison, which may explain why a large portion of literature puts less emphasis on conceptual and



theoretical aspects of customer segmentation and instead focuses on it as a technical problem or ‘task’.

We identified 14 unique evaluation metrics for the quality of customer segmentation, all technical. The many metrics are the consequence of multiple algorithms: because the statistical and mathematical properties of different algorithms vary, one metric cannot be applied to evaluate the (internal) success of the modeling task. However, it would be valuable to have more centralized evaluation metrics for customer segmentation success; as the internal evaluation is affected by computational specificities, perhaps researchers could shift their focus on external (ecological) evaluation metrics, focusing on the business outcomes and the customer dynamics of applying customer segmentation rather than the segments’ creation process.

Business orientation separates customer segmentation from other clustering tasks. In addition to ML-dependent technical hyperparameters (e.g., number of segments, distance function), researchers utilize business-specific hyperparameters (e.g., length of customer involvement, profitability) for customer segmentation inputs. However, most studies reviewed (82%) only applied segment size as the hyperparameter.

Considering the historical development of the use of clustering in marketing, the post-2000 sample we analyzed shows some progress compared to the previous analysis made by Punj and Stewart in 1983. Specifically, the researchers indicated then that there is a “failure of numerous authors in the marketing literature to specify what clustering method is being used.” (p. 134). In our more recent sample, this condition does not take place as the authors are more explicit on the precise method being used—so, there has been progress in the reporting of clustering details in marketing work.

However, our results confirm the fundamental challenges of clustering (segmentation) as stated by Punj and Stewart (1983): “choice of an appropriate metric, selection of variables, cross-validation, and external validation” (p. 134). As then, these challenges remain topical and fundamentally unresolved. There are many metrics to choose from. The selection of segmentation variables is arbitrary. Cross-validation and external validation (ecological validity) are difficult to execute and thus often omitted. So, the fundamental nature of segmentation has not changed with the novel AI technologies, at least yet.

Interestingly, there can be seen as a continuation of knowledge. That is, the analysis by Punj and Stewart from 1983 (see their Table 4 on pp. 141–142) indicates K-Means as one of the most popular algorithms. Forty years later, this algorithm still maintains its position as the leading segmentation algorithm. We can interpret this finding as either proof of its superiority in this problem, or as traditionalism. However, either way, the conclusion remains that the novel

AI-based approaches have not been able to replace the “old AI” approaches, at least when it comes to K-Means.

In addition to the above highlights, in the following subsection, we provide a taxonomy of algorithms for customer segmentation.

Central goals and directions for future research

From our review, there are multiple avenues for future research to advance customer segmentation research. In the following, we outline seven key goals (KG) for future work:

- *KG01: Providing taxonomies of algorithms and metrics.* There is a need for conceptual frameworks, classifications, and taxonomies that help address the undeniable plurality of algorithms in the domain of customer segmentation research, which includes at least (a) algorithm selection plurality, (b) segmentation criteria plurality, (c) hyperparameter plurality, and (d) evaluation metric plurality. Our taxonomy of algorithms for customer segmentation provides a starting point and an example of outputs that can help address this gap. We invite other researchers to provide conceptual work (not only empirical!) that systemically categorizes the extant work on customer segmentation. Additionally, consensus on some fundamental concepts is much needed—for instance, how can ‘high-quality’ customer segmentation be distinguished from ‘low-quality’ customer segmentation? Should we focus on the quality of the process, the quality of the evaluation metrics, or the quality of the actual customer segments? Propositions (or even discussions) concerning these matters are direly needed.
- *KG02: Providing empirical evidence on customer segmentation outcomes.* Based on our literature review, we observed that few studies provide an empirical analysis of the actual application of ACS in organizations. Algorithmic studies tend to stop at the stage of creating the customer segments; their application in companies is not explored. On the one hand, this casts shadows on whether the potential benefits of customer segmentation mentioned in the studies are rather hypothetical, or whether they can be backed up with empirical evidence. Therefore, we encourage researchers to shift their focus from creating segments to applying them in firms and other organizations. This not only represents an exciting research gap but efforts in this regard can help further enhance and incentivize research projects on customer segmentation, as they would be more strongly linked to key performance metrics that firms and other organizations value. To this end, case studies, field studies, A/B



tests, experiments, and longitudinal studies would be welcome to address this vast knowledge gap.

- *KG03: Integrating algorithms into customer segmentation systems.* One central direction for segmentation research is developing more comprehensive pipelines that can handle multiple different data types (i.e., customer segmentation criteria of different types) and changing business requirements. A key direction in this regard would be merging customer segmentation research with *intelligent systems* research (i.e., research that focuses on developing systems that can think, reason, and make decisions independently, without human intervention, interacting with the environment and making decisions to optimize quantitative outcomes (Bauer and Dey 2016)) to generate and empirically investigate more comprehensive customer segmentation systems that stakeholders can interact with, not merely isolated attempts of testing how ‘Algorithm X’ fares with the customer segmentation task. Again, this corresponds to our postulation that, based on our review, customer segmentation research would benefit from a higher degree of ambition and scope, as the current body of work focuses on developing and testing algorithms instead of systems.
- *KG04: Proposing a standardized framework for evaluation.* Due to methodological plurality, there is a lack of consensus on what constitutes ‘quality’ in customer segmentation. However, the consensus from prior research indicates that quality is perceived as models’ internal consistency and evaluated using technical performance metrics focused on this internal consistency. ACS has inherited its evaluation metrics from the ML research tradition, essentially adopting the metrics used in other clustering and data dimensionality tasks to customer segmentation as well. While we do not deny the merits of these technical metrics—they provide useful information about the model’s fit with the customer dataset—we call for further extensions and contributions to broaden the hierarchy of evaluating customer segmentation outputs. In addition to technical metrics, we ought to consider other metrics as well, such as stakeholder perception (“Are these segments useful? How useful?”) and organizational outcomes (e.g., the “ROI of customer segmentation”, i.e., how much does the implementation of customer segmentation improve the profit of the company?). A more holistic and nuanced, hierarchical way of measuring the quality of customer segmentation would further develop the field in academic and practical circles.
- *KG05: Exploring organizational challenges of customer segmentation.* Challenges in customer segmentation arise, on the one hand, from organizational realities such as culture, capabilities, and individual experience and, on the other hand, from technical rationale such as data availability, selection of algorithms, and validating the quality of customer segmentation. The organizational aspect of applying customer segmentation in decision-making is a clear and present research gap. Again, qualitative studies can help generate rich insights into how customer segmentation algorithms transit into organizational adoption. So, we need a more collaborative approach to customer segmentation research. This means not giving up on developing better algorithms, but in addition to that, engaging with social science researchers in a pursuit to discover the impact of applying these algorithms in the real world. Such extension presents an exciting new field of study.
- *KG06: Providing more critical analyses.* Researchers typically do not discuss the fundamental limitations of using AI and ML in customer segmentation, such as the fact that clustering algorithms were not originally developed for customer segmentation. Questions such as “What are the downsides of using AI for customer segmentation?” are not asked. Yet, they *should* be asked, as critical analyses can reveal insights into transformative improvements in this field. As a starting point for such analyses, we offer some ideas: first, the use of algorithms may distance the stakeholders from the data, especially if the creation of the segments is not a participatory process but outsourced to a group of analysts/data scientists who will not be part of the eventual use of the segments. So, there can be a problem of detachment and silos. Second, algorithms may cause segmentation to become overly rigid and resistant to change, so the same criteria used to group customers may not be relevant after a certain amount of time. Third, customers may resent being “put into a box” and may not appreciate feeling like just another anonymous statistic; thus, customer perceptions regarding customer segmentation could be explored. These and other negative aspects of ACS should be studied much more.
- *KG07: Making the role of humans explicit.* Currently, humans play a central role in ACS; they select the algorithms and evaluation metrics, program the experiments, interpret the results, and choose final hyperparameter combinations. They also evaluate the results and make the final judgment of whether an algorithm did a good job at the segmentation. Finally, humans apply the segments in practice and make decisions based on them. Yet, we speak of ‘algorithmic customer segmentation’ and seem to delegate a lot of responsibility for the segmentation process to algorithms, often obfuscating the role of humans. As the field of customer segmentation is increasingly reliant on AI to do this work, a vital question is, what can AI learn from non-AI-reliant customer segmentation? In other words, there is a need to better understand human factors in the segmentation process



and how these can support or bias the process when co-existing with algorithms.

Practical implications

There are three main practical implications (PI) for researchers and practitioners:

- *PI01: Segmenting Beyond K-means:* Although nearly fifty algorithms were identified, k-means or a derivative is by far the most popular (27.6%) algorithmic approach for customer segmentation. This situation presents an opportunity and a need for exploring and comparing these algorithmic approaches to determine the most effective approach for different business scenarios—earlier research has partially done this (Punj and Stewart 1983), but not in a completely systematic way and not considering business scenarios. To do this comparing, it would entail both algorithmic investigation and business context research to identify characteristics for different business scenarios and how these relate to the selection of the best scenarios to address such analysis.
- *PI02: Sensemaking of the Segments:* In addition to investigating approach algorithms given the attributes of business scenarios, there is a general lack of evaluation of the resulting segment, with most of the evaluations done using technical measures related to the chosen algorithm. However, given the algorithmic plurality, this results in a plurality of evaluation techniques as well. This calls for standard evaluation criteria for customer segments extending across the various algorithms. Finally, there is a critical and unmet need for evaluating segmentation beyond the algorithmic nuances, using broader criteria of accuracy, fairness, diversity, or coverage. For these, investigations of segmentation hyperparameters would most likely be needed.
- *PI03: Finding the “X” in [Segmentation + X]:* The richness of segmentation criteria implies a nearly infinite number of ways of dividing a customer population into segments and these segments into sub-segments. Also, the segmentation criteria can be modified by adding or removing specific criteria at any time. Each of these segmentations offers one view of the segmentation that most likely has many possible views. So, which one is correct? This question is probably impossible without an ‘X’, i.e., some criteria external to the data for which to evaluate the segmentation. For this, the segmentation results need to be placed within the given business scenario, as discussed above, and the segments validated using external feedback, organizational key performance indicators, or achievement of business scenario goals.

Overall, customer segmentation, especially when applying AI and ML algorithms, is a socio-technical problem. In other words, both ‘technical’ (algorithm choice, data availability) and ‘social’ aspects (culture, goals) affect the success of customer segmentation projects. Therefore, mere technical solutions or more sophisticated algorithms are not adequate for ensuring successful customer segmentation projects. Thus, it is vital to understand customer segmentation projects as long-term processes that require stakeholder buy-in and effective implementation plans (i.e., the segmentation process does not end with creating the segments but only begins).

The current body of ACS literature focuses on the technical application of algorithms but largely omits the role of humans in this process, whether the role deals with various aspects of using the AI/ML algorithms (i.e., using judgment for the hyperparameter selection), evaluating the results, and eventually applying the results of ACS. Given the predominant focus on technical metrics and algorithms, there is a need to go beyond these aspects into the realm of inspecting the technology’s impact on actual organizations. According to our understanding, this can best be achieved by cross-disciplinary collaboration with social scientists, marketers, and other stakeholders who understand the qualitative side of customer segmentation and have access to organizational performance data. So, while this step of expanding ACS research into the realm of application is likely to involve a certain exit from the ML paradigm’s “comfort zone”, it is a necessary step to establish true scientific progress in this domain. Fernández-Delgado et al. (2014) famously asked in the context of classification, “Do we need hundreds of classifiers to solve real world classification problems?” (cited 3594 times at the time of writing this); we can paraphrase this question: *Do we really need 46 different customer segmentation algorithms?* Given the discrepancy between the large number of algorithms and the scarce number of articles applying subject matter experts for the evaluation of successful customer segmentation, the answer might be negative.

Study limitations

We did not include articles published before the year 2000 in our sample. We did not include keywords dealing with market segmentation or consumer segmentation (e.g., Kamakura and Russell 1989)—conceptually, these are different goals, as ‘market’ includes non-customers as well. However, it could be interesting to compare the methods and variables used in these different segmentation tasks. We leave this for future work.



Conclusion

Customer segmentation has been a major focus in academic literature for many years and continues to be one. It is also of high value to marketers in industry. However, due to a myriad of different approaches, the field suffers from the lack of clarity that we aimed to address with this study. We found that researchers have used 46 different algorithms for customer segmentation. Interestingly, around 80% of them utilized a single algorithm for this purpose. On average, they created about 5.7 customer segments, deciding the exact number inductively based on quantitative evaluation metrics. Surprisingly, few articles offer empirical evidence of the benefits of customer segmentation. Our results point the way for future research, as addressing the proposed key goals helps successfully develop customer segmentation algorithms, make sense of the customer segments, and evaluate the impact of the segmentation.

Funding Open Access funding provided by University of Vaasa (UVA). Funding was provided by Liikesivistysrahasto.

Declarations

Conflict of interest On behalf of all authors, the corresponding author states that there is no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Ballestar, M.T., P. Grau-Carles, and J. Sainz. 2018. Customer Segmentation in E-Commerce: Applications to the Cashback Business Model. *Journal of Business Research* 88: 407–414.
- Bauer, C., and A.K. Dey. 2016. Considering Context in the Design of Intelligent Systems: Current Practices and Suggestions for Improvement. *Journal of Systems and Software* 112: 26–47.
- Bell, J. 2014. *Machine Learning: Hands-on for Developers and Technical Professionals*, 1st ed. Hoboken: John Wiley & Sons.
- Böttcher, M., M. Spott, D. Nauck, and R. Kruse. 2009. Mining Changing Customer Segments in Dynamic Markets. *Expert Systems with Applications* 36 (1): 155–164.
- Cooil, B., L. Aksoy, and T.L. Keiningham. 2008. Approaches to Customer Segmentation. *Journal of Relationship Marketing* 6 (3–4): 9–39.
- Coussement, K., F.A.M. Van den Bossche, and K.W. De Bock. 2014. Data Accuracy's Impact on Segmentation Performance: Benchmarking RFM Analysis, Logistic Regression, and Decision Trees. *Journal of Business Research* 67 (1): 2751–2758.
- Deng, Y., and Q. Gao. 2020. RETRACTED ARTICLE: A Study on e-Commerce Customer Segmentation Management Based on Improved K-Means Algorithm. *Information Systems and e-Business Management* 18 (4): 497–510.
- Fernández-Delgado, M., E. Cernadas, S. Barro, and D. Amorim. 2014. Do We Need Hundreds of Classifiers to Solve Real World Classification Problems? *The Journal of Machine Learning Research* 15 (1): 3133–3181.
- Güçdemir, H., and H. Selim. 2015. Integrating Multi-Criteria Decision Making and Clustering for Business Customer Segmentation. *Industrial Management & Data Systems* 115 (6): 1022–1040.
- Hiziroglu, A. 2013. A Neuro-Fuzzy Two-Stage Clustering Approach to Customer Segmentation. *Journal of Marketing Analytics* 1 (4): 202–221.
- Hong, T., and E. Kim. 2012. Segmenting Customers in Online Stores Based on Factors That Affect the Customer's Intention to Purchase. *Expert Systems with Applications* 39 (2): 2127–2131.
- Hosseini, M., and M. Shabani. 2015. New Approach to Customer Segmentation Based on Changes in Customer Value. *Journal of Marketing Analytics* 3: 110–121.
- Jansen, B., S. Jung, and J. Salminen. 2021. The Effect of Hyperparameter Selection on the Personalization of Customer Population Data. *International Journal of Electrical and Computer Engineering Research*. <https://doi.org/10.53375/ijecer.2021.31>.
- Jiang, T., and A. Tuzhilin. 2009. Improving Personalization Solutions through Optimal Segmentation of Customer Bases. *IEEE Transactions on Knowledge and Data Engineering* 21 (3): 305–320.
- Joung, J., and H. Kim. 2023. Interpretable Machine Learning-Based Approach for Customer Segmentation for New Product Development from Online Product Reviews. *International Journal of Information Management* 70: 102641.
- Kamakura, W.A., and G.J. Russell. 1989. A Probabilistic Choice Model for Market Segmentation and Elasticity Structure. *Journal of Marketing Research* 26 (4): 379–390.
- Kashwan, K.R., and C. Velu. 2013. Customer Segmentation Using Clustering and Data Mining Techniques. *International Journal of Computer Theory and Engineering* 5: 856–861.
- Khajvand, M., and M.J. Tarokh. 2011. Estimating Customer Future Value of Different Customer Segments Based on Adapted RFM Model in Retail Banking Context. *Procedia Computer Science* 3: 1327–1332.
- Kitchenham, B., O.P. Brereton, D. Budgen, M. Turner, J. Bailey, and S. Linkman. 2009. Systematic Literature Reviews in Software Engineering—a Systematic Literature Review. *Information and Software Technology* 51 (1): 7–15.
- Kühl, N., M. Schemmer, M. Goutier, and G. Satzger. 2022. Artificial Intelligence and Machine Learning. *Electronic Markets* 32 (4): 2235–2244.
- Kuo, R.J., C.H. Mei, F.E. Zulvia, and C.Y. Tsai. 2016. An Application of a Metaheuristic Algorithm-Based Clustering Ensemble Method to APP Customer Segmentation. *Neurocomputing* 205: 116–129.
- Lee, Y., and S. Cho. 2021. User Segmentation via Interpretable User Representation and Relative Similarity-Based Segmentation Method. *Multimedia Systems* 27 (1): 61–72.
- Lee, J.H., and S.C. Park. 2005. Intelligent Profitable Customers Segmentation System Based on Business Intelligence Tools. *Expert Systems with Applications* 29 (1): 145–152.
- Li, K.J., D.K.H. Fong, and S.H. Xu. 2011. Managing Trade-in Programs Based on Product Characteristics and Customer Heterogeneity in Business-to-Business Markets. *Manufacturing & Service Operations Management* 13 (1): 108–123.



- Liu, D.-R., C.-H. Lai, and W.-J. Lee. 2009. A Hybrid of Sequential Rules and Collaborative Filtering for Product Recommendation. *Information Sciences* 179 (20): 3505–3519.
- Lu, T.-C., and K.-Y. Wu. 2009. A Transaction Pattern Analysis System Based on Neural Network. *Expert Systems with Applications* 36 (3, Part 2): 6091–6099.
- Manidatta, R., M. Ray, K. Muduli, A. Banaitis, and A. Kumar. 2021. Integrated Approach of Fuzzy Multi-Attribute Decision Making and Data Mining for Customer Segmentation. *E a M: Ekonomie a Management* 24: 174–188.
- Manjunath, Y.S.K., and R.F. Kashef. 2021. Distributed Clustering Using Multi-Tier Hierarchical Overlay Super-Peer Peer-to-Peer Network Architecture for Efficient Customer Segmentation. *Electronic Commerce Research and Applications* 47: 101040.
- Munusamy, S., and P. Murugesan. 2020. Modified Dynamic Fuzzy C-Means Clustering Algorithm—Application in Dynamic Customer Segmentation. *Applied Intelligence* 50 (6): 1922–1942.
- Mustak, M., J. Salminen, L. Plé, and J. Wirtz. 2021. Artificial Intelligence in Marketing: Topic Modeling, Scientometric Analysis, and Research Agenda. *Journal of Business Research* 124: 389–404.
- Nemati, Y., A. Mohaghar, M.H. Alavidoost, and H. Babazadeh. 2018. A CLV-Based Framework to Prioritize Promotion Marketing Strategies—Case Study of Telecom Industry. *Iranian Journal of Management Studies*. <https://doi.org/10.22059/ijms.2018.242492.672837>.
- Nie, D., P. Cappellari, and M. Roantree. 2020. A Methodology for Classification and Validation of Customer Datasets. *Journal of Business & Industrial Marketing* 36 (5): 821–833.
- Nisan, S., Ş Elif, and Ç.U. Banu. 2022. Discrete Event Simulation Model Performed with Data Analytics for a Call Center Optimization. *Istanbul Business Research* 51 (1): 189–208.
- Peker, S., A. Kocyigit, and P.E. Eren. 2017. A Hybrid Approach for Predicting Customers' Individual Purchase Behavior. *Kybernetes* 46 (10): 1614–1631.
- Punj, G., and D.W. Stewart. 1983. Cluster Analysis in Marketing Research: Review and Suggestions for Application. *Journal of Marketing Research* 20 (2): 134–148.
- Repschlaeger, J., K. Ereik, and R. Zarnekow. 2013. Cloud Computing Adoption: An Empirical Study of Customer Preferences among Start-up Companies. *Electronic Markets* 23 (2): 115–148.
- Safari, F., N. Safari, and G.A. Montazer. 2016. Customer Lifetime Value Determination Based on RFM Model. *Marketing Intelligence & Planning* 34: 446–461.
- Salminen, J., V. Yoganathan, J. Corporan, B.J. Jansen, and S.-G. Jung. 2019. Machine Learning Approach to Auto-Tagging Online Content for Content Marketing Efficiency: A Comparative Analysis between Methods and Content Type. *Journal of Business Research* 101: 203–217.
- Salminen, J., K. Chhirang, S.-G. Jung, S. Thirumuruganathan, K.W. Guan, and B.J. Jansen. 2022. Big Data, Small Personas: How Algorithms Shape the Demographic Representation of Data-Driven User Segments. *Big Data*. <https://doi.org/10.1089/big.2021.0177>.
- Seret, A., S. Maldonado, and B. Baesens. 2015. Identifying next Relevant Variables for Segmentation by Using Feature Selection Approaches. *Expert Systems with Applications* 42 (15): 6255–6266.
- Simões, D., and J. Nogueira. 2022. Learning about the Customer for Improving Customer Retention Proposal of an Analytical Framework. *Journal of Marketing Analytics*. <https://doi.org/10.1057/s41270-021-00126-7>.
- Spoor, J.M. 2022. Improving Customer Segmentation via Classification of Key Accounts as Outliers. *Journal of Marketing Analytics*. <https://doi.org/10.1057/s41270-022-00185-4>.
- Sun, Z.-H., T.-Y. Zuo, D. Liang, X. Ming, Z. Chen, and S. Qiu. 2021. GPHC: A Heuristic Clustering Method to Customer Segmentation. *Applied Soft Computing* 111: 107677.
- Symeonidis, S., D. Effrosynidis, and A. Arampatzis. 2018. A Comparative Evaluation of Pre-Processing Techniques and Their Interactions for Twitter Sentiment Analysis. *Expert Systems with Applications* 110: 298–310.
- Thirumuruganathan, S., N.A. Emadi, S. Jung, J. Salminen, D.R. Robillos, and B.J. Jansen. 2023. Will They Take This Offer? A Machine Learning Price Elasticity Model for Predicting Upselling Acceptance of Premium Airline Seating. *Information & Management* 60: 103759.
- Tsai, C.-F., Y.-H. Hu, and Y.-H. Lu. 2015. Customer Segmentation Issues and Strategies for an Automobile Dealership with Two Clustering Techniques. *Expert Systems* 32 (1): 65–76.
- Wang, S.-C., Y.-T. Tsai, and Y.-S. Ciou. 2020. A Hybrid Big Data Analytical Approach for Analyzing Customer Patterns through an Integrated Supply Chain Network. *Journal of Industrial Information Integration* 20: 100177.
- Warner, L.A. 2019. Using Homeowners' Association Membership to Define Audience Segments for Targeted Local Social Marketing Interventions: Implications From a Statewide Study. *Social Marketing Quarterly* 25 (4): 291–307.
- Wong, E., and Y. Wei. 2018. Customer Online Shopping Experience Data Analytics: Integrated Customer Segmentation and Customised Services Prediction Model. *International Journal of Retail & Distribution Management* 46 (4): 406–420.
- Wu, T., and X. Liu. 2020. A Dynamic Interval Type-2 Fuzzy Customer Segmentation Model and Its Application in E-Commerce. *Applied Soft Computing* 94: 106366.
- Wu, S., W.-C. Yau, T.-S. Ong, and S.-C. Chong. 2021. Integrated Churn Prediction and Customer Segmentation Framework for Telco Business. *IEEE Access* 9: 62118–62136.
- Yao, Z., P. Sarlin, T. Eklund, and B. Back. 2014. Combining Visual Customer Segmentation and Response Modeling. *Neural Computing and Applications* 25 (1): 123–134.
- Zhou, J., J. Wei, and B. Xu. 2021. Customer Segmentation by Web Content Mining. *Journal of Retailing and Consumer Services* 61: 102588.
- Aryuni, M., E.D. Madyatmadja, and E. Miranda. 2018. Customer Segmentation in XYZ Bank Using K-Means and K-Medoids clustering. In: 2018 International Conference on Information Management and Technology (ICIMTech), 412–416, 2018. IEEE.
- Bhade, K., V. Gulalkari, N. Harwani, and S.N. Dhage. 2018. A Systematic Approach to Customer Segmentation and Buyer Targeting for Profit Maximization. In: 2018 9th International Conference on Computing, Communication and Networking Technologies (ICCCNT), 1–6, 2018. IEEE.
- Dzulhaq, M.I., K.W. Sari, S. Ramdhan, and R. Tullah. 2019. Customer Segmentation Based on RFM Value Using K-Means Algorithm. In: 2019 Fourth International Conference on Informatics and Computing (ICIC), 1–7, 2019. IEEE.
- Jung, S.-G., J. Salminen, J. An, H. Kwak, and B.J. Jansen. 2018. Automatically Conceptualizing Social Media Analytics Data via Personas. In: Proceedings of the International AAAI Conference on Web and Social Media (ICWSM 2018), June 25, 2018. San Francisco, California, USA.
- Kandeil, D.A., A.A. Saad, and S.M. Youssef. 2014. A Two-Phase Clustering Analysis for B2B Customer Segmentation. In: 2014 International Conference on Intelligent Networking and Collaborative Systems, 221–228, 2014. IEEE.
- Pramono, P.P., I. Surjandari, and E. Laoh. 2019. Estimating Customer Segmentation Based on Customer Lifetime Value Using Two-Stage Clustering Method. In: 2019 16th International Conference on Service Systems and Service Management (ICSSSM), 1–5, 2019. IEEE.



- Ranjan, A., and S. Srivastava. 2022. Customer Segmentation Using Machine Learning: A Literature Review. In: AIP Conference Proceedings, 2022. vol. 2481. AIP Publishing LLC, p. 020036.
- Syakur, M.A., B.K. Khotimah, E.M.S. Rochman, and B.D. Satoto. 2018. Integration K-Means Clustering Method and Elbow Method for Identification of the Best Customer Profile Cluster. In: IOP Conference Series: Materials Science and Engineering, 2018. vol. 336, 012017. IOP Publishing.
- Xu, X., N. Yuruk, Z. Feng, and T.A. Schweiger. 2007. Scan: A Structural Clustering Algorithm for Networks. In: Proceedings of the 13th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 824–833, 2007.
- Zhu, B., C. He, and X. Jiang. 2015. *A Consistency-Based Validation for Data Clustering—IOS Press* [Online]. <https://content.iospress.com/articles/intelligent-data-analysis/ida727>. Accessed 3 Mar 2023.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Joni Salminen is an Associate Professor (tenure track) at the University of Vaasa, Vaasa, Finland. His research focuses on personas and customer segmentation, along with overlapping topics such as

Human-Centered AI, Quantitative UX, Interactive Systems, Algorithms in Marketing, and so on.

Mekhail Mustak is an Assistant Professor (tenure track) at Hanken School of Economics, Helsinki, Finland. His research focuses on the application of artificial intelligence in marketing, B2B marketing, and services marketing.

Muhammad Sufyan is a post-doctoral researcher at the strategy and entrepreneurship research group of Jyväskylä University School of Business and Economics in Jyväskylä, Finland. His research interests include digital entrepreneurship, transnational migrant entrepreneurship, new venture creation, and micro-foundations in entrepreneurship.

Bernard J. Jansen is a Principal Scientist at the Qatar Computing Research Institute, Doha, Qatar. He is a graduate of West Point and has a Ph.D. in computer science from Texas A&M University, along with master degrees from Texas A&M (computer science) and Troy State (international relations). Dr. Jim Jansen served in the U.S. Army as an Infantry enlisted soldier and communication commissioned officer.

