SEGMENTATION OF THE AUDIENCE BASED ON BEHAVIORAL AND DEMOGRAPHIC DATA: PROSPECTS FOR PERSONALIZED MARKETING

A.U. Makhtibekov, *Master's degree*
Bay Atlantic University
(USA, Washington)

Abstract. *The article examines audience segmentation methods based on behavioral and demographic data to enhance the effectiveness of personalized marketing. It analyzes machine learning technologies and algorithms, such as clustering and regression, which help create accurate consumer profiles for targeted marketing campaigns. Special attention is given to data integration from various sources, as well as the impact of a personalized approach on business metrics such as engagement, conversion, and customer retention. The study demonstrates that the use of modern methods can significantly increase companies' competitiveness in the modern market.*

Keywords: *audience segmentation, personalized marketing, behavioral data, demographic data, machine learning, marketing campaigns.*

Modern digital marketing is developing in the context of volumes of consumer data that keep on growing fast. Companies increasingly use tailor-made approaches with the aim to maximize the success of ad campaigns, increase user engagement, and optimize conversion. One of the most significant tools for personalization is audience splitting based on combining demographic and behavior data.

Conventional segmentation techniques, for example, by age, sex, or geographic location, are very far from being a guarantee of the saliency and validity of a message. Behavioral data, recorded browsing history, clicks, purchases, and time spent with content, outline deep user preferences and allow for the forecasting of his further actions. It is these data, in conjunction with the consumer's demographic characteristics, that form the basis for building highly detailed profiles of the latter.

The aim of this paper is to examine the effectiveness of personalized segmentation through the integration of behavioral and demographic data, explore data analysis methods and machine learning (ML) algorithms, and assess their impact on marketing performance. Additionally, the study analyzes the prospects for the development of personalization in digital marketing and considers potential limitations related to user data processing.

Data sources for personalized segmentation

Personalized audience segmentation relies on a wide range of data, which is formally divided into two levels: demographic and behavioral. The data gives the basis for the correct classification of users and establishes the proper manner of elaboration of adequate marketing strategies.

Demographic data is one of the traditional sources of user information. The parameters included in this section are age, gender, location, income level, education, and marital status. This is normally gathered through questionnaires, website registration forms, and social media. Despite their wide use, demographic characteristics are limited to a certain degree of accuracy. They only give general context to a user, and it's hard to draw deep insights about the behavior and needs of a user using this data alone [1].

At the same time, great significance is taken by **behavioral data** in the segmentation process: information about user actions, for example, the frequency and quantity of visits to the site, purchase behavior, clicking on ads, time spent on pages, interaction with content. This is the data that gives a much more objective picture of what actually is of interest to users. Working with such data lets you create more complex and detailed consumer profiles, which raises the effectiveness of marketing campaigns several times over. Behavioral data is often taken from website analytics systems, mobile applications, and social platforms such as **Google Analytics** or **Facebook Insights*** (принадлежит компании Meta, признанной экстремистской и запрещённой на территории РФ).

The use of application programming interfaces (API) and integration with external information sources, such as CRM systems and e-commerce platforms, allows combining different types of data for more precise segmentation. For example, when combined with demographic data, behavioral information helps determine which users with certain characteristics are more likely to make purchases or respond to advertising. Predictive models built on this data allow predicting future user actions and thus planning marketing campaigns more effectively [2].

It is also about integrating unstructured data, like user reviews and comments in social networks, which opens up new opportunities for increasing personalization. NLP and sentiment analysis let one extract from texts certain useful information that afterwards can be used to enhance user profiles and increase the accuracy of their segmentation.

Data sources for personalized segmentation become more and more diverse and complex.

Let's point out that the quality and completeness of data directly affect the accuracy and effectiveness of marketing strategies. Only the comprehensive approach, using demographic and behavioral data in complex with new methods of information processing, allows achieving maximum results in personalization and increasing conversion.

Methods of data analysis and integration

Effective personalized audience segmentation requires not only collecting different types of data, but also using special methods for analyzing and integrating them. Methods for processing and combining data play a key role in creating accurate consumer profiles that can be used for targeted marketing campaigns. It is important to consider that data can be both structured (e.g., numeric indicators and categories) and unstructured (texts, images), which requires using different approaches to analysis (table 1).

Table 1. Data analysis methods and their integration for personalized marketing [3]

Data analysis method	Algorithms/tools	Application	Advantages	Limitations
Clustering	K-means, hierarchical clustering	Grouping the audience into segments based on similar behavior.	Helps uncover hidden audience segments.	Requires large data volumes for accuracy.
Regression	Linear regression, logistic regression	Predicting the likelihood of purchase.	High accuracy for predicting based on available data.	Requires clean and complete data.
Neural networks	Deep learning, recurrent neural networks	Building complex models to predict customer preferences.	High flexibility and ability to detect complex patterns.	High computational cost.
Time series analysis	ARIMA, exponential smoothing	Forecasting purchases and activity based on time.	Well-suited for trend and seasonality analysis.	May be inaccurate with unstable data.
Association rules	Apriori algorithm, FP-growth	Finding frequent patterns in customer behavior.	Helps uncover hidden behavioral patterns.	Requires high-quality data.

One of the major methods of data analysis is **clustering**, which enables the grouping of users by characteristics. Clustering typically uses such algorithms **as k-means** and **DBSCAN**. The k-means method is one of the most popular approaches to segmentation, since it divides users into k groups in accordance with their characteristics, whereby you can find different segments according to behavioral and demographic data. In turn, the DBSCAN algorithm is applicable in cases when it is necessary to identify segments with an uneven distribution of data and does not require a pre-defined number of clusters. This is useful for the segmentation of the audience in the case of complicated and non-standard data structures.

Another useful method of analysis involves **regression**, which can allow forecasting future user actions based on available data. Using linear or logistic regression methods, one can model the probability of a user taking certain actions-like making a purchase-given certain demographic and behavioral attributes. This approach can raise the accuracy of predictions manifold and make offers more personalized.

Feature engineering serves as a key approach for integrating multiple data types, such as demographic and behavioral information. This process involves the creation of new features that more accurately represent user needs and interests. For example, combining age and purchase frequency into a single feature, such as «buyer activity», can help identify high-value customer segments. This approach contributes to the precision of user classification and improves behavioral prediction accuracy.

Also of interest are methods of processing unstructured data, such as free-text commentaries, that can be processed by means of applying NLP technology. Using semantic information extraction and sentiment analysis algorithms, it is possible to extract useful information from user reviews and comments, which further improves the quality of segmentation [4].

Thus, effective data analysis and integration are an integral part of personalized audience segmentation. The use of various analysis methods, such as clustering, regression and feature engineering, in combination with new technologies for processing unstructured data, allows to significantly increase the accuracy of marketing forecasts and create more relevant offers for consumers.

The impact of personalized marketing on business performance

Personalization directly influences the performance of a business within an organization, resulting not only in growth in user interaction but also in growth in conversion, customer loyalty and, ultimately, business profitability. Personalization enables you to more effectively customize marketing activity to users' needs and interests, greatly enhancing the level of interaction with the audience as well as resulting in the formation of loyalty. It is also worth considering that users have become more likely to expect personalized content from brands (fig.1).
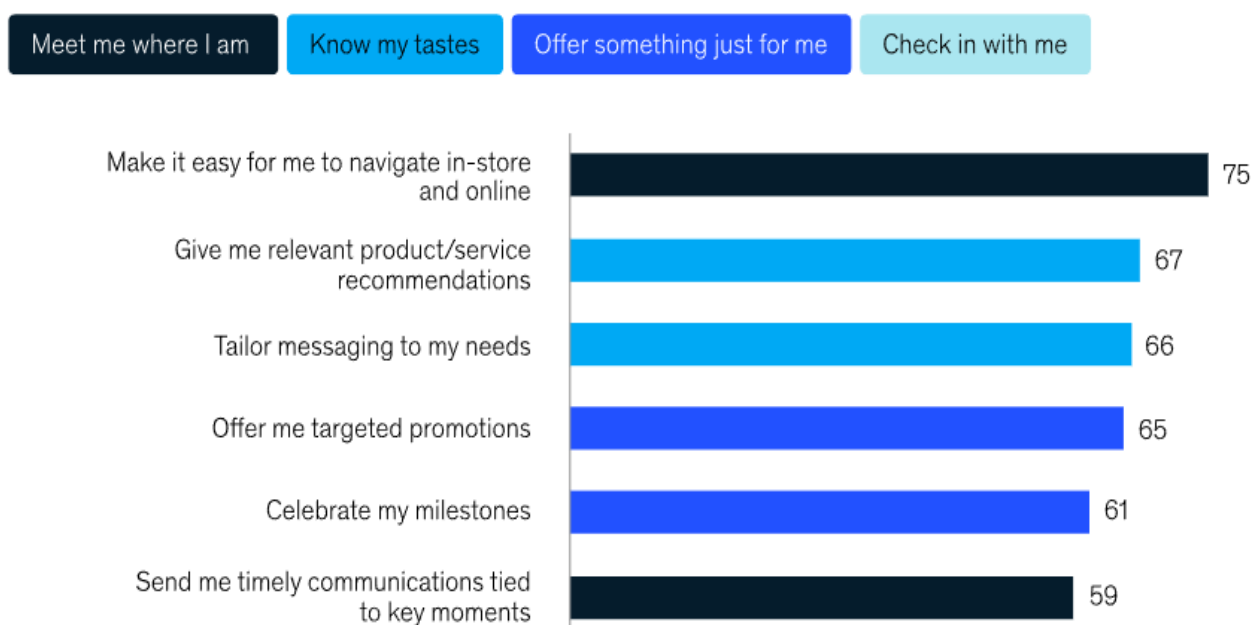

Fig. 1. Consumers' expectations for different types of personal content, % [5]

The simplest effect of personalized marketing is an **increase in user response**. If a marketing message corresponds to the expectations and interests of a customer, the possibility of responding with his or her offer increases significantly. This allows companies to achieve noticeable results at relatively lower costs for marketing campaigns.

Conversion is another key indicator that improves the personalized approach. The use of ML methods and behavioral data analysis allows us to accurately predict which users are more likely to make a purchase. Predicted consumer behavior allows us to create targeted offers, which reduces the number of non-targeted requests and increases the effectiveness of advertising [6].

Another critical question of personalized marketing success is **customer retention.** It is always more expensive to acquire a new customer than to retain an old one. Personalized offers and recommendations, received from the service based on his previous experience, help increase

satisfaction and loyalty [7]. By implementing personalized marketing strategies, it's not just increasing the number of new customers that can be achieved but also improving relationships with the currently existing ones for long-term business growth. Besides, personalization in marketing means fewer customer churns. Simultaneously, creating personalized offers and improving customer experience allows competing for users' attention more effectively, which then is transformed into a stronger position in the market [8].

Thus, the very influence of personalized marketing on business performance is reflected in the improvement of key indicators such as engagement, conversion, customer retention, and churn reduction. The implementation of personalized strategies allows companies not only to increase competitiveness but also to increase profitability by reducing the costs of mass advertising.

Ethical and legal aspects of user data processing in personalized marketing

The use of targeted marketing strategies is intrinsically linked with the collection, analysis, and utilization of huge volumes of user data, both structured and unstructured information drawn from demographic and behavior-based inputs. With increased processing power for data, particularly through the incorporation of machine learning methods and heightened analytics, the ethical and legal dimensions to data use have become as integral to the discussion as model performance of segmentation models or campaign reports.

One of the most significant concerns here is how to render **international and regional regulatory compliances,** such as the European Union's General Data Protection Regulation (GDPR), the California Consumer Privacy Act (CCPA), and other data privacy laws, secure. These legal instruments formulate general principles of data processing like the necessity of an informed user's consent, right of access, modify, and delete personal data, and obligation to provide efficient protection against unauthorized access or loss of data.

One of the greatest challenges for marketers is to **find the right degree of personalization within the limits of digital privacy**. Behavior analysis that could include tracking user clicks, session duration, purchase history, and content interaction across websites becomes extremely sensitive when combined with demographic pro-

files. Multidimensional integration increases the likelihood of re-identification even when data sets are technically anonymized, thereby undermining user privacy and potentially leading to trust violation.

A second critical ethical dimension is **explainability and transparency of personalized decisions.** As algorithms intervene more in what users see, what offers they are exposed to, and how they are categorized, there is a growing need for companies to ensure that these activities are not only technically sound but also intelligible to non-experts. The principle of algorithmic accountability demands that automatic decisions are interpretable, contestable, and auditable. Consumers must be notified when automated profiling has a considerable effect on their online experience and be provided with ways to opt out or pursue redress. This guarantees that personalization continues as a tool of relevance as opposed to manipulation, maintaining user autonomy and equity.

Additionally, the widespread application of artificial intelligence in advertising has ushered in the **issue of algorithmic opacity**. Labeled as «black box» models, they operate in a way that is not so easy to comprehend even for their designers. That these operations are opaque complicates regulatory audits and ethical judgment of fairness, bias propensity, and compliance with regulations.

Perhaps the highest ethical risk of personalized marketing is **algorithmic discrimination**. If data-driven algorithms are used to design offers as a function of socio-economic attributes, the risk is that prevailing inequalities will be entrenched by limiting access to service or setting prices to specific user groups. Low-quality or biased training data can also lead to social stereotypes, which stereotype individuals away by excluding them from exact targeting marketing campaigns.

Regulatory compliance is now a significant aspect of consumer trust as well as corporate reputation. Organizations which actively adopt ethical data practices like privacy by design, data minimization, and user control of personal data evidence digital responsibility and are more likely to exhibit long-term user engagement. Ethically treating user information not only safeguards legal liability but is also a major differentiator in a competitive market, where customers are ever

more aware of their online rights and right to privacy.

Successful personalized marketing, therefore, requires more than technical expertise with information analysis. It requires extensive knowledge of the legal and ethical infrastructure on which digital behavior rests. Embedding data ethics as a culture in organizational structures, developing processes to make algorithms answerable, renewing privacy policies in light of current standards, and providing open channels of communication with users are all critical steps. Through a comprehensive and ethically informed approach to data-driven personalization alone can sustainable development and legitimacy be attained in the data-driven marketing landscape of today.

Conclusion

In time, personalized audience segmentation will surely become one of the key tools for optimizing marketing strategy results. ML algorithms, combined with different types of integrated data, manage to precisely recognize users' needs and preferences more and more, enabling value propositions to become more relevant and attractive. Engagement, conversion, and customer loyalty are all deeply influenced by that.

Nevertheless, personalized marketing is far from being that easy to be realized. First, high accuracy of segmentation depends on good quality and completeness of the data. Poor or distorted data result in wrong conclusions and low effectiveness of marketing campaigns. Second, using algorithms requires a great number of computational resources that may not be available to every small company with limited budgets.

Personalized marketing is, therefore, a tool of effectiveness only in the case when this area is complex. Above all, when working with the modern approaches of data analysis and processing, it is necessary to consider not just opportunities but also possible limitations linked with the access to high-quality data and computing resources. A company that is capable to handle such challenges will significantly raise its competitiveness on the market and reduce its mass advertising expenses.

References

1. Kedi W.E. AI software for personalized marketing automation in SME: Enhancing customer experience and sales // World Journal of Advanced Research and Reviews. – 2024. – Vol. 23. – № 1. – P. 1981-1990.

2. Smirnova E. Developing digital transformation strategies in enterprise operations management // Professional Bulletin: Economics and Management. – 2024. – №2/2024. – P. 11-15.

3. Adeniran I.A. Transforming marketing strategies with data analytics: A study on customer behavior and personalization // International Journal of Management & Entrepreneurship Research. – 2024. – Vol. 6, № 8. – P. 41-51.

4. Tretiakov I. AI as a tool for enhancing customer experience and increasing consumer loyalty // Danish scientific journal. – 2024. – № 90. – P. 138-141.

5. The value of getting personalization right – or wrong – is multiplying / McKinsey's. – URL: https://www.mckinsey.com/capabilities/growth-marketing-and-sales/our-insights/the-value-of-getting-personalization-right-or-wrong-is-multiplying.

6. Iyelolu T.V. Leveraging artificial intelligence for personalized marketing campaigns to improve conversion rates // International Journal of Engineering Research and Development. – 2024. – Vol. 20. – № 8. – P. 253-270.

7. Kuznetsov I. Hybrid recommender systems for mobile apps: using collaborative filtering and content personalization // International Research Journal of Modernization in Engineering Technology and Science. – 2024. – Vol. 6. – № 12. – P. 1152-1156.

8. Todorova A., Ruskova S. Empathy in Marketing: A Customer-Oriented Approach based on Authenticity // Proceedings of the international conference on economics and social sciences. Bucharest University of Economic Studies, Romania. – 2024. – Vol. 6. – № 1. – P. 1076-1085.

СЕГМЕНТАЦИЯ АУДИТОРИИ НА ОСНОВЕ ПОВЕДЕНЧЕСКИХ И ДЕМОГРАФИЧЕСКИХ ДАННЫХ: ПЕРСПЕКТИВЫ ДЛЯ ПЕРСОНАЛИЗИРОВАН-НОГО МАРКЕТИНГА

А.У. Махтибеков, *магистр*
Университет Бэй Атлантик
(США, г. Вашингтон)

Аннотация. В статье рассматриваются методы сегментации аудитории на основе поведенческих и демографических данных с целью повышения эффективности персонализированного маркетинга. Анализируются технологии и алгоритмы машинного обучения, такие как кластеризация и регрессия, которые помогают создать точные профили потребителей для создания таргетированных маркетинговых кампаний. Особое внимание уделяется интеграции данных из различных источников, а также влиянию персонализированного подхода на бизнес-показатели, такие как вовлеченность, конверсия и удержание клиентов. Исследование показывает, что использование современных методов может значительно повысить конкурентоспособность компаний в условиях современного рынка.

Ключевые слова: сегментация аудитории, персонализированный маркетинг, поведенческие данные, демографические данные, машинное обучение, маркетинговые кампании.