Jurnal Aplikasi Manajemen dan Bisnis, Vol. 10 No.1, January 2024
Permalink/DOI: http://dx.doi.org/10.17358/jabm.10.1.209

Available online
http://journal.ipb.ac.id/index.php/jabm

CUSTOMER SEGMENTATION FOR DIGITAL MARKETING BASED ON SHOPPING PATTERNS

Juhaini Alie*)1, Rendra Gustriansyah)**

*) Department of Management, Faculty of Economics, Universitas Indo Global Mandiri
Jl. Jenderal Sudirman No. 629 Palembang 30129, Indonesia
**) Department of Informatics, Faculty of Computer Science, Universitas Indo Global Mandiri
Jl. Jenderal Sudirman No. 629 Palembang 30129, Indonesia

Article history:
Received
21 March 2023

Revised
30 July 2023

Accepted
20 August 2023

Available online
31 January 2024

Abstract: Customer segmentation is the customer grouping based on similar shopping behavior or patterns. Inappropriate customer segmentation can have negative impacts, such as lost marketing opportunities, resource inefficiencies, loss of potential customers, and decreased performance, and business profits, especially in customer satisfaction. Therefore, this study aims to develop a customer segmentation model for digital marketing. This model is based on customer shopping patterns using the Recency-Frequency-Monetary (RFM) model and the Partitioning Around Medoids (PAM) method. The research data is historical customer purchase data consisting of 18,535 transactions and 541,909 transaction details from 4,339 customers for 3,665 product items over two years. The research variables focus on the model used: recency, frequency, and monetary. The five customer segments generating from this study are main, potential, general, minimum, and prospective customer. The internal validation results show that the minimum C-Index value is 0.1429 (close to zero), and the maximum Calinski-Harabasz Index value is 512.9553. It shows that the quality of customer segmentation results is good. In other words, the model can identify correlations between customer segments and shopping patterns and preferences. In this way, marketers can optimize services, adjust strategies, and offer the right products for each customer segment. Further research can be directed at product segmentation.

Keywords: partitioning around medoids, digital marketing, shopping pattern, recency-frequency-monetary, customer segmentation

Abstrak: *Segmentasi pelanggan adalah pengelompokan pelanggan berdasarkan perilaku atau pola belanja yang serupa. Segmentasi pelanggan yang tidak tepat dapat berdampak negatif seperti hilangnya peluang pemasaran, inefisiensi sumber daya, kehilangan pelanggan yang potensial, penurunan kinerja dan keuntungan bisnis, terutama dalam hal kepuasan pelanggan. Oleh karena itu, penelitian ini bertujuan untuk mengembangkan model segmentasi pelanggan untuk digital marketing. Model ini didasarkan pada pola belanja pelanggan menggunakan model Recency-Frequency-Monetary (RFM) dan metode Partitioning Around Medoids (PAM). Data penelitian merupakan data histori pembelian pelanggan yang terdiri dari 18.535 transaksi dan 541.909 detil transaksi dari 4.339 pelanggan untuk 3.665 item produk selama periode dua tahun. Variabel penelitian terfokus pada model yang digunakan yaitu recency, frequency, dan monetary. Lima segmen pelanggan yang dihasilkan dari penelitian ini adalah segmen utama, potensial, umum, minimum, dan calon pelanggan. Hasil validasi internal menunjukkan bahwa nilai C-Index minimum adalah 0,1429 (mendekati nol) dan nilai Calinski-Harabasz Index maksimum adalah 512,9553. Hal ini menunjukkan bahwa kualitas hasil segmentasi pelanggan adalah baik. Dengan kata lain, model dapat mengidentifikasi korelasi antara segmen pelanggan terhadap pola belanja dan preferensi. Dengan demikian, para pemasar dapat mengoptimalkan layanan, menyesuaikan strategi, dan penawaran produk yang tepat untuk setiap segmen pelangan tersebut. Adapun, studi di masa depan dapat diarahkan untuk segmentasi produk.*

Kata kunci: *partitioning around medoids, pemasaran digital, pola belanja, recency-frequency-monetary, segmentasi pelanggan*

[1] Corresponding author:
Email: ruditriyono83@gmail.com

INTRODUCTION

In the marketing world, every customer is unique because they have different behaviors and preferences. Therefore, customer segmentation is needed as a key marketing strategy (Riza, Seminar and Maulana, 2018). Customer segmentation can help marketers to find correlations between certain customer segments with their shopping patterns and preferences so that marketers can adjust strategies, messages, promotions and special offers for these customer segments (Tabianan, Velu and Ravi, 2022). It can retain valuable customers by providing a customized experience for each segment, increasing customer loyalty, providing more efficient time, cost and labor resources, increasing sales and optimizing business profits. Improper customer segmentation can cause lost marketing opportunities, create resource inefficiencies, trigger loss of potential customers, and even reduce business performance and profits, especially in terms of customer satisfaction (Wang, 2022).

In the e-commerce industry, customer segmentation is a competitive and challenging field (Joung and Kim, 2023). This motivates researchers to identify effective customer segmentation models. Generally, customer segmentation uses the RFM model to analyze customer shopping patterns in the past (Abbasimehr and Bahrini, 2022; Handojo et al. 2023; Monalisa et al. 2023). He and Cheng stated that the RFM model can also provide some important information for customer segmentation (He and Cheng, 2021). The RFM model can be used for the development of an effective marketing strategy (Monalisa, Nadya and Novita, 2019). In addition, the RFM model can also be widely implemented in the sales field (Gustriansyah, Suhandi and Antony, 2020; Gustriansyah, Ermatita and Rini, 2022).

RFM is a simple but effective marketing analysis model to measure customer shopping patterns based on the last shopping time interval (recency), spending intensity (frequency), and total money spent (monetary) (Gustriansyah, Suhandi and Antony, 2020)Frequency, and Monetary. RFM is a simple but effective method that can be applied to market segmentation. RFM analysis is used to analyze customer's behavior which consists of how recently the customers have purchased (recency. Traditional customer segmentation methods involve manual analysis and limited sample sizes. However, the combination of RFM models with machine learning methods can increase the scalability of data processing in real time, identify patterns, correlations and trends that human analysts may miss, automate processes, save time and resources (Verma et al. 2021). The application of the machine learning method allows marketers to respond and adapt to the dynamics of changing shopping patterns and customer preferences so that marketers can deliver messages to the right customers and at the right time, thereby increasing the effectiveness of company promotions (Christy et al. 2021). The combination of RFM and machine learning methods can generate various strategies for customer segmentation (Wong and Wei, 2018; Rahim et al. 2021; Serper, Şen and Çalış Uslu, 2022; Salminen et al. 2023).

Machine learning enables marketers to analyze customer shopping history data in real-time, follow patterns and dynamics of changes in customer preferences (Aryuni, Madyatmadja and Miranda, 2018; Tabianan, Velu and Ravi, 2022; Sun, Liu and Gao, 2023). Riza, et al. used the k-means and apriori methods to segment potential customers based on customer characteristics and transaction patterns with certain types of merchants (Riza, Seminar and Maulana, 2018). Lee, et al. applied several machine learning approaches such as k-mean, fuzzy c-mean, and Wald's method for customer segmentation (Lee et al. 2021). The approach using the machine learning method also allows marketers to analyze customer shopping history data in real-time, follow patterns and dynamics of changes in customer preferences so as to enable marketers to understand their customers, provide personalized experiences by conveying messages, recommendations and offers that are aligned with each customer's preferences (Zhang, Lin and Simeone, 2022). This personalization can grow customer loyalty, increase sales, and increase a competitive advantage (Li et al. 2021).

Based on a review of some literature, it can be concluded that research on RFM models integrated with machine learning methods in the last five years is still minimal, so that it provides anopportunity to conduct research in this field by adopting other methods. Therefore, this study aims to (1) develop a customer segmentation model for digital marketing on e-commerce platforms based on customer shopping patterns or behavior by integrating the classic RFM model with one of the machine learning (PAM) methods. The use of the PAM method is intended to update the standard segmentation method of the RFM model in order to obtain more proportional and optimal segmentation results. The

PAM method was selected because it has the ability to cluster datasets that contain outliers (Pierpaolo et al. 2022). (2) identify the right customer segments so that marketers can deliver messages, recommendations and promotions according to customer preferences. This level of personalization can foster customer loyalty, increase engagement, and drive higher shopping intent. Meanwhile, the contribution of this research is to increase sales by streamlining promotions and shopping experiences according to customer segments. In addition, this research also provides an understanding of how machine learning (PAM) technology enables marketers to personalize messages, recommendations, and offers that meet customer preferences.

METHODS

This research is quantitative and uses an e-commerce dataset derived from the Kaggle repository (Daniel, 2019). This dataset contains customer purchase history data which includes 18,535 transactions and 541,909 transaction details from 4,339 customers for 3,665 product items over a two-year period. Customer shopping patterns were analyzed and segmented based on the RFM model and PAM method.

Meanwhile, this study hypothesizes that machine learning (PAM) technology can update the standard segmentation method of the RFM model to produce more proportional and optimal segmentation. Therefore, machine learning allows marketers to analyze customer shopping history data in real time, following patterns and dynamics of changes in customer preferences. Figure 1 shows the customer segmentation framework for identifying shopping patterns and customer preferences using the RFM model with a limited sample size. However, combining the RFM model with the PAM method can increase the scalability of data processing in real time, identifying patterns, correlations, and trends that human analysts may miss. Furthermore, the research stages consist of (1) data preparation; (2) customer segmentation; (3) segment quality measurement; (4) interpretation of segmentation results. This study uses R-Studio as a tool for data processing and visualization.

Before analyzing the dataset further, it is necessary to transform the dataset into a certain format so that the processing becomes easier and more effective, the segmentation results become more accurate, and the computation time is minimized. The data preparation stage includes: (1) cleaning incomplete, noisy (data error or outlier), inconsistent, and duplicative data; (2) data normalization using the z-score formulation in equation (1) and the natural logarithm (log base 10) (Gustriansyah, Ermatita and Rini, 2022). Z-score (z) is a measure of how far the sample (x) is from the sample mean ($\overline{x}$) and is inversely proportional to the standard deviation (SD).

$$z = (x-\overline{x})/SD$$

For the normalized dataset, the average value of each RFM variable is calculated. Furthermore, the dataset will be segmented into five customer segments: main, potential, minimum, general, and prospective customer (Wang, 2022). Customer segmentation in this study uses the PAM method (Schubert and Rousseeuw, 2021) which is included in the factoextra package (Kassambara and Mundt, 2020) for R programming. This segmentation will produce customer clusters with a specific average RFM value for each segment.

Segments that have average F and M values that are higher than the average F and M values of tha whole data will be given an indicator ↑. Meanwhile, segments with a lower average value will be given the indicator ↓. For R, if the average value of each segment is higher than the R value for the whole data, the symbol will be ↓. Conversely, if the average value is lower, it will be given the symbol ↑. This is because, the shorter the last shopping interval, the higher the R value.

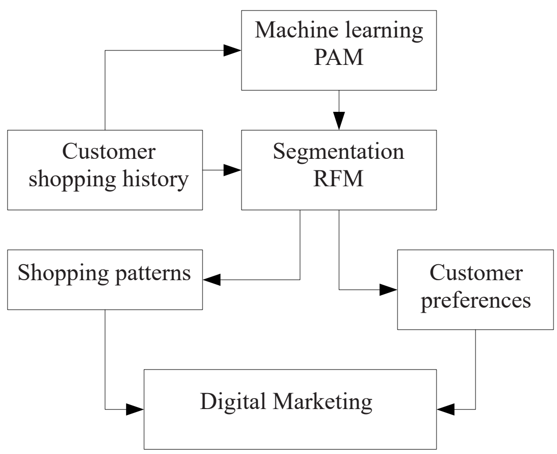

Figure 1. The customer segmentation framework

Segment quality will be validated using the C-Index (CI) and Calinski-Harabasz Index (CHI) available in the NbClust package (Charrad et al. 2014) for R-programming. The C-Index measures intra-segment variance against the ideal segmentation as shown in equation (2) (Bezdek et al. 2016). CI values are in the interval 0 – 1, where the smaller the CI value, the better the quality of the segmentation results.

$$CI=(S_w - S_{min})/(S_{max} - S_{min}); S_w \neq S_{min}; CI \in (0,1)$$

Where S_w is the sum of all within-cluster dissimilarities; S_{min} is the sum of smallest within-cluster dissimilarity; S_{max} is the sum of largest within-cluster dissimilarity.

Meanwhile, CHI measures segmentation results based on the proportion of the inter-segment variance (BSS) to the intra-segment variance (WSS), which is then multiplied by a normalization factor. This normalization factor represents the proportion of the difference in the amount of data to the number of segments, compared to the number of segments minus one (Gustriansyah, Ermatita and Rini, 2022). The greater the CHI value, the better the quality of the segmentation results (Charrad et al. 2014).

$$CHI(k) = (BSS/(k-1))/(WSS/(n-k))$$

Where n = the number of all objects studied and k = the number of segments.

RESULTS

The data preparation stage resulted in a dataset that had been cleaned from incomplete, duplicate, outlier, and inconsistent data. This filtering process left 18,533 transaction and 397,923 detailed transaction data from 4,337 customers. If this transaction data is visualized in 3D, the data distribution is based on RFM as shown in Figure 2. The data distribution was uneven. Therefore, the dataset was normalized using z-score and natural logarithm scaling. The normalization results show a more even (normal) distribution of data as shown in Figure 3.

A normally distributed dataset is more representative and easier to segment. Data with the highest RFM value are is located in the upper right corner of the graph. Meanwhile, the data with the lowest RFM value is located in the lower left corner of the graph. The average value for each RFM variable after being analyzed based on the RFM model is listed in Table 1. Furthermore, the dataset was segmented using the PAM method. The segmentation results with the average RFM value for each segment are listed in Table 2. Meanwhile, the RFM indicators and customer criteria for each segment are shown in Table 3. The visualization of customer segmentation results based on the RFM model and the PAM method is shown in Figure 4. Meanwhile, an example of customer segmentation is listed in Table 4.

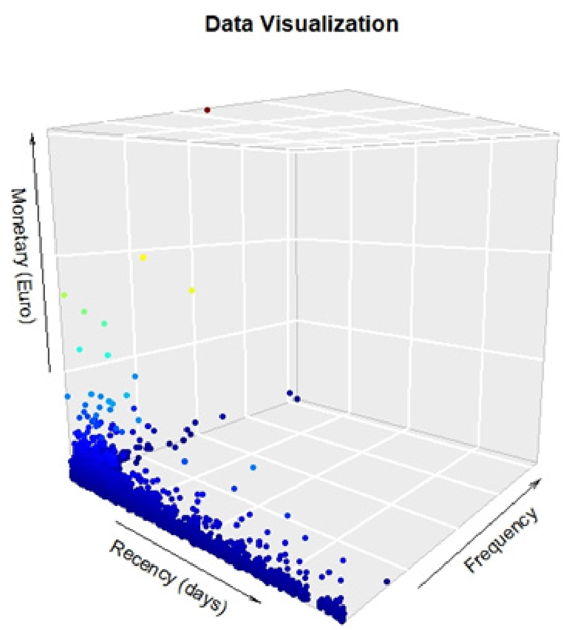

Figure 2. The data distribution based on RFM

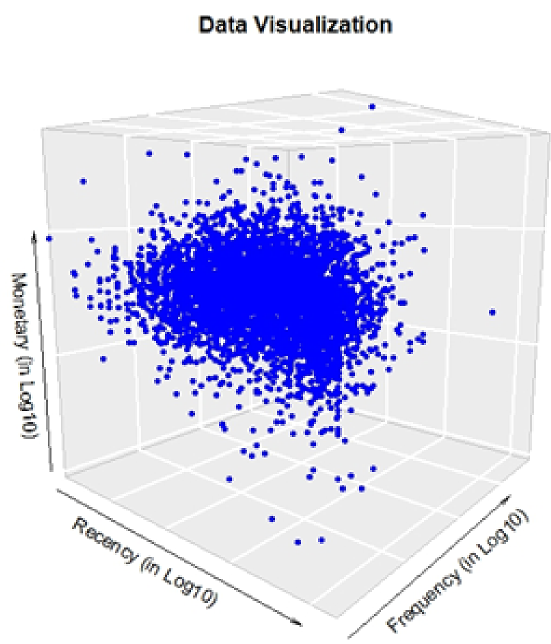

Figure 3. The data distribution of RFM after normalization

Table 1. The average RFM value of all data

R	F	M
93.05	4.27	407.96

Table 2. The average RFM value of each segment

Segment	R	F	M
1	53.61	4.07	1,274.11
2	14.63	17.30	389.31
3	295.01	1.34	302.66
4	158.73	2.06	290.23
5	36.80	3.00	267.14

Table 3. Customer criteria based on RFM indicators

Segment	Total	RFM Indicator	Customer Criteria
1	512	R↑F ↓M↑	Main
2	466	R↑F ↑M↓	Potential
3	2,037	R↑F ↓M↓	General
4	702	R↑F ↓M↓	Minimum
5	620	R↑F ↓M↓	Prospective

Table 4. Example of customer segmentation based on RFM and PAM methods

Customer ID	Recency	frequency	Monetary	Segment
18102	1	60	9,639	1
⋮	⋮	⋮	⋮	⋮
16775	11	11	232	2
14935	298	1	1,785	3
15146	165	2	834	4
14753	88	1	563	5

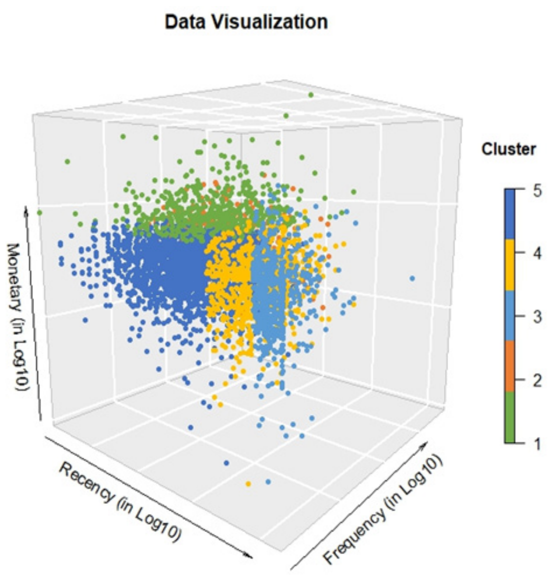

Figure 4. Customer segmentation based on the RFM model and the PAM method

Key customers are included in Segment 1. They are the company's valuable assets based on the RFM model. Although their purchase frequency is directly proportional to the average spending of other customers, their recency value and shopping transactions are grander than others. Therefore, the right marketing strategy for customers in this segment is to focus on maintaining customer engagement and increasing their purchasing frequency. Strategies that can be implemented include: (1) Increase promotional value. Customers in this segment are usually willing to pay more for the products they want. Therefore, the company can increase the value of promotions and discounts to attract customers to shop more frequently. The company may offer additional packages or discounts for large purchases. (2) Offer loyalty programs. Loyalty programs can help to increase customer shopping frequency by providing various benefits, such as points that can be exchanged for products or services, special discounts, cash-back, or access to premium services to encourage further purchases. (3) Increase personalization. Marketers can increase the personalization of company marketing communications for customers in this segment. It can be done by sending exclusive messages and offers that are relevant to customers' preferences and needs based on their shopping history.

Meanwhile, Segment 2 is a potential customer based on their recency value and shopping frequency. The proportion of shopping transactions in this segment is quite high, and among the most frequent, but they are not yet willing to pay more. The right marketing strategy for customers in this segment can be focused on increasing the customer's monetary value. This strategy can be done by: (1) Inform customers about the benefits and urgency of more expensive products. Customers may not realize the benefits of more expensive products. Therefore, companies can provide information to customers about these benefits so that they can see the added value of the product. The company can also offer packages of complementary products at special prices. (2) Offers a loyalty program that provides benefits based on transaction value. Loyalty programs can help to increase the customer's monetary value by providing benefits, such as points or cash-back for each transaction that can be exchanged for more expensive products, special discounts for loyal customers, or access to exclusive services for customers who frequently shop for special offers or products not available to other customers. Time-limited promotions with additional discounts or bonuses for purchases within a certain period may also be applied to customers. (3) Increase the personalization of marketing communications. The company can increase the personalization of its marketing communications for these customers. It can be done by sending messages that are relevant to preferences and needs as well as offering specific offers and promotions to customers. If analyzed further, the number of primary and potential customers is 978, or 22.5% of the total customers. It shows that these two customer segments have the potential to generate 80% of the company's revenue based on the 80/20 rule (Pareto principle) (Tanabe, 2018). Therefore, the management must focus on developing specific marketing strategies to retain customers in both segments.

The next segment is Segment 3, or general customers. Customers in this segment do not often make transactions or shop. This customer segment also has the lowest average RFM value compared to other segments. Customers in this segment have not shown a high interest in the product or satisfaction with the company. These customers are potential customers to be converted into more valuable customers. Therefore, the right marketing approach for these customers is to focus on increasing customer interest, engagement, and experience. The strategy can be carried out in the following way: (1) Deliver relevant education and tips about the product. Customers may not be aware of the benefits of a company's products or services. Therefore, the company can provide product usage guides and useful tips to customers so that they can see the added value of the product and increase customer engagement. (2) Offer limited offers and promotions. Attractive offers and sales promotions with special prices or bonuses for purchases within a certain period can attract customers' attention and encourage them to try the company's products or services. Surprises with limited offers can also encourage spontaneous purchases. (3) Increase the personalization of marketing communications. The company can increase the personalization of its marketing communications for these customers. It may be accomplished by involving customers in discussions or providing feedback regarding a product. The company can build a sense of exclusivity to increase customer engagement by providing rewards or small incentives as appreciation for customer participation. (4) Offers a simple loyalty program. The company can offer easy-to-understand loyalty programs with rewards or discounts every few purchases and provide incentives to increase customer purchase frequency.

Meanwhile, the minimum customers are in Segment 4. This customer segment is identical to the general customer segment. However, this segment has slightly better recency and frequency values than the average transaction value proportion. Therefore, the appropriate marketing strategy for this customer segment is identical to Segment 3.

The last segment is prospective customers, who are included in Segment 5. This segment has moderate shopping activity and purchasing power but with a high shopping recency value. The company can develop focused marketing strategies to maintain interest and increase shopping frequency. Customers in this segment have the potential to become potential customers and even main customers. Therefore, marketing strategies that can be used for this customer segment are (1) Strategies to maintain customers' interest. Enhancing consumer experience, providing special perks through loyalty programs, and personalizing marketing messaging are some ways to carry it out. (2) Strategies to boost shopping frequency. Offering alluring discounts and offers, personalizing marketing messages more, and providing loyalty programs with advantages based on transaction value are some ways to go about it. Furthermore, the internal validation results of the customer segmentation show that the minimum CI value is 0.1429 (close to zero) and the maximum CHI value is 512.9553. These results indicate that the intra-segment and inter-segment variance is high. In other words, the quality of the resulting customer segmentation is good (ideal).

Managerial Implications

This study analyzes customer segments that have similar shopping patterns. This is done to assist company management in identifying and focusing on the most profitable customer segments (based on the Pareto principle). In addition, the findings and interpretations of this study can guide company management to: (1) respond and adapt to the dynamics of changing customer preferences and shopping patterns in real-time; (2) understand customers and provide a personalized experience by conveying messages, recommendations and offers according to the preferences of each customer segment; (3) increase the effectiveness of product promotions by implementing various strategies according to customer segments; (4) increase resources efficiency in the form of time, cost and effort; (5) increase sales and the company's competitive advantage; (6)

growing customer loyalty; and (7) identify potential prospects, new customers, and customers who will be lost.

CONCLUSIONS AND RECOMMENDATIONS

Conclusions

Customer segmentation in business is very important. Companies must be able to identify customers who have a higher priority than others. One of the analysis models used in customer segmentation is RFM (Recency, Frequency and Monetary). This model considers when the customer last time made a purchase transaction (Recency), how often the transaction was made (Frequency), and how much the customer spends (Monetary). This study introduces customer segmentation for digital marketing based on the RFM model and the PAM (Partitioning Around Medoids) method. Initially, the dataset was pre-processed before being analyzed based on RFM. Furthermore, the dataset is segmented using the PAM method. The result consists of five customer segments: main, potential, minimum, general, and prospective customer. Then, the segmentation results are validated using the C-Index and the Calinski-Harabasz Index. Based on the internal validation results, it is concluded that the quality of the segmentation results is good or ideal. Thus, the results of this customer segmentation can assist marketers improve the company's business growth by implementing appropriate marketing strategies for each customer segment.

Recommendations

This study only has limited access to a customer transaction dataset and does not consider other variables, such as location, company scale, psychographics, and number of personnel. Therefore, this study cannot identify more specific segments. However, by considering a combination of these variables, marketers can develop a more diversified segmentation strategy for products positioning and target marketing. Furthermore, the dataset used is customer transaction data for two years. Using data for a longer period, deeper insights into customer shopping patterns can be obtained. Future studies could also be directed towards product segmentation. This product segmentation can be correlated with customer segmentation to obtained a shopping list for each customer segment. This can optimize the marketing of certain products for customers in the same segment. The study is very important to see the opportunities for other customers to purchase similar products.

REFERENCES

Abbasimehr H, Bahrini A. 2022. An analytical framework based on the recency, frequency, and monetary model and time series clustering techniques for dynamic segmentation. *Expert Systems with Applications* 192: 1–11. https://doi.org/10.1016/j.eswa.2021.116373

Aryuni M, Madyatmadja ED, Miranda E. 2018. Customer segmentation in xyz bank using k-means and k-medoids clustering, in 2018 International Conference on Information Management and Technology (ICIMTech), Jakarta, Indonesia, 2018, pp. 412-416, https://doi.org/10.1109/ICIMTech.2018.8528086.

Bezdek JC et al. 2016. The generalized c index for internal fuzzy cluster validity, *IEEE Transactions on Fuzzy Systems* 24(6): 1500–1512. https://doi.org/10.1109/TFUZZ.2016.2540063

Charrad M et al. 2014. NbClust: an r package for determining the relevant number of clusters in a data set. *Journal of Statistical Software* 61(6): 1–36. https://doi.org/10.18637/jss.v061.i06

Christy AJ et al. 2021. RFM ranking – An effective approach to customer segmentation. *Journal of King Saud University - Computer and Information Sciences* 33(10): 1251–1257. https://doi.org/10.1016/j.jksuci.2018.09.004

Daniel F. 2019. *Customer segmentation*, *Kaggle Datasets*. https://www.kaggle.com/code/fabiendaniel/customer-segmentation/notebook. [15 June 2023].

Gustriansyah R, Ermatita E, Rini DP. 2022. An approach for sales forecasting. *Expert Systems with Applications* 207: 1–8. https://doi.org/10.1016/j.eswa.2022.118043

Gustriansyah R, Suhandi N, Antony F. 2020. Clustering optimization in RFM analysis based on k-means. *Indonesian Journal of Electrical Engineering and Computer Science* 18(1): 470–477. https://doi.org/10.11591/ijeecs.v18.i1.pp470-477

Handojo A et al. 2023. A multi layer recency frequency monetary method for customer priority segmentation in online transaction. *Cogent Engineering* 10(1): 1–18. https://doi.org/10.1080/

23311916.2022.2162679

He Y, Cheng Y. 2021. Customer Segmentation and Management of Online Shops Based on RFM Model, in *Advances in Intelligent Systems and Computing*. pp. 34–41. https://doi.org/10.1007/978-3-030-51431-0_6

Joung J, Kim H. 2023. Interpretable machine learning-based approach for customer segmentation for new product development from online product reviews. *International Journal of Information Management* 70: 1–12. https://doi.org/10.1016/j.ijinfomgt.2023.102641

Kassambara A, Mundt F. 2020. *factoextra: Extract and Visualize the Results of Multivariate Data Analyses*, *cran.r-project*. https://cran.r-project.org/web/packages/factoextra/index.html. [15 July 2023].

Lee Z-J et al. 2021. Clustering and Classification Based on Distributed Automatic Feature Engineering for Customer Segmentation. *Symmetry* 13(9): 1–14. https://doi.org/10.3390/sym13091557

Li Y et al. 2021. Customer segmentation using K-means clustering and the adaptive particle swarm optimization algorithm. *Applied Soft Computing* 113: 1–22. https://doi.org/10.1016/j.asoc.2021.107924

Monalisa S et al. 2023. Customer segmentation with RFM models and demographic variable using DBSCAN algorithm. *TELKOMNIKA (Telecommunication Computing Electronics and Control)* 21(4): 742–749. https://doi.org/10.12928/telkomnika.v21i4.22759

Monalisa S, Nadya P, Novita R. 2019. Analysis for customer lifetime value categorization with RFM model. *Procedia Computer Science* 161: 834–840. doi: 10.1016/j.procs.2019.11.190.

Pierpaolo D et al. 2022. Community mobility in the European regions during COVID-19 pandemic: A partitioning around medoids with noise cluster based on space–time autoregressive models. *Spatial Statistics* 49: 1–22. https://doi.org/10.1016/j.spasta.2021.100531.

Rahim MA et al. 2021. RFM-based repurchase behavior for customer classification and segmentation, *Journal of Retailing and Consumer Services* 61: 1–9. https://doi.org/10.1016/j.jretconser.2021.102566.

Riza M, Seminar KB, Maulana A. 2018. Pembentukan target pasar berdasarkan data stream transaksi kartu kredit (clustering dan association rule) pada PT Bank Bukopin. *Jurnal Aplikasi Bisnis dan Manajemen* 4(1): 86–95. https://doi.org/10.17358/jabm.4.1.86.

Salminen J et al. 2023. How can algorithms help in segmenting users and customers? A systematic review and research agenda for algorithmic customer segmentation. *Journal of Marketing Analytics* 1:1–16. https://doi.org/10.1057/s41270-023-00235-5.

Schubert E, Rousseeuw PJ. 2021. Fast and eager k-medoids clustering: o(k) runtime improvement of the PAM, CLARA, and CLARANS Algorithms. *Information Systems* 101: 1–19. https://doi.org/10.1016/j.is.2021.101804.

Serper N, Şen E, Çalış Uslu B. 2022. Discrete event simulation model performed with data analytics for a call center optimization. *Istanbul Business Research* 51(1): 189–208. https://doi.org/10.26650/ibr.2022.51.951646.

Sun Y, Liu H, Gao Y. 2023. Research on customer lifetime value based on machine learning algorithms and customer relationship management analysis model. *Heliyon* 9(2): 1–16.https://doi.org/10.1016/j.heliyon.2023.e13384.

Tabianan K, Velu S, Ravi V. 2022. K-Means clustering approach for intelligent customer segmentation using customer purchase behavior data. *Sustainability* 14(12): 1–15. https://doi.org/10.3390/su14127243.

Tanabe K. 2018. Pareto's 80/20 rule and the Gaussian distribution. *Physica A: Statistical Mechanics and its Applications* 510: 635–640. https://doi.org/10.1016/j.physa.2018.07.023.

Verma S et al. 2021. Artificial intelligence in marketing: Systematic review and future research direction. *International Journal of Information Management Data Insights* 1(1): 1–8. https://doi.org/10.1016/j.jjimei.2020.100002.

Wang C. 2022. Efficient customer segmentation in digital marketing using deep learning with swarm intelligence approach. *Information Processing & Management* 59(6): 1–13. https://doi.org/10.1016/j.ipm.2022.103085.

Wong E, Wei Y. 2018. Customer online shopping experience data analytics. *International Journal of Retail & Distribution Management* 46(4): 406–420. https://doi.org/10.1108/IJRDM-06-2017-0130.

Zhang J, Lin P, Simeone A. 2022. Information mining of customers preferences for product specifications determination using big sales data. *Procedia CIRP* 109: 101–106. https://doi.org/10.1016/j.procir.2022.05.221.