

Implementation Of Data Mining In Digital Marketing Of Knit Bag Msmes West Bandung District

Yusep Saepullo¹, Moch Mumin Amirullo², Christina Juliane³

LIKMI College of Information and Computer Management, Bandung, Indonesia

e-mail: yusepa165@gmail.com

ABSTRACT

Knitted Bag Micro, Small and Medium Enterprises MSMEs in West Bandung Regency face challenges in exploiting the potential of the digital market. This research aims to apply data mining techniques to understand consumer patterns and increase the effectiveness of digital marketing strategies. In the methods section, the approach used in the clustering analysis is explained in detail. The data used, the variables observed, as well as the clustering techniques and statistical tests applied are described. In addition, the data processing and analysis procedures carried out are also explained to provide a clear understanding of the research methodology. This section presents the results of the clustering analysis and statistical tests performed. This includes the results of the clustering process using K-Means, the resulting cluster centers, as well as the results of the variance test showing the differences in means between cluster groups. Patterns in the data, differences between cluster groups, as well as the potential use of analysis results for developing marketing strategies and customer management are discussed in depth. This research succeeded in grouping customer data into five different clusters based on observed variables. Clustering analysis has helped in understanding patterns in data and identifying different customer groups. The implications of these findings for customer management and marketing strategy development are discussed, and suggestions for further research and development are provided.

Keywords: Data Mining, Digital Marketing, MSMEs, Knitted Bags, Consumer Segmentation, West Bandung Regency.

INTRODUCTION

Digitalization has turned data into something of value for an organization. The wealth of data, known as big data, contains valuable information that can be used for various purposes. Inseparable from the business world, data has become a valuable asset for companies and businesses to run their business processes as not only a tool for data analysis, but also a key to unlocking new opportunities and improving your company and its performance. By using data mining effectively, companies can improve customer understanding, increasing the ROI (Return On Investment) of marketing campaigns. One part of the business world that is greatly impacted by the use of internal data is Digital Marketing.

This has brought many fundamental changes to business processes within companies, especially in the marketing aspect. The emergence of the Internet and various e-commerce platforms and social media has also influenced business and marketing life today. Digitalization has created new opportunities for businesses to reach consumers that were previously impossible for them achieved but with digital marketing a brand can develop rapidly and quickly by using the use of data in Business Intelligence in their company. By using digital marketing effectively, businesses can increase brand awareness, reach a wider audience, increase conversions and sales, save marketing costs, and can increase adaptability and innovation. Digital marketing has become a vital instrument in today's business world.

MSMEs have a vital role in national economic development. In monetary times,

MSMEs are more resilient in facing crises. Especially since the economic crisis occurred, the role of small, medium and micro businesses can be said to be a safety net for the national economy. West Regency Knitted Bag MSMEs, as part of the creative sector, face challenges in competing in a wider market and utilizing existing information technology. The use of data mining in digital marketing can provide competitive advantages by understanding consumption patterns, customer preferences, and optimizing marketing campaigns. However, implementation of this technology is still limited among MSMEs due to lack of resources, technical knowledge, and access to effective analytical tools. MSMEs can identify market trends, consumer behavior and the effectiveness of the marketing strategies used. This is especially important in West Regency, where the knit bag industry has great potential for product development and market expansion. However, the application of data mining to optimize digital marketing in this sector has not been widely explored, which shows that there is a gap that needs further research.

LITERATURE REVIEW

Data Mining

Data mining is a field of science for handling information retrieval from large databases, which combines various techniques such as statistics, machine learning, data visualization, and others (Werdiningsih, 2020).

Data Mining is a powerful tool that can help businesses make better decisions, increase efficiency, increase customer retention, increase sales, and improve security (Fayez Yousif, Zaki Mohamed, 2016). One of the roles of data mining in the business world is to increase efficiency, especially in the field of promotion, so that campaigns built by companies can run more effectively and efficiently with lower promotional costs.

Based on the definition above, data mining has a large function and role in technological progress and digitalization, as in the business world today, many people use data mining as a tool to carry out analysis in new product or brand campaigns that will be sold to the general market. Overall, data mining helps businesses improve decision making, increase efficiency, increase customer loyalty, increase sales, and increase security.

Digital Marketing

In Marketing, we are familiar with the marketing mix process which consists of the classic marketing mix, namely Product, Price, Place, Promotion and the modern marketing mix of Product, Price, Place, Promotion, People, Process and Physical Evidence. Promotion cannot be eliminated in a marketing process because it is one of the important factors in marketing itself. In classic marketing, advertising activities are often only carried out through print or audio-visual media, but developments in digital technology are influencing the way advertising is carried out. The process is often called digital marketing. Digital marketing itself includes the use of devices connected to the internet to carry out marketing through various digital strategies and media, with the aim of keeping businesses connected with consumers through online media channels. (Andy Gunawan Chakti, 2019), According to the American Marketing Association (AMA), digital marketing is an activity, institution and process that is supported by digital technology. The goal is to create, communicate, deliver, and provide value to consumers and other interested parties.

(Kannan & Hong Sang, 2016). In its application, digital marketing uses the internet and social media as a platform or tool for product promotion.

Social media is one of the digital marketing tools that is easy to use, especially in Indonesia, Micro, Small and Medium Enterprises tend to market their products through social media. This helps boost business growth and the sustainability of the business itself. Because the costs are quite cheap and no specific skills are required to start, social media is considered to have the ability to attract potential consumers directly. Based on the characteristics of social media relationships, it is divided into two parts, namely profile-based, namely social media which focuses on individual members, this group encourages relationships that occur because an individual is interested in the personality of social network users, for example Facebook, X, WhatsApp. Content-based, namely social media focuses on the content of the content, discussions and comments on the content displayed. The main goal is to connect individuals with content provided by creators because these individuals like that content, for example YouTube, Instagram, Pinterest.

Marketing through digital marketing will have a wider reach and cheaper costs. The existence of social media is a medium that consumers can use to disseminate information in the form of text, images, audio and video with many parties, both between business actors and with consumers or customers within the company (Kotler, 2012). Digital marketing can reach all groups, anytime, anywhere and in any way. Of course, it is far superior to conventional marketing which is limited in time, place and ability to reach users because it only promotes products directly in stores.

RESEARCH METHODS

The subject of this research is the Implementation of Data Mining in Digital Marketing for MSMEs with the case study taken, namely Knitted Bag MSMEs in West Bandung Regency. This is done to determine the level of effectiveness and efficiency of Digital Marketing with the help of Data Mining so that it has a good impact on the business processes carried out by existing MSMEs.

The method used in this research uses the Clustering Method, with a Partitional Clustering Approach. Clustering is a data mining technique that groups data objects into clusters in such a way that the data belonging to each cluster is different. Grouping is an analysis method that groups data objects into the data groups required for each group. This grouping can be classified as clustering hierarchical. The most commonly used clustering method is K-Means. The K-Mean algorithm is easy to implement, relatively quick to run, easy to complete, and generally applicable in practice. K-means is used to develop decision support systems that improve safety using a multi-step process. Therefore, cluster analysis usually provides data in clusters with high expected values, and data in other clusters with low expected values.

The data set processed in this research is District Knitted Bag MSME Consumer Data West Bandung in the 2019 – 2023 Time Period with the Data Set Attributes used, including:

1. Customer Data

Table 1. Customer Data Attributes

No	Data Name	Data Attributes	Period
1	Customer Data	1.1. Customer ID	2019 - 2023
		1.2. Customer Name	
		1.3. Gender	
		1.4. Age	
		1.5. Address	
		1.6. Tel	
		1.7. E-mail	
		1.8. Facebook	
		1.9. Instagram	
		1.10. TikTok	
		1.11. X	

2. Product Data

Table 2. Product Data Attributes

No	Data Name	Data Attributes	Period
2	Product Data	2.1. Product Code	2019 - 2023
		2.2. Product name	
		2.3. Color	
		2.4. Size/Size	
		2.5. Quantity	

3. Transaction Data

Table 3. Transaction Data Attributes

No	Data Name	Data Attributes	Period
3	Transaction Data	3.1. Order Number	2019 - 2023
		3.2. Order Date	
		3.3. Order Item ID	
		3.4. Payroll Amount	
		3.5. OrderStatus	
		3.6. Statement Numbers	
		3.7. Payout Status	

4. Product Insights Data

Table 4. Product Insight Data Attributes

No	Data Name	Data Attributes	Period
		4.1. Date	
		4.2. Product Visitors (Visits)	
		4.3. Product Page Viewed	
		4.4. Products Visited	
		4.5. Visitors View Without Buying	
		4.6. Visitor Viewing Rate Without Purchasing	
		4.7. Click Search	
		4.8. Like	

No	Data Name	Data Attributes	Period
4	DataProduct Insights	4.9. Product Visitors (Adding Products to Cart)	2019 - 2023
		4.10. Added to Cart (Product)	
		4.11. Product Added to Cart Conversion Rate	
		4.12. Total Buyers (Orders Created)	
		4.13. Products (Order Created)	
		4.14. Ordered Products	
		4.15. Total Sales (Orders Created) (IDR)	
		4.16. Conversion Rate (Orders Created)	
		4.17. Total Buyers (Orders Ready to Ship)	
		4.18. Products (Orders Ready to Ship)	
		4.19. Products Ready to Ship	
		4.20. Sales (Orders Ready to Ship) (IDR)	
		4.21. Conversion Rate (Orders Ready to Ship)	
		4.22. Conversion Rate (Orders Ready to Ship divided by Orders Created)	

5. Visit Origin Data

Table 5. Data attributes of origin of visit

No	Data Name	Data Attributes	Period
5	Visit Origin Data	5.1. Website	2019 - 2023
		5.2. TikTok	
		5.3. Facebook	
		5.4. Instagram	
		5.5. X (Twitter)	
		5.6. WA Business	

Based on the attributes in the Data Set above, it is then processed using K-Means Clustering using the help of statistical applications. The K-means algorithm is very easy to implement especially when implementing very large data sets, and can also overcome the difficulties of the required data. The basic process of K-means clustering consists of six steps. The first step is initialization. That is, determine the value of k as the desired number of clusters and the desired metric (distance). For example, if you want to group into 5 clusters then $k = 5$. The second step is to determine the cluster center randomly. Then, in the third step, all data reaches the cluster center mathematically, using the formula below:

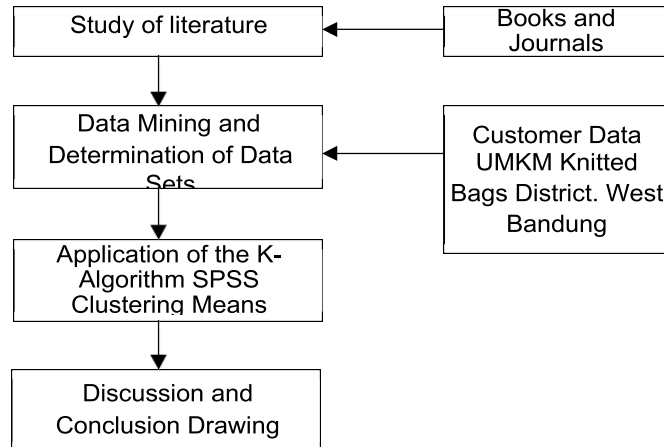
$$d_{ik} = \sqrt{\sum_j^m (C_{ij} - C_{kj})^2}$$

The fourth step groups the data into clusters with the shortest distance and calculates the new cluster center, using the following formula:

$$C_{kj} = \frac{\sum_{i=1}^p X_{ij}}{p}$$

The final step is step 6. Repeat steps 2 to 4 to ensure no more data is moved to another cluster. The sample data used was 60 samples, the date range used in the research This has several implementation stages, including:

Figure 1. Research Stages



The analysis carried out is clustering analysis, which is one of the methods in multivariate data analysis. Clustering, or clustering, is a clever technique in the world of statistics that helps us understand data better. Imagine you have a large pile of data, such as customer information or survey results. Clustering allows you to group this data automatically based on what they have in common, like sorting clothes by color or type. In this context, data is grouped into clusters that have similarities based on certain observed features. Specifically, the method used in this analysis is the application of the SPSS K-Means Clustering Algorithm. The K-Means algorithm is one of the most commonly used clustering methods. This method groups data into K clusters, where K is a predetermined number of clusters. The K-Means process involves selecting starting points (initial cluster centers) randomly or based on initial information, grouping data into clusters based on proximity to the starting point, recalculating cluster centers based on the average of each cluster, and iterating this process until convergence is achieved. Apart from K-Means, there is also analysis of variance carried out to assess the differences in means between cluster groups for each observed variable. These two methods, K-Means Clustering and ANOVA, are used together to identify patterns in data and gain a better understanding of the groups that form based on observed characteristics.

RESULTS AND DISCUSSION

Initial cluster centers for five clusters formed from various indicators. Each row in the table shows the average value of each indicator for each cluster.

Table 6. Initial Cluster Centers

	Clusters				
	1	2	3	4	5
Conversion_Rate	1.64	3.16	3.67	.00	5.60
Income	5,224	3,745	2,335	.00	5,695
Visitors	61.00	158.00	109.00	132.00	125.00
Buyer	1.00	5.00	4.00	.00	7.00
Order	1.00	5.00	4.00	.00	9.00
Show_Page	184.00	317.00	271.00	293.00	396.00
Unit_Sold	1.00	5.00	4.00	.00	10.00

	Clusters				
	1	2	3	4	5
Revenue_Per_Buyer	5,224	7,484	5,824	.00	8.134
Visitor_Value	855.85	2368.39	2134.82	.00	4553.47
User_Add_To_Cart	5.00	12.00	5.00	10.00	5.00
Units_Added_To_Cart	5.00	18.00	5.00	12.00	6.00
Wishlists	1.00	6.00	4.00	1.00	11.00
User_Wishlists	1.00	6.00	4.00	1.00	8.00
Revenue_Per_Order	5,224	7,484	5,824	.00	8.134
Number_of_Products_Per_Order	1.00	1.00	1.00	.00	1.11
Number_of Cancellations	1,895	.00	4,034	.00	.00
Return_or_Refund_Cost	.00	.00	5,274	.00	.00

Source: data processed by SPSS.

The table above is a representation of the initial cluster centers resulting from the clustering process. This table consists of 5 clusters, each of which has an average value for each indicator. Each row represents one indicator, while each column represents one cluster.

Cluster 1 has a fairly low average value for the indicators Conversion Rate, Revenue, Buyers, Orders, and Return Costs, but has moderate values for the indicators Pages Viewed, Units Sold, Revenue per Buyer, Users Add to Cart, and Wishlists User. Cluster 2 has a high average value for almost all indicators, except for the number of cancellations and low return costs. Cluster 3 has relatively low average values for most indicators, but high for Visitor Value and Number of Cancellations. Cluster 4 has zero values for several indicators such as Revenue, Buyers, Orders, and Return Costs, but has quite high values for the Units Added to Cart and Wishlists indicators. Cluster 5 has high average values for most indicators, with the highest values for the Conversion Rate, Revenue, Page Views and Visitor Value indicators.

The table above provides an overview of the characteristics of each cluster in terms of the average values of the various indicators used in the clustering process. By understanding the distribution of these values, we can understand how the data has been grouped into different clusters and understand the characteristics of each of these clusters.

Table 7. Iteration History

Iteration	Change in Cluster Centers				
	1	2	3	4	5
1	5.7044	3.2744	7.6904	8.9874	5.5794
2	,000	2.6254	8.2003	8.9573	,000
3	,000	,000	4.0603	1.7193	,000
4	,000	,000	,000	,000	,000

Source: data processed by SPSS.

Iteration History shows the journey of iteration when the clustering process was carried out. At each iteration, changes in cluster centers are measured to see whether convergence occurs. In this case, in the first iteration, there was a significant change in cluster centers, with the largest change being 8.9874. However, as the iteration

continues, these changes become smaller. In the second iteration, there was a decrease in changes in the cluster center, and in the third iteration, the changes became smaller until they reached zero in the fourth iteration. Convergence is achieved due to no or very small changes in cluster centers. The minimum distance between the initial centers of the cluster is 158050.471, indicating that the initial points are located quite far from each other. This shows that the clustering process has succeeded in achieving convergence with cluster centers that are stable and have not experienced significant changes.

Table 8. Final Cluster Centers

	Clusters				
	1	2	3	4	5
Conversion_Rate	2.08	3.62	3.29	1.23	5.58
Income	8,054	3,175	1,965	6,964	5,375
Visitors	70.00	127.00	103.47	102.21	107.50
Buyer	1.50	4.40	3.24	1.21	6.00
Order	1.50	4.60	3.29	1.21	7.50
Show_Page	178.00	298.20	274.88	244.82	337.50
Unit_Sold	1.50	5.40	3.53	1.21	8.00
Revenue_Per_Buyer	5,334	7,264	6,264	4,974	9,114
Visitor_Value	1115.99	2626.44	1983.58	711.73	5081.33
User_Add_To_Cart	5.50	7.80	8.24	7.62	5.00
Units_Added_To_Cart	6.50	11.20	9.71	8.94	5.50
Wishlists	1.50	5.20	3.59	1.44	7.50
User_Wishlists	1.50	4.80	3.53	1.44	6.00
Revenue_Per_Order	5,334	7,264	6,264	4,974	9,114
Number_of_Products_Per_Order	1.00	1.19	1.11	.85	1.06
Number_of Cancellations	1,405	1,024	5820.29	1626.32	4,344
Return_or_Refund_Cost	.00	.00	6196.24	.00	.00

Source: data processed by SPSS.

The Final Cluster Centers table describes the final cluster centers after the clustering process is complete. Each row in the table represents the average value of the various features or variables observed for each cluster formed. There are five clusters identified. Cluster 1 has a cluster center with a different value from the other clusters for each variable. The same applies to other variables such as Revenue, Visitors, and so on.

Detailed analysis of the table shows that each cluster has different characteristics in terms of observed variables. Cluster 1 has lower values for the Income and Visitor Value variables compared to other clusters, while Cluster 5 has higher values for these variables. The number of products per order also varies between clusters, with cluster 3 having a higher number of products per order compared to the others.

This table provides valuable insight into how the data is divided into different groups based on observed characteristics. This can help in further understanding of patterns in the data, as well as in better decision making.

Table 9. Distances between Final Cluster Centers

Clusters	1	2	3	4	5
1		2.7145	1.7775	1.3875	4.6975
2	2.7145		1.2225	2.5005	2.2365
3	1.7775	1.2225		1.2815	3.4535
4	1.3875	2.5005	1.2815		4.7295
5	4.6975	2.2365	3.4535	4.7295	

Source: data processed by SPSS.

The Distances between Final Cluster Centers table shows the distance between final cluster centers. Each cell in the table shows the Euclidean distance between two cluster centers. In this case, five clusters are identified, and the distance between each pair of cluster centers is measured.

The distance between the centers of clusters 1 and 2 is 2.7145, 2 and 3 is 1.7775, 3 and 4 is 1.3875, 4 and 5 4.6975 which shows how far the two clusters are separated in the observed feature space. The greater the distance, the more different the characteristics between the two clusters. From the table, we can also see that the distance between cluster centers in one group is usually smaller than the distance between cluster centers of different groups. This shows that clusters that have more similar characteristics tend to be closer to each other in the observed feature space.

Table 10. ANOVA

	Clusters		Error		F	Sig.
	Mean Square	df	Mean Square	df		
Conversion_Rate	21,328	4	,692	55	30,841	,000
Income	1.75411	4	1.4109	55	124,375	,000
Visitors	1272,860	4	811.424	55	1,569	,006
Buyer	26,466	4	,515	55	51,404	,000
Order	34,828	4	,533	55	65,403	,000
Show_Page	11060.911	4	4619.891	55	2,394	,061
Unit_Sold	44,839	4	,536	55	83,615	,000
Revenue_Per_Buyer	1.4859	4	3.5188	55	4,220	,005
Visitor_Value	1.4177	4	220649.525	55	64,216	,000
User_Add_To_Cart	7,315	4	12,407	55	,590	,002
Units_Added_To_Cart	17,110	4	19,658	55	,870	,008
Wishlists	36,608	4	1,024	55	35,763	,000
User_Wishlists	27,254	4	,762	55	35,760	,000
Revenue_Per_Order	1.4859	4	3.5188	55	4,220	,005
Number_of_Products_Per_Order	,261	4	.102	55	2,563	,048
Number_of_Cancellations	9.6469	4	3.3148	55	29,112	,000
Return_or_Refund_Cost	1.1698	4	8.9007	55	1,314	,006

Clusters have been chosen to maximize differences between cases in different clusters. The observed significance level is not corrected for this in that the cluster means are the same if $\text{sig} < 0.05$.

Source: data processed by SPSS.

The table above shows the results of variance analysis for each variable in each cluster which is significant at the 0.05 significance level. ANOVA is used to test differences in means between several groups. This table shows the F-statistic value and

p-value (Sig.) for each variable in different cluster groups.

A variable that is significant at the 0.05 significance level indicates that there is a significant mean difference between cluster groups for that variable. For the Conversion Rate variable, the F-statistic value is 30.841 with a p-value of less than 0.05, indicating that there is a significant difference in Conversion Rate between each cluster. The same applies to other variables such as Income, Buyers, and so on. However, some variables may not be significant at the 0.05 significance level. For the Viewed Pages variable, the p-value is 0.061, which is greater than 0.05, so there is not enough evidence to conclude that there is a significant difference in Viewed Pages between cluster groups.

Table 11. Number of Cases in Each Cluster

Clusters	1	2,000
	2	5,000
	3	17,000
	4	34,000
	5	2,000
Valid		60,000
Missing		.000

Source: data processed by SPSS.

The "Number of Cases in Each Cluster" table shows the number of cases or data included in each identified cluster. Each row in the table represents a particular cluster, and the number next to that cluster indicates the number of cases belonging to that cluster. There are 17 cases included in Cluster 3, while Cluster 4 has 34 cases. The total number of valid cases is 60, which is the total of all cases distributed into different cluster groups. Distribution of data within each cluster and can help in evaluating the homogeneity or heterogeneity of each cluster group. This can also provide an idea of the proportion of data that falls into each group, which can be used for further analysis or further understanding of patterns in the data.

The results of the clustering analysis show a series of tables that provide deep insight into the data grouping process. First, the Initial Cluster Centers table describes the initial centers of each cluster before the clustering process begins, which functions as a starting point in placing data into the appropriate clusters. Next, Iteration History shows how cluster centers change over the course of iteration, with convergence achieved when the change in cluster centers decreases over time. Next, Final Cluster Centers displays the final results of the clustering process, with each row representing the average value of the observed variables for each cluster. Distances between Final Cluster Centers provides an idea of the distance between final cluster centers, helping to understand how close or far each cluster is from the others in feature space. Then, the ANOVA table sig < 0.05 provides the results of analysis of variance for each variable in each cluster which shows significant mean differences between cluster groups. Finally, the Number of Cases in Each Cluster table describes the distribution of data in each cluster. These analyzes together provide a comprehensive understanding of the clustering process and the characteristics of each resulting cluster. By understanding the results of each of these tables, we can more deeply understand the patterns in the data and better identify different groups.

CONCLUSIONS AND RECOMMENDATIONS

From the results of the clustering analysis carried out using the K-Means method and ANOVA test, it can be concluded that the data has been successfully grouped into five clusters which have different characteristics based on the observed variables. The cluster centers have stabilized and convergence is achieved, indicating that the clustering results are consistent and reliable. The results of the ANOVA test show that there are significant mean differences between cluster groups for most of the observed variables, indicating that clustering has succeeded in identifying groups that differ significantly in the data. From these results, several suggestions can be given for further development. First, clustering results can be used to further understand the characteristics and patterns in the data, as well as to identify customer segmentation or behavioral patterns that may be hidden. In addition, information from clustering results can be used for better decision making, such as adjusting marketing strategies or improving services based on a better understanding of different customer groups. In addition, for more advanced analysis, you can consider using more complex clustering methods or adding additional variables to gain a more holistic understanding of the data. Thus, the results of this clustering can become a strong basis for further strategy development in various contexts, from business to research.

REFERENCE

- Adiwinoto et al. (2020). The Digital Platform Promotion Strategy To Support MSMEs In Bangka Belitung Province.
- Astuti et al. (2019). Determining Promotion Strategies for Micro, Small and Medium Enterprises (MSMEs) Using the CRISP-DM Method with the K-Means Clustering Algorithm.
- Azhar et al. (2022). Sentiment Analysis of Cryptocurrency Based on Python TextBlob Using the Naïve Bayes Algorithm.
- Budi Utomo et al. (2023). Analysis of the Effectiveness of Integrated Digital Marketing Communication Strategies in Building MSMEs Brand Awareness Through Social Media.
- Faisal Nugraha et al. (2022). Application of Data Mining with Classification Methods Using the C4.5 Algorithm.
- Hasanah & Juliana (2021). Digital Marketing a Marketing Strategy for MSME Products in The Digital Era.
- Lestari & Saifuddin (2020). Implementation of Product Promotion Strategies in the Purchasing Decision Process Through Digital Marketing During the Covid'19 Pandemic.
- Qudus et al. (2021). Implementation of Digital Marketing in MSMEs in the Era of Industrial Revolution 4.0 Study in MSMEs in Sidoarjo Regency.
- Rahmatica et al. (2020). Implementation of K-Means Clustering and Weighted Products in Determining Crime-Prone Locations.

- Rozinah & Meiriki (2020). Utilization of Digital Marketing in Small and Medium Enterprises (MSMEs) in South Tangerang City.
- Saura (2021). Using Data Sciences in Digital Marketing: Framework, methods, and performance metrics.
- Sufi & Suharti (2021). The Influence of Intrinsic Motivation and ICT Knowledge on Digitalization ReadinessMSMEs with Demographic Factors as Moderating Variables Study on Culinary MSMEs in Salatiga.
- Zamani & Kamaruddin (2023). Crypto-sentiment Detection in Malay Text Using Language Models with an Attention Mechanism.