Improving Marketing Strategies
Using a Clustering and AR-MBA Methods at Indomaret Kaliurang

Rifki Nurul Mukarim, Dandy Fri Subagja, Demas Emirbuwono Basuki, Ratna Agil Apriani

ABSTRACT

Marketing is the main activity carried out by entrepreneurs to maintain the viability of their business, develop the company, and get maximum profit. Companies need to know the right and appropriate marketing strategy so that the products to be sold on the market. Marketing strategy is one of the strategies applied to the retail industry in Indonesia, one of which is Indomaret. The purpose of this research is to increase market competitiveness so that it can survive and develop, and Indomaret in Kaliurang hopes to make a strategy to increase sales by conducting an analysis of consumer segments and product categories so that they can create sales strategies that can increase sales. The method used is clustering to find potential target markets that have similar characteristics and association rules & Market basket analysis (AR-MBA) is used to find best practices for product marketing. The results of the Cluster Method with Non-Hierarchical Algorithm K-Means are 3 clusters. Based on the analysis conducted, it was found that the results of the AR-MBA method, there could be 6 rules for goods purchased simultaneously. Marketing strategies that can be applied by Indomaret Kaliurang such as promotions can be created by making products that are rarely purchased as bonuses for buying products that are often purchased. In addition, vouchers can be created as a promotional tool to encourage greater consumer purchases.

Keywords: AR-MBA, clustering, sales strategy, strategy marketing

Introduction

At this time, the retail industry is experiencing very fierce competition in increasing company sales. The retail industry continues to strive to develop innovation and find business strategies so as to increase its sales rate. Based on data from the Yogyakarta Central Bureau of Statistics, the growth rate of Gross Regional Domestic Product (GRDP) in the third quarter of 2017 for the business sector in accommodation, food and beverage providers including retail is 62% [1]. This is now a must that companies, especially the retail industry, must do in order to maintain a competitive advantage in a market that continues to change. In addition, the marketing sector in the retail industry must also be able to compete by adjusting to customer needs [2]. Marketing is the main activity carried out by entrepreneurs to maintain the viability of their business, develop the company, and get the maximum profit. According to Permata, et al (2020), marketing is a business strategy that must protect all activities in a company, including all processes, creating, offering, exchanging value from a project or a company [3]. Marketing is part of the company’s management and is also one of the most critical factors in the company’s success. This marketing activity can focus on concepts, prices, and promotions in creating a great opportunity in accordance with the goals of a company [4]. Companies need to know the right and appropriate marketing strategy for the products to be sold in the market [5]. One marketing strategy that can increase customer attraction is sales promotion. Sales promotion has a big role in increasing customer desire to buy products suddenly without any plan or usually referred to as impulse buying [6]. Sales promotion is one of the strategies applied to the retail industry in Indonesia, one of which is Indomaret.

Indomaret is a franchise retail network in Indonesia. Indomaret is one of the subsidiaries of Salim group. Indomaret is a minimarket network that includes large companies and has been established for a long time that also provides basic necessities and daily necessities with a sales area of less than 200 m². The first store was opened in Ancol, North Jakarta, in 1988, managed by PT. Indomarco Prismatama [7]. Tahun 1997 perusahaan mengembangkan bisnis gerai waralaba pertama di Indonesia, setelah memiliki lebih dari 230 gerai. Sampai dengan awal tahun 2016, jumlah gerai sebanyak 12.100 toko [8]. Indomaret Kaliurang hopes to always make strategies to increase competitiveness to survive and develop. Indomaret in Kaliurang hopes to always make strategies to increase...
sales. Indomaret must have the ability to develop standard procedures that analyze consumer segments and product categories so as to increase sales. It can provide analytical insights that can be important for making marketing decisions [9]. Therefore, Indomaret Kaliurang will conduct a customer segmentation analysis so that several variables are needed to conduct this customer segmentation analysis. Indomaret Kaliurang wants to know the character of customers in addition to customer shopping patterns in order to provide the right product promotion. In addition, pertinent questions will be asked in the questionnaire consumer shopping habits, such as what products are frequently purchased simultaneously, total transactions, payment methods, frequency of arrivals in one month, and shopping experience (satisfied or dissatisfied) as well as whether to be purchased again or not. Therefore, data mining methods are used to conduct customer analysis.

Data mining is a process to look for interesting patterns or information in selected data using certain techniques or methods. [10]. Data Mining is a data collection technique based on data obtained from several sources, and then the data will be processed into information that is very useful using various data processing methods. Very large data sets and their processing can help someone to do more detailed and clear data analysis as in the business world, business activities are similar to competition in particular, namely the minimarket business or indomaret [11]. One of the most useful data mining techniques in performing analysis for customer segmentation is clustering by grouping of customers similarity and behavior and customers grouped into homogeneous groups [12]. In addition, a technique that is often used to find out buying patterns is Market basket analysis which has a focus on identifying relationships between the products that customers buy. Understanding customer intent during a visit allows retailers to provide satisfactory service that is personalized according to their visit.

Dalam beberapa penelitian sebelumnya digunakan clustering untuk mendapatkan deskripsi dari sekumpulan data dengan cara mengungkapkan kecenderungan setiap individu data untuk berkelompok dengan individu-individu data lainnya. Research conducted by Malik, et al (2023), data mining techniques called Market Basket Analysis (MBA) and Clustering are able to contribute an important role to determine the placement of goods, and design business strategies for retailers to attract consumer attraction and hence increase business [13]. According to Kalijaga, et.al (2021), clustering dan ARMBA mampu memberikan strategi untuk dapat menarik customer dan memenangkan pasar dengan mengetahui pola pembelian barang oleh konsumen [14]. In Majid and Pramudyo's research (2021), it shows some recommendations based on the analysis of association rules include the removal of some parts of the product adjacent to the corresponding part and the use of rarely purchased products for price discount offers [15]. The K-Means algorithm is useful for data clusters, so the FP-Growth algorithm is useful for the association process so that it is able to provide product recommendations to customers more accurately [16]. Research conducted by Syahputra, et.al (2022), explained that clustering is a good choice that can be used to detect similarities in customer buying patterns to provide the right marketing strategy [17]. Market basket analysis is very helpful for business owners in improving their marketing strategy [18]. Research conducted by Qisman, et.al (2020), explained that the priori algorithm can analyze transaction data at Mizan Computer Retail Store by finding association rules that meet the minimum support and minimum confidence requirements, based on the items in the transaction [19]. Other studies also mention that the K-Means method can be used in assessing customer loyalty [20]. Therefore, in improving the marketing strategy for Indomaret Kaliurang, data mining methods are used, namely clustering and AR-MBA. Clustering is used to find potential target markets that have similar characteristics. Finally, association rules & market basket analysis (AR-MBA) were used to find the best practices for product marketing.

Research Methods

Research Subjects and Object

The subjects of this study were customers of Indomart Kaliurang without age and gender restrictions. The number of customer respondents was 100. In addition, the object studied is customer segmentation, and the types of products that customers buy. Data collection uses an open questionnaire design in the form of a gform so that it can facilitate data filling in by respondents. This study also uses several variables such as the type of product purchase, total transaction, payment method, frequency of arrival, shopping satisfaction, and the likelihood of coming back. This variable was chosen because it can help in finding customer segmentation so that researchers can provide strategy recommendations that suit customer needs. Thus, the implementation of sales strategies in Indomaret Kaliurang can increase sales.

Research Flow

The research has a flow that includes problem identification, data collection, data processing, and analysis. The following is the research flow used in this research:
Based on the research flow above, the research begins with stages in the form of problem identification, which is carried out in two ways: literature study and field study. The data was collected by distributing questionnaires to 100 respondents who made transactions at Indomaret Kaliurang. The collected data was then processed using three data mining methods: clustering and AR-MBA. Data processing results will become marketing strategies that Indomaret can implement to increase competitiveness. The final step is to conclude the research that has been carried out.

Research Methods

Clustering is an algorithm for obtaining a description of a data set by revealing the tendency of each data individual to group with other data individuals [21]. Algorithm K-means is one of the algorithms with partitional because K-means is based on determining the initial number of groups by defining their initial centroid values. K-means clustering is one of the data methods clustering non-hierarchy, which groups data in one or more forms of cluster/group [22]. Through clustering, this will be known information or knowledge from a large data set [23].

Association Rule is often known as shopping cart analysis (market basket analysis). Market basket analysis provides information on the relationship between items. This method can be used in business processes, including sales. Processed sales data includes information on data relationships with customers so that it will produce customer purchasing patterns, provide recommendations, and promote product needs market basket analysis starting from the accuracy and benefits built in the form of association rules (association rule) and from data patterns linked in the database [24]. Market basket analysis helps to analyze the relationship between all the items included in the transaction to obtain information to benefit the marketing strategy so that the opportunity to increase the sale of items simultaneously can be done [25]. Besides, Market basket analysis It can also be used to determine and predict customer behavior patterns when shopping based on the shopping habits of previous customers [26]. MBA provides information on what consumers are buying at the same time and the reasons why they buy these products. There are three steps to how an MBA works, namely determining the minimum score support and confidence, determining a frequent item set that meets the minimum frequency, and results in a frequent item set will produce association rules [27].

Results and Discussion

Clustering Analysis

According Ernawati, et.al (2020), customer segmentation divides a market into different groups. Where each group has the same or almost the same characteristics of its consumers [28]. Customer segmentation is divided into demographics, psychographics, and behavior. Customer segmentation is a necessity, and in approaching each market segment that is used as a target market, you must have a particular strategy.

Based on the Customer Satisfaction questionnaire results at Indomaret Kaliurang, which had previously been distributed to 100 respondents, data was obtained, and the completion of consumer grouping was done using a non-hierarchical method (K-Means). The indicator used in this research is knowing consumer shopping patterns to provide the correct type of product promo. The variables used are:

a. Product Purchase Frequency
   This product purchase frequency variable determines the number of products frequently purchased by respondents simultaneously in one transaction. This variable is divided into 5, namely,
   1. Basic ingredients
2. Home equipment  
3. Snack  
4. Drink  
5. Instant food.  

b. Total Transactions  
Total transactions are the total purchases in one transaction, which are then accumulated based on units of time. The amount of money consumers can measure is the nominal amount of a transaction spent to carry out a transaction. This variable is divided into six types, namely,  
1. <Rp. 49.999  
2. Rp. 50.000 - Rp. 149.999  
3. Rp. 150.000 – Rp. 249.999  
4. Rp. 250.000 – Rp. 499.999  
5. Rp. 500.000 – Rp. 999.999  
6. > Rp.1,000,000  

c. Methods of payment  
Payment method is a variable used to determine the transaction method carried out by consumers in the product purchasing process. This variable is divided into two types, namely,  
1. Cash  
2. Non Tunai (Emoney/Qiris, Debit, Kredit)  

d. Arrival Frequency  
Arrival frequency is a variable used to determine the total arrival of visitors to Indomaret locations to make purchases within one month. This variable is divided into five types, namely,  
1. 1 time  
2. 2 times  
3. 3 times  
4. 4 times  
5. > 5 times.  

e. Shopping Satisfaction  
Customer satisfaction is the key to creating customer loyalty. Apart from increasing customer loyalty, customer satisfaction can also prevent customer turnover, reduce customer sensitivity to price, reduce the costs of marketing failure, reduce operating costs caused by an increase in the number of customers, increase advertising effectiveness, and improve business reputation. This variable is divided into two types, namely,  
1. Satisfied  
2. Not Satisfied  

f. Desire to Return or Not  
desire to return or not is a variable used to determine customer loyalty to the company. This variable is divided into two types, namely,  
1. Yes  
2. No  

The following are the results of data processing from 100 respondents who made transactions at Indomaret Kaliurang:
It can be seen from the iteration table that out of 100 data, it takes nine iterations to obtain cluster exact, and the minimum value distance between the cluster values is 5.746. Next, data processing is done to find the final cluster centers.

**Table 1. Final Cluster Centers**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Cluster 1</th>
<th>Cluster 2</th>
<th>Cluster 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zscore (Purchase_Frequency)</td>
<td>-0.22137</td>
<td>-0.42304</td>
<td>1.66637</td>
</tr>
<tr>
<td>Zscore (Total_Transactions)</td>
<td>-0.60579</td>
<td>0.59078</td>
<td>0.18896</td>
</tr>
<tr>
<td>Zscore (Purchase_Method)</td>
<td>-0.64166</td>
<td>0.38759</td>
<td>0.79558</td>
</tr>
<tr>
<td>Zscore (Frequency_Arrival)</td>
<td>-0.48289</td>
<td>0.44265</td>
<td>0.22132</td>
</tr>
<tr>
<td>Zscore (Shopping_Satisfaction)</td>
<td>0.20875</td>
<td>-0.14657</td>
<td>-0.20764</td>
</tr>
<tr>
<td>Zscore (Desire to_Return)</td>
<td>0.46364</td>
<td>-0.55</td>
<td>0.1</td>
</tr>
</tbody>
</table>

**Figure 3 Output Final Cluster Centers and Anova**

**Figure 4 Output Number of Cases in Each Cluster**

This output shows the number of customers entering each cluster. It can be seen that in cluster 1 there are 44 customers; in cluster 2 there are 40 customers; and in cluster 3 there are 16 customers, so the total number of customers who fill out the Google form is 100. With the overall output of data processing using SPSS, the average calculation for each cluster criterion can be calculated. The formula used is as follows:

\[ x = \mu + z \cdot \sigma \]  \hspace{1cm} (1)

where:
- \( x \) : sample mean (inner variable cluster)
- \( \mu \) : population mean
- \( z \) : standardization value
- \( \sigma \) : standard deviation

Information on the average value of each cluster can be seen in the following table:

**Table 2. Cluster Average Results**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Cluster 1</th>
<th>Cluster 2</th>
<th>Cluster 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Purchase Frequency (t)</td>
<td>2.02</td>
<td>1.80</td>
<td>3.92</td>
</tr>
<tr>
<td>Payment Method (t)</td>
<td>2.08</td>
<td>3.01</td>
<td>2.74</td>
</tr>
<tr>
<td>Arrival Frequency (t)</td>
<td>1.88</td>
<td>3.02</td>
<td>3.39</td>
</tr>
<tr>
<td>Shopping Satisfaction (t)</td>
<td>3.60</td>
<td>4.38</td>
<td>4.23</td>
</tr>
</tbody>
</table>
In non-hierarchical methods, results crosstab can be seen from the final output cluster centers. Final Cluster Centers, the information can be seen as follows:

1. Cluster 1
   Cluster 1 consists of customers with characteristics of purchase frequency below the total average, total transactions below the total average, payment methods below the total average, arrival frequency below the total average, shopping satisfaction above the average total, and desire returns above the total average. So, cluster 1 is one of the most potential loyal customer groups.

2. Cluster 2
   Cluster 2 consists of customers with purchase frequency below the total average, total transactions above the total average, payment methods above the total average, arrival frequency below the total average, shopping satisfaction below the total average, and return desires below the total average. So, cluster 2 is a group of potential loyal customers.

3. Cluster 3
   Cluster 3 consists of customers with characteristics of purchase frequency above the total average, total transactions above the total average, payment methods above the total average, arrival frequency above the total average, shopping satisfaction below the total average, and a return desire above the total average. So cluster 3 includes a group of loyal customers who have less potential.

**Analisis AR-MBA**

The method used to determine consumer purchasing patterns by knowing the products often purchased together at Indomaret Kaliurang is Association Rule-Market Basket Analysis (AR-MBA) using the FP-Growth algorithm. In carrying out the AR-MB analysis, the first stage is to group departments according to product type groups determined based on the questionnaire that has been distributed.

**Table 3. Types of Products**

<table>
<thead>
<tr>
<th>Department</th>
<th>Types of Goods</th>
</tr>
</thead>
<tbody>
<tr>
<td>Department 1 (Basic Materials)</td>
<td>Oil, Sugar, Rice, Salt, Royko, Margarine, Spices, etc</td>
</tr>
<tr>
<td>Department 2 (Home Appliances)</td>
<td>Stella, Soap, Dish Soap, Toilet Cleaner, etc</td>
</tr>
<tr>
<td>Department 3 (Snacks)</td>
<td>Chocolatos, Nuts, Pilus, Tango, Wafer, Kuaci, etc</td>
</tr>
<tr>
<td>Department 4 (Drinks)</td>
<td>Mineral Water, Syrup, Pocari Sweat, Buavita, Fruit Tea, etc</td>
</tr>
<tr>
<td>Department 5 (Instant Food)</td>
<td>Mie Instan, Nuget, Frozen Kentang, Puding Instan, dll</td>
</tr>
</tbody>
</table>

Next, the AR-MBA method process design was carried out. At this stage, it is carried out using the software Rapid Minner Based on the results of data processing, the following is the output rule from AR-MBA:

**Table 4. Output Rules**

<table>
<thead>
<tr>
<th>No</th>
<th>Rules</th>
<th>Support</th>
<th>Confidence</th>
<th>Lift</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Dept 5, Dept 1 → Dept 2</td>
<td>12%</td>
<td>75%</td>
<td>1.786</td>
<td>2.320</td>
</tr>
<tr>
<td>2</td>
<td>Dept 2, Dept 1 → Dept 5</td>
<td>12%</td>
<td>75%</td>
<td>1.630</td>
<td>2.160</td>
</tr>
<tr>
<td>3</td>
<td>Dept 3, Dept 2 → Dept 5</td>
<td>13%</td>
<td>76.5%</td>
<td>1.662</td>
<td>2.295</td>
</tr>
</tbody>
</table>

Rule valid ones are those that have value Lift Ratio > 1, until the whole output rules the form in this research is valid. From the output results above, the following rules can be seen:

1. Department 5 and Department 1 will be purchased at the same time as Department 2, with a confidence percentage of 75% and 12% support from the overall data and a Lift Ratio obtained of 1.786.
2. Department 2 and Department 1 will be purchased at the same time as Department 5, with a confidence percentage of 75% and supported by 12% of the overall data and Lift Ratio The result obtained was 1,630.
3. Department 3 and Department 2 will be purchased at the same time as Department 5, with a confidence percentage of 76% and supported by 13% of the overall data and Lift Ratio The result obtained was 1,662.

From the rule above, it is obtained knowledge in the form of:

1. The choice of shopping place layout can be adjusted based on the rules formed, where products often purchased simultaneously are placed closer to where they are placed.
2. Products often purchased can also be brought closer together in catalog preparation.
3. To overcome the problem of rarely purchased products, discounts or promotions can be created by making rarely purchased products as a bonus for buying products that are frequently purchased.
4. Vouchers can be created as a promotional tool to encourage larger purchases from consumers.
Marketing Strategy

The following are marketing strategies that can be provided to Indomaret Kaliurang Yogyakarta:

   In the non-hierarchy method, results crosstab can be seen from the output of the final cluster centers. Final Cluster Centers, the information can be seen as follows:
   a. Cluster 1
      Cluster 1 consists of customers with characteristics of purchase frequency below the total average, total transactions below the total average, payment methods below the total average, arrival frequency below the total average, shopping satisfaction above the average total, and desire returns above the total average. So, cluster 1 is one of the most potential loyal customer groups.
   b. Cluster 2
      Cluster 2 consists of customers with purchase frequency below the total average, total transactions above the total average, payment methods above the total average, arrival frequency below the total average, shopping satisfaction below the total average, and return desires below the total average. So, cluster 2 is a group of potential loyal customers.
   c. Cluster 3
      Cluster 3 consists of customers with characteristics of purchase frequency above the total average, total transactions above the total average, payment methods above the total average, arrival frequency above the total average, shopping satisfaction below the total average, and a return desire above the total average. So cluster 3 includes a group of loyal customers who have less potential

   Based on the rules formed in data processing using data mining clustering techniques, the total learning variable is chosen as the most considered variable in determining whether consumers will return or not. Thus, Indomaret Kaliurang will pay more attention to this total expenditure variable. So, a marketing strategy that can increase customer satisfaction in shopping is to provide maximum service and pay more attention to loyal customers who enter the type of customers in cluster 1. In addition, the types of customers in the cluster can be used as marketing targets by providing several promotions so as to increase customer loyalty and increase sales. Marketing strategies that can be given such as providing loyalty cards to customers who have high loyalty who are able to provide special discounts so that Indomaret Kaliurang can maintain loyal customers. In addition, the development of product bundle packages that have been tailored to customer clusters and the placement of favorite products near the cashier so as to increase purchase impulses.

2. AR-MBA with Algorithms FP-Growth
   Based on the AR-MBA method, marketing strategies and frequently purchased items that can be provided to consumers are as follows:
   a. Department 5 which contains (instant ingredients) and department 1 which contains (basic ingredients) have a relationship with department 2 which contains (home supplies), so they are placed close to each other.
   b. Department 2 which contains (household supplies) and department 1 which contains (staple ingredients) have a relationship with department 5 which contains (instant ingredients), so they are placed close to each other.
   c. Department 3 which contains (light foods) and department 2 which contains (household equipment) have a connection with department 5 which contains (instant ingredients), so they are placed close to each other.

   Based on the results of data processing using AR-MBA, several marketing strategies are given such as, Products that are often purchased simultaneously are brought closer to the preparation of the catalog. In addition, to solve the problem of rarely purchased products, promotions can be made by making rarely purchased products as a bonus for buying frequently purchased products. Vouchers can also be created as a promotional tool to encourage greater consumer purchases. Some of these strategies have been applied to previous research which states that the application of the AR-MBA technique to find association rules so that it can provide a suitable marketing strategy can increase sales in the retail sector [29], [30].

Conclusion

Based on the research conducted, from the results of the Cluster Method with K-Means Non Hierarchical Algorithm that there are 3 clusters. Cluster 1 is one of the most potential loyal customer groups among the other two clusters. Marketing strategies that can be given are tailored to the most potential customer clusters such as providing loyalty cards to customers who have high loyalty, developing product bundle packages that have been tailored to customer clusters and placing favorite products near the cashier. With the
implementation of this strategy, Indomaret Kaliurang is expected to increase their sales. Meanwhile, based on the results of the AR-MBA method, it can be concluded that the goods purchased simultaneously are instant food and staples have a relationship with home supplies. In addition, home supplies and staples have a relationship with instant food. Snacks and home supplies have a relationship with instant food. So that the three rules can be placed close to each other. Based on the results of the clustering method and AR-MBA, Indomaret Kaliurang can make marketing strategies such as procuring product discounts according to products that are often purchased simultaneously by determining target customers based on the results of clustering that has been done. In addition, Indomaret Kaliurang will pay more attention to this total spending variable. So, the marketing strategy that can increase customer satisfaction in shopping is to provide maximum service and pay more attention to loyal customers. Suggestions that can be given for future research are to use similar methods in providing marketing strategies in different sectors so as to increase the feasibility or effectiveness of applying similar methods.

Bibliography


