Determining the retail sales strategies using association rule mining

Roaida Yanti, Prita Nurkhalisa Maradjabessy, Qurtubi, Ira Promasanti Rachmadewi
Department of Industrial Engineering, Faculty of Industrial Technology, Universitas Islam Indonesia, Yogyakarta, Indonesia

ABSTRACT

Competitive competition in the retail industry requires retailers to maintain improvements and formulate accurate strategies to maintain their competitiveness. A small number of daily visitors visit retail store Y if compared to other retail stores, which leads to decreased store revenue due to the small number of products sold. Therefore, it is crucial to formulate the right business strategy to increase sales by utilizing customer shopping behavior derived from transaction data. The method used is association rule mining (ARM) with a frequent pattern growth (FP-growth) algorithm to determine consumer buying patterns. Data processing results generate five valid rules that meet the specified criteria for an association relationship. Utilization rules are acknowledged by determining retail sales strategies by recommending store layouts, shopping catalogs, and voucher discounts to attract customers.

Keywords:
Data mining
FP-growth algorithm
In-store logistics
Market basket analysis
Store layout

1. INTRODUCTION

As time goes by, the increase in human needs and desires triggers the presence of various shopping places, such as the retail industry, that sell products to meet the multiple needs of society. The retail sector has become essential due to its significant economic impact [1], [2]. Retail is the final stage in the distribution channel, which incorporates all the activities related to providing consumers with goods and services for personal, family, or home usage [3].

Building on the existing literature, the research acknowledges the rapid growth of the retail business in Indonesia [4]-[6]. This is shown by the expanding number of retail companies from year to year and many traditional retailers that have altered to modern retail [7], [8]. Despite the rise of digital commerce, physical retail stores in Indonesia retain significant appeal. The number of retail outlets in 2018 has grown to 40,000 outlets spread throughout Indonesia. The market offers substantial opportunities for local and international companies, with retail sales valued at a staggering $350 billion in 2017, potentially surpassing other Southeast Asian nations [9]. Thus, the retail business has been recognized as one of the industries with very high competition [10]. Therefore, every company must be able to create several effective marketing plans and maintain and grow them over time to deal with a growing, competitive, solid environment [11]. This strategy is carried out to intensify income, where income increases from the value of consumer transactions [10]. Strategy determination can be performed by utilizing data mining implementation.

To gain profits and compete in a competitive environment, the retail industry actors realize that implementing data mining is a must, among others. Additionally, adopting data mining techniques has been highlighted as a crucial strategy for retail industry actors to gain insights into product sales trends, consumer
preferences, and sales predictions [12]. One of the data mining techniques commonly used to determine strategy in retail is association rule mining (ARM) [13], [14].

The ARM technique, also known as market basket analysis (MBA), is well-known due to its ability to identify patterns and spending habits of consumers in their shopping baskets by identifying associations and correlations among items. Many unique and hidden relationships among products in shopping malls cannot be detected easily [15]. The literature emphasizes the relevance of this technique in uncovering hidden relationships among products in shopping malls. MBA could help retail organizations in their analysis system to determine goods placement and distinguish design sale promotions for different customers to improve satisfaction and benefits [16]. Through the association rule-market basket analysis (AR-MBA) method, retailers can discover the relationship between one item and another that was previously unidentified [17]. Information from the association patterns can be used to find solutions and provide the best recommendations, including structuring layout and creating an effective and targeted marketing strategy. According to Ünvan [18], some purchases are made on impulse in the retail business, so utilization of MBA can provide clues as to what customers might buy if a product idea emerges.

Retail store Y is a modern retail store in Yogyakarta, Indonesia. The outlined problem revolves around the intense competition within the retail industry, specifically focusing on the challenges retail store Y faces. The research aims to address how retail store Y can optimize its layout and marketing approaches to not only attract customers but also outperform its neighboring competitors in the retail sector.

The unsolved problems and improvement areas focus on effective layout styles to guide customers across different store sections [19]. While literature acknowledges the importance of layouts to support sales and profits [20], there is a gap in understanding how retail store Y, amidst fierce competition, can optimize its layout and marketing strategies for tremendous success. Additionally, the specific challenges of retail store Y being near other retail outlets warrant exploration in terms of distinct marketing strategies and layout designs. The accurate and captivating layout will stimulate the consumers’ interest in shopping [21]. Sales promotion must also be considered since it plays a vital role in commercial sales practice. The retail business has become the area with the most significant sales promotion activities, and many retailers have implemented sales promotions magnificently [22]. Sales promotion is an essential part of a sales campaign, a collection of incentive tools designed to boost product buying or faster services [23]. Several retail promotion genres include lottery, product bundling offers, discounts, and coupons [24]. This study looked into the effects of employing AR-MBA to determine shelf layout and promotional policies in the retail industry. While previous studies investigated the impact of product placement in supermarkets on sales and revenue [18], they did not explicitly address the utilization of AR-MBA for optimizing shelf layout and promotional strategies in retail settings.

The issues center on optimizing retail store Y’s layout and marketing strategies amid fierce competition and proximity to other retail outlets. Existing literature recognizes the importance of layouts for sales and profits but needs more insight into retail store Y’s unique challenges. To fill this gap, the research introduces the AR-MBA method to uncover product associations and optimize layouts and marketing for success on Kaliurang Street, Yogyakarta. The subsequent sections of the research paper will systematically detail the research methodology, encompassing data collection, pre-processing, and analysis utilizing the AR-MBA method. This, in turn, will demonstrate how the identified associations and patterns can be used to inform layout designs, improve sales promotions, and enhance overall marketing strategies. The relevance of each section will be evident in showcasing the practical application of the AR-MBA method in addressing the specific challenges faced by retail store Y.

2. LITERATURE REVIEW

Data mining is a process that uses mathematical, artificial intelligence, machine learning, and statistical methods to extract and identify relevant information originating from various vast databases. Researchers stated that the availability of enormous amounts of data has triggered the interest of the information industry and the public for the last several years due to the need to transform information data into usable knowledge [25], [26]. Generally, data mining is described as detecting patterns in vast raw materials so that significant knowledge can be revealed [27].

ARM to relate one or more attributes in a data group with other attributes [28], uncover hidden yet significant relationships among attributes [29], and generate the if-then statement related to attributes’ value in the form of rules [30]. The number of association rules will be formed based on the pre-determined support and confidence value [31]. The support value determines how frequently a specific item or combination appears in the database. In contrast, the confidence value measures the accuracy ratio of item set correlation under a particular condition [32]. The lift ratio is a value that demonstrates the validity of a formed rule. A rule is confirmed as valid if a lift ratio is rated more than 1 [33]. The formulation for three parameters is presented as (1) to (4).

---

*Determining the retail sales strategies using association rule mining ... (Roaida Yanti)*
The calculation for support 1 item is obtained from the (1).

\[
\text{Support (A)} = \frac{\text{Number of Transactions that Contains } A}{\text{Total of Transactions}}
\]  

(1)

While the value of the support two items can be obtained under the (2).

\[
\text{Support (A, B)} = \frac{\text{Number of Transactions Containing A and B}}{\text{Total of Transactions}}
\]  

(2)

The calculation of the confidence value is generated from the (3).

\[
\text{Confidence (B)} = \frac{\text{Total Transaction containing A and B}}{\text{Total Transaction Containing A}}
\]  

(3)

The calculation of the lift ratio value is generated from the (4).

\[
\text{Lift Ratio} = \frac{\text{Support (A \& B)}}{\text{Support (A)} \cdot \text{Support (B)}}
\]  

(4)

MBA is a modeling technique to analyze customers’ shopping patterns by learning the association among products in their shopping basket in a particular transaction [34]. In brief, MBA shows the most frequently bought product combination for a specific order. According to Yanti et al. [35], market basket analysis is a widely used data mining method to gain attractive business strategies. It provides information for retailers to understand customer purchasing behavior, which can help retailers. This association can be used to improve profitability through recommendation, promotion, cross-selling, and even the placement of the item on the menu or store [16]. Moreover, MBA enables the company to identify the essential products that separate it from others in the market, which might lead to business loss due to the absence of these products or price increases.

The frequent pattern growth (FP-growth) algorithm is an association rule technique that can identify the dataset’s most prevalent product set [36]. The development of this apriori algorithm lies in database scanning and rule accuracy. FP-growth is more beneficial since scanning the database takes only once or twice. On the other hand, apriori takes multiple scans [37]. The FP-growth algorithm doesn’t have to take the item combination from the data set, so the candidate items are undeveloped, as in the apriori algorithm [38]. The FP-growth algorithm extracts common itemsets by creating a frequent pattern tree (FP-tree) structure [39].

Randhawa and Saluja [40] explained the correlation between impulsive buying behavior and various visual merchandises. One way to boost impulsive buying is by implementing an MBA. This research recommends the following study to gather data from other areas for alternative results. Zhao et al. [41] suggested future researchers study the combination of display layout determination and promotional policy, such as discount or product bundling with discount. Azhra et al. [42] implied that future research could combine the existing promotional strategies. Kurnia and Lestari [43] proposed the following study to assess the layout recommendation to improve sales.

This research presents recommendations on retail sales marketing by determining the shelf layout and promotional policies: discount and catalog. This strategy is taken to increase sales and benefits in the retail industry by employing ARM-MBA. Leveraging the insights from MBA, this research proposes data-driven recommendations for shelf layout and promotional policies, offering retailers a scientific and objective approach to maximizing sales and profitability.

3. RESEARCH METHOD

This research was conducted using data collection by directly observing retail store Y. The data collection involves extracting valuable information from shopping receipts, blueprints, and product layout details. Specifically, the term “layout” is defined within the context of the store, referring to the arrangement of goods. The pre-processing stage is meticulously executed to ensure data integrity, involving the removal of duplicates and noise. Moreover, a critical transformation of purchased product data into binary numbers facilitates subsequent analysis, with one value assigned to products bought in a specific department and zero for unpurchased items. Using Excel Software and Rapidminer facilitates efficient data processing, incorporating the AR-MBA method and explicitly employing the FP-growth algorithm.
The direct observation of the retail store Y provides a real-world context for data collection, ensuring that the information gathered is relevant and reflects actual consumer behavior. By thoroughly describing the data pre-processing steps, including cleaning duplication and noise reduction, the methodology ensures the integrity of the dataset. The transformation of product data into binary numbers is a crucial step, providing a clear understanding of how the AR-MBA is applied. The FP-growth algorithm’s effectiveness in revealing patterns and relationships within sizable datasets is the foundation for using AR-MBA. The FP-growth algorithm develops an FP-tree to find frequent itemsets and search for the most common itemsets in the studied data [44]. This method is well-suited for extracting meaningful insights from the complex shopping behavior of customers, offering a robust foundation for decision-making.

4. RESULTS AND DISCUSSION

This research employs transaction data from the company’s history in September 2022. Next, to get rules formed, it is necessary to determine the minimum value of support and minimum confidence. The value of minimum support shows the percentage of the minimum requirement for the probability percentage of item combinations that appear in transactions. This study determined the value of minimal support as 6%. Meanwhile, the minimum confidence value indicates the minimum requirement of the confidence percentage level on the combination of the items purchased simultaneously. In this study, the value of minimal confidence is 40%, as previously determined by researchers.

4.1. Condition of retail store Y

Retail store Y has a small number of daily customers compared to other retail stores. Inconvenient and less strategic product layout placement can cause this condition. Without addressing this issue, retail store Y risks losing significant revenue and market share due to customer frustration and missed sales opportunities.

Product type-department layout is implemented in department 6 (D6), loaded with ice cream, and department 3 (D3), which includes flavored drinks and pudding in the same space since they must be stored in the freezer. Department 2 (D2), with cleaning products, and department 14 (D14), with fragrance products, are in the same corridor. Generally, food products are stored at the inner right side of the store entrance, while non-food products are usually positioned at the left side of the store. The sales department must also execute sales strategies such as issuing promo catalogs or giving voucher discounts. Therefore, the ARM is essential to solving the existing problem. The initial layout of retail store Y is illustrated in Figure 1.

![Initial layout of retail store Y](image)

Figure 1. The initial layout of retail store Y

4.2. Association rule result analysis

Value determination of minimal support and minimum confidence that has previously been performed helps determine the most appropriate way to make decisions. To be able to be analyzed, two values are entered into the transaction data processing on Rapidminer software. The output results are in the form of rules, as illustrated in Table 1.

---

Determining the retail sales strategies using association rule mining ... (Roaida Yanti)
Based on Table 1, five rules are formed with the following analysis: i) If products in department 15 are purchased, there is a 40% chance that products in department 13 will also be bought, supported by 6.2% of all data and the lift ratio obtained by 2.064; ii) If products in department 9 are purchased, then there is a 40.9% possibility that products in department 5 will also be bought, supported by 7% of all data and lift ratio obtained by 1.426; iii) If products in department 15 are purchased, there is a 45% chance that products in department 5 will also be bought, supported by 7% of all data and lift ratio obtained by 1.569; iv) If products in department 11 are purchased, there is a 47.1% possibility that products in department 3 will also be bought, supported by 6.2% of all data and lift ratio obtained by 1.641; and v) If products in department 11 are purchased, there is a 47.1% chance that products in department 13 will also be bought, supported by 6.2% of all data and lift ratio obtained by 2.426.

From the results above, the provisions of minimum confidence of 40%, minimal support of 6%, and lift ratio > 1 have been fulfilled by all rules. Therefore, all five rules are valid and related. This shows that the purchase of products in specific departments correlates with the likelihood of purchasing products in other departments. The proposed method in this study tended to have an inordinately higher proportion of associated purchases, indicating solid correlations between specific product categories. These findings demonstrate clear patterns of association between product categories within the retail environment, suggesting that strategic placement and promotional efforts in one department can influence purchases in related departments. Armed with these validated rules, retail store Y can derive actionable insights. The rules provide a data-driven foundation for making informed choices, optimizing resource allocation, and enhancing operational efficiency.

These findings indicate that meeting stringent criteria such as minimum confidence, minimal support, and lift ratio in ARM leads to reliable insights for decision-making in the retail industry. Unlike some previous studies by Randhawa and Saluja [40], which focused solely on product placement or promotional strategies without considering the interplay between various departments, our approach emphasizes the holistic understanding of customer buying patterns. This allows retail store Y to optimize its product layout, design visually appealing catalogs, and offer promotions and discounts strategically without negatively affecting overall sales performance.

4.3. Recommendations for output
Retailers must be more competitive, considering the number of retailers offering their uniqueness and leaving consumers with more purchasing choices. In light of the data processing results that validate the robustness of the identified rules, retail store Y can strategically position itself in the competitive landscape by leveraging these insights to tailor marketing strategies, optimize product placements, and enhance the overall shopping experience for consumers. The implications extend beyond the immediate operational enhancements, fostering long-term competitiveness and sustainability in the dynamic retail landscape. Based on the analysis of the data processing result, the recommendations are suggested for retail store Y as follows.

4.3.1. Layout determination
The store layout is a significant variable affecting shopper behavior and a primary determinant of the overall store image [45]. A retail establishment’s physical design has been shown to impact customers’ attitudes and actions and store performance. Exposing merchandise to customers to assist their contemplation, which ultimately results in purchasing the exposed things, is a crucial function of store layout. This effort aims to ascertain shelf layout to maximize product exposure to buyers [46].

A well-designed layout is essential since it greatly influences the design of the in-store movement, the shopping environment, shopping behavior, and operational productivity. In addition, the product layout in a retail store can influence consumer shopping behavior, so the right strategy is necessary for its design. Layout recommendations for retail store Y are shown in Figure 2.

### Table 1. Result of association rule

<table>
<thead>
<tr>
<th>No.</th>
<th>Premises</th>
<th>Conclusion</th>
<th>Supp</th>
<th>Conf</th>
<th>Lift</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Dept 15</td>
<td>Dept 13</td>
<td>0.062</td>
<td>0.400</td>
<td>2.064</td>
</tr>
<tr>
<td>2</td>
<td>Dept 9</td>
<td>Dept 5</td>
<td>0.070</td>
<td>0.409</td>
<td>1.426</td>
</tr>
<tr>
<td>3</td>
<td>Dept 15</td>
<td>Dept 5</td>
<td>0.070</td>
<td>0.450</td>
<td>1.569</td>
</tr>
<tr>
<td>4</td>
<td>Dept 11</td>
<td>Dept 3</td>
<td>0.062</td>
<td>0.471</td>
<td>1.641</td>
</tr>
<tr>
<td>5</td>
<td>Dept 11</td>
<td>Dept 13</td>
<td>0.062</td>
<td>0.471</td>
<td>2.428</td>
</tr>
</tbody>
</table>
Determining the retail sales strategies using association rule mining

( Roaida Yanti )

Figure 2. Store layout recommendation

The recommendations for changing the layout, as depicted in Figure 2 and derived from the Rapidminer software output, hold substantial implications for influencing consumer behavior. Placing frequently-purchased products adjacent to rarely-purchased items can optimize resource utilization and capitalize on consumers’ purchasing patterns. This strategic layout is anticipated to encourage customers to buy more products, leveraging the convenience of accessing related items quickly. The practical application of these recommendations aligns with the overarching goal of retail store Y, which is to increase sales and enhance customer satisfaction and loyalty.

4.3.2. Promotion catalog

The findings regarding the positive influence of sales promotions on impulsive purchases and the potential increase in sales through promotional catalogs, as indicated by the research [47], hold significant ramifications for the future strategies of retail store Y. Making a promotional catalog can be completed by observing the rarely purchased products, which can later be bundled with rarely-purchased products by providing a small discount. An example of an offer in a rule-based promotional catalog formed in the customers’ shopping cart of retail store Y by offering product bundling is that the buyer will only pay IDR 12,000 for one preferred drink product and one snack product. It is considered as the utilization of the 4th rule, in which the consumers who buy products in department 11 (snacks) also have the potential to purchase products in department 3 (drinks) by displaying attractive product photos as the catalog cover and bundling the less-purchased beverage products with snack products which are often bought or vice versa.

In the future, these insights will come in handy as retail store Y continues to refine and innovate its sales and marketing strategies. The practice of creating rule-based promotional catalogs can be extended to other departments and product categories, providing a versatile approach to boosting sales. Understanding the associations between different departments and products allows customized promotions tailored to specific customer preferences, maximizing the effectiveness of promotional efforts. Furthermore, the future utility of these findings extends to adapting to evolving consumer preferences and market dynamics. As retail store Y continually analyzes transaction data and customer behavior, it can refine its promotional strategies to align with changing trends and demands. This adaptability is crucial for long-term success in the retail industry, ensuring that promotional efforts remain relevant and appealing to the target audience.

4.3.3. Shopping voucher

As highlighted in the research [48], the insight into the effectiveness of price discounts emphasizes the significance of pricing value in shaping customer perceptions. Applying shopping vouchers that provide discounts on specific products, such as dairy products, can be a powerful incentive for customers to purchase. This tactic aligns with consumer preferences for value and strategically targets products that require higher sales at retail store Y.

Targeted discounts on specific product categories, contingent on a minimum purchase threshold, encourage customers to buy more and focus on rarely purchased products. The example of offering a 25%
discount on dairy products for a minimum purchase of IDR 100,000 reflects a thoughtful approach to boosting sales while maintaining consideration for pricing and profit margins. Furthermore, applying limited-time purchase promotions creates a sense of urgency and can lead to impulsive buying decisions. This aligns with consumer psychology, as highlighted in the research, wherein time-limited discounts can cause a feeling of regret if customers do not take advantage of the offer promptly. This sense of urgency can be harnessed in future voucher strategies to drive immediate purchases and capitalize on customers’ desire to secure discounts. The future utility of these findings extends to the ongoing adaptation of retail store Y to changing market dynamics and consumer behaviors. By incorporating the lessons learned from the effectiveness of shopping vouchers, the store can continually refine its promotional strategies, keeping them aligned with evolving customer preferences and competitive pricing landscapes.

5. CONCLUSION

The identified association rules from retail store Y’s transaction data analysis provide practical insights for current decision-making and future strategies. Based on the transaction data processing and analysis at retail store Y, five valid association rules were obtained that met the criteria of the inter-items relationship, which are the value of minimum confidence of 40%, minimal support of 6%, and lift ratio > 1 that are achieved by each output rule. Meeting stringent criteria such as minimum confidence, minimal support, and lift ratio, these rules offer a reliable foundation for informed choices in product placement, layout design, and promotions. Aligning with these rules, optimizing the retail product layout, creating visually appealing catalogs, and offering promotions and voucher discounts are recommended to enhance customer shopping patterns and foster long-term engagement. These findings serve as a valuable resource for guiding retail store Y in understanding evolving customer behaviors and refining strategies, ensuring adaptability and sustained competitiveness in the retail market. In conclusion, these insights enable retail store Y to make data-driven decisions that boost current sales and position the store for long-term success and adaptability. Additionally, external factors like season variations or economic trends were not considered, potentially impacting purchasing behavior. Further research across diverse retail environments is needed to validate and enhance the effectiveness of these association rules.

REFERENCES


Determining the retail sales strategies using association rule mining … (Roaida Yanti)
BIOGRAPHIES OF AUTHORS

Roaida Yanti is an industrial engineering graduate from the Department of Industrial Engineering, Faculty of Industrial Technology, Universitas Islam Indonesia, Yogyakarta. She was a Data Mining Laboratory Assistant. She actively conducts research and publishes articles. She is interested in operation research, logistics, halal supply chain, and industrial engineering. She can be contacted at email: roaida.yanti@alumni.ui.ac.id.

Prita Nur Khalisa Maradjabessy is an undergraduate at the Department of Industrial Engineering, Faculty of Industrial Technology, Universitas Islam Indonesia. Currently, she serves as an Assistant in the Data Mining Laboratory. Her research interests are marketing logistics, customer segmentation, and industrial engineering. She can be contacted at email: prita.maradjabessy@students.ui.ac.id.

Qurtubi is an Associate Professor in the Department of Industrial Engineering, Faculty of Industrial Technology, Universitas Islam Indonesia, Yogyakarta. He obtained his doctoral degree in the area of logistics performance. His research interests are marketing logistics, supply chain, and industrial engineering. He can be contacted at email: qurtubi@ui.ac.id.

Ira Promasanti Rachmadewi is an Assistant Professor in the Department of Industrial Engineering, Faculty of Industrial Technology, Universitas Islam Indonesia, Yogyakarta. She is Secretary of the International Undergraduate Program (IUP) of the Industrial Engineering, Universitas Islam Indonesia. She is interested in information engineering, applied machine learning, health informatics, and industrial engineering. She can be contacted at email: irarachmadewi@ui.ac.id.