IJCCS (Indonesian Journal of Computing and Cybernetics Systems)
Vol.x, No.x, July xxxx, pp. 1~5
ISSN (print): 1978-1520, ISSN (online): 2460-7258
DOI: 10.22146/ijccs.xxxx

SMART PRODUCT RECOMMENDATIONS IN THE E-COMMERCE WEBSITE: UTILIZING THE APRIORI ALGORITHM FOR MARKET BASKET ANALYSIS

Hendra[1], Aditiya Hermawan*[2], Edy[3]
[1]Bachelor Program of Informatics Engineering, Buddhi Dharma University, Banten, Indonesia
[2]Deparment of Informatics Engineering, Buddhi Dharma University, Banten, Indonesia
[3]Department of Software Engineering, Buddhi Dharma University, Banten, Indonesia
e-mail: [1]hendra@ubd.ac.id, ***2aditiya.hermawan@ubd.ac.id** , [3]edy.edy@ubd.ac.id

Abstrak

Dunia perdagangan online menjadi semakin kompetitif dan untuk sukses di bidang ini tidak cukup hanya dengan menampilkan produk kepada calon pembeli. Sangat penting untuk menawarkan beragam produk dan menjaga agar rekomendasi produk tetap mutakhir terutama bagi pelanggan yang cenderung membeli banyak barang. Untuk mengatasi tantangan ini, diperlukan sistem cerdas yang dapat secara otomatis menghasilkan rekomendasi produk yang sedang tren berdasarkan data penjualan. Pada penelitian ini, metode Market Basket Analysis (MBA) digunakan untuk menganalisis data transaksi konsumen dan mengidentifikasi produk yang sering dibeli bersamaan. Algoritma apriori diterapkan untuk menghasilkan aturan asosiasi dan parameter Lift Ratio digunakan untuk mengevaluasi kekuatan aturan-aturan tersebut. Penelitian ini diimplementasikan pada sebuah situs web e-commerce, dan dari sistem tersebut aturan asosiasi yang dihasilkan akan diterapkan untuk memberikan rekomendasi produk otomatis berdasarkan tren data penjualan beberapa waktu terakhir. Hasil penelitian menunjukkan bahwa sistem rekomendasi produk otomatis yang dibuat pada website e-commerce sangat membantu pengguna dalam meningkatkan pengalaman berbelanja online mereka. Penggunaan parameter Lift Ratio dalam validasi aturan asosiasi memberikan bukti kuat tentang relevansi dan akurasi dari rekomendasi produk yang dihasilkan, yang pada gilirannya dapat meningkatkan tingkat kepuasan pelanggan dan potensi penjualan.

Kata kunci*— Apriori, Data Mining, E-commerce, Market Basket Analysis (MBA), Website*

Abstract

The world of online commerce is becoming increasingly competitive, and to succeed in this field, it is not enough to showcase products to potential buyers. It is crucial to offer various products and keep product recommendations up-to-date, especially for customers who buy multiple items. To address this challenge, an intelligent system is needed that can automatically generate trending product recommendations based on sales data. In this research, the Market Basket Analysis (MBA) method analyzes consumer transaction data and identifies products often purchased together. The apriori algorithm is applied to generate association rules, and the Lift Ratio parameter is used to evaluate the strength of these rules. This research is implemented on an e-commerce website, and the generated association rules will be applied to provide automatic product recommendations based on recent sales trends. The results show that the automatic product recommendation system developed for the e-commerce website significantly helps users enhance their online shopping experience. Using the Lift Ratio parameter in validating association rules provides strong evidence of the relevance and accuracy of the generated product recommendations, which can increase customer satisfaction and sales potential.

Received June 1st,2012; Revised June 25th, 2012; Accepted July 10th, 2012

Keywords— *Apriori, Data Mining, E-commerce, Market Basket Analysis (MBA), Website*

1. INTRODUCTION

In modern times, technology plays an important role. With the emergence of technology, it has become an essential medium for sharing information and conducting business transactions. The business transaction process requires electronic media to support the speed of information in communication and business using electronic devices such as notebooks, smartphones, and personal computers (PCs) [1].

One form of technological development is the presence of the Internet in the community. The Internet is a crucial intermediary for users to share information and make transactions in the company's business processes. The development of the Internet in the industrial world has resulted in a new type of business form, e-commerce, which makes it easy for customers to buy goods online [2].

Integrating e-commerce into our daily lives has proved to be an essential aspect of modern society. This technological advancement has revolutionized transacting, offering unparalleled convenience and transforming our shopping experiences. However, this convenience has created intense competition, with e-commerce platforms offering products at lower prices than their brick-and-mortar counterparts. As such, it has become increasingly challenging for traditional stores to keep up with the convenience and affordability of e-commerce platforms. According to the Bank Indonesia Annual Meeting Report in 2021, shown in Figure 1, Indonesia's e-commerce transaction value has steadily increased over the past four years. In 2018, it amounted to 106 trillion Rupiah, followed by 206 trillion Rupiah in 2019, 266 trillion Rupiah in 2020, and 401 trillion Rupiah in 2021. It is expected to continue growing in the coming years [3].

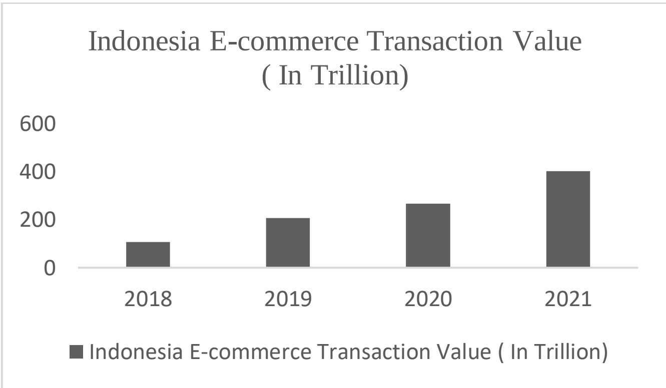

Figure 1 Indonesia Total E-commerce Transaction Value

The escalation in the value of e-commerce transactions necessitates a multifaceted approach beyond mere product display. The imperative lies in developing a sophisticated product recommendation system that entices visitors to increase purchasing activities, thereby augmenting the company's revenue streams. The proliferation of diverse product offerings and the absence of real-time updates about product variations engender potential confusion among consumers engaged in multi-item shopping endeavors. Fortunately, an intelligent system can be developed to provide automatic product recommendations based on current sales trends. This system can improve the customer experience by suggesting items that complement their original purchase. Using data mining in this system can assist users in making informed decisions about the products they plan to buy [4].

The market basket analysis (MBA) method, also known as the association rules method, can be applied to the commercial field so that it is possible to discover the behavior of customers when shopping and can also be used by companies to make product recommendations aimed at various types of customers [5]. Market basket analysis is a technique that analyzes the contents of consumer shopping baskets through all consumer transaction data to determine which products are most frequently purchased by consumers. The availability of accurate and efficient types of transaction data in the transaction data analysis process will help companies find interesting shopping patterns so that the results of the analysis can be used to make strategies for running a business, such as making product layout arrangements simultaneously [6]–[10] and creating a product recommendation system for visitors [1], [2], [4].

One algorithm that proves helpful in identifying correlations between sales itemsets is apriori. It generates association rules that demonstrate a strong connection. To derive these rules, we must establish the minimum values for support and confidence [2]. The Lift Ratio parameter is used to evaluate the strength of each rule and determine whether it exhibits a strong, weak, or nonexistent correlation [11]–[15].

This research follows a structured process to develop an effective product recommendation system: Data Provision involves collecting and preparing data from Kaggle.com. Data Cleaning and Transformation ensures the data is clean and in a suitable format for analysis. The Apriori Algorithm is applied to generate association rules by identifying frequent item sets and calculating support and confidence values. Evaluation Using Lift Ratio measures the strength and validity of these rules. Finally, implementing Recommendations applies the validated rules on the e-commerce website to provide automatic product recommendations based on recent sales trends. This structured approach aims to enhance the online shopping experience and provide valuable insights for e-commerce businesses. The characteristics used to analyze transactions in this research include the frequency of item purchases, co-occurrence of items in transactions, and repeated purchase patterns. The research identifies patterns and correlations in consumer purchasing behavior by analyzing these characteristics, leading to practical product recommendations.

Based on the explanation given above. Our research uses the apriori algorithm to create an advanced recommendation system for e-commerce platforms. The system will provide dynamic product recommendations that align with current market trends by analyzing sales transaction data, empowering buyers to make informed and confident purchasing decisions.

2. METHODS

To carry out our research, we have determined that the most effective approach is to employ the CRISP-DM (Cross-Industry Standard Process for Data Mining) methodology. This particular methodology is widely acknowledged for its ability to consistently and efficiently guide data mining processes, and it has been successfully implemented across various industries [5]. By employing this methodology, we can ensure that our research will be thorough and reliable, allowing us to draw accurate and meaningful conclusions from our data. CRISP-DM methodology consists of six significant phases, which can be seen in Figure 2 below:

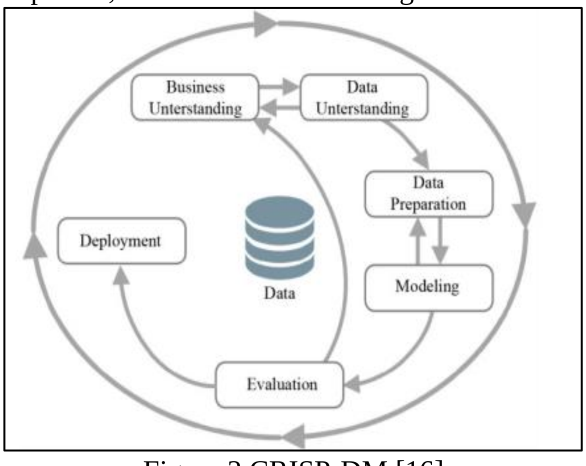

Figure 2 CRISP-DM [16]

The CRISP-DM methodology consists of six primary phases. First, the Business Understanding phase involves defining the project objectives and requirements from a business perspective and translating them into a data mining problem definition, which, in our case, is enhancing the product recommendation system on an e-commerce platform. Second, the Data Understanding phase includes initial data collection and exploration to understand data properties and identify quality issues using transaction data from Kaggle.com. Third, the Data Preparation phase involves cleaning, transforming, and structuring the data to handle missing values, remove noise, and format it for the apriori algorithm. Fourth, the Modeling phase applies the apriori algorithm to generate association rules for the recommendation system. Fifth, the Evaluation phase assesses the model's effectiveness using support, confidence, and lift ratio metrics to ensure it meets business objectives. Finally, the Deployment phase integrates the validated association rules into the e-commerce website to provide real-time product recommendations. By following these phases systematically, we ensure a structured and comprehensive approach to data mining, leading to reliable and actionable insights.

Figure 3 demonstrates how the obtained database undergoes the data preparation stage to assist the researcher in processing data by removing unnecessary attributes. This stage includes selecting tables, records, data attributes, and data transformation. Once the data preparation stage is complete, the data is ready for processing using the apriori algorithm to generate rules by determining the minimum support value. The next step is determining the minimum confidence value to create association rules. The obtained association rules will be utilized as product recommendations on e-commerce websites.

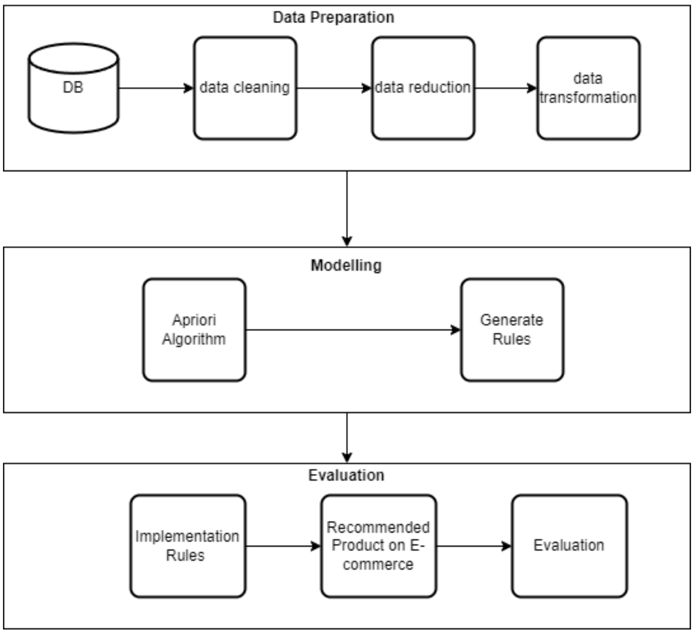

Figure 3 Stages in the process of creating association rules

2.1 Business Understanding

The business understanding stage is to understand the problems, goals, and needs from a business perspective. It is necessary to understand the data mining activities that will be carried out. The purpose of this research is to find purchasing patterns from consumer transactions to find out what types of products are purchased simultaneously so that it can facilitate e-commerce owners in making product recommendations to customers when shopping on the e-commerce web. To get new information, this research utilizes secondary data from purchase transaction data made by consumers accessed by researchers on the Kaggle.com website. Using data mining techniques, stored transaction data can be explored to get new information about consumer purchasing patterns (association rules) in customers' purchasing of items. Using the market basket analysis method can help find out the habits of purchasing goods by consumers, which products are purchased together to find combination association rules to create a product recommendation system on an e-commerce website that makes it easier for customers when shopping on the website.

2.2 Data Understanding

At this stage, the data used in this study is secondary data obtained from the online repository website Kaggle.com. The secondary data consists of 2,019,500 entries containing 200,000 transaction IDs and 134 types of products with 12 attributes. The attributes listed in Table 1 provide an overview of the structure and critical components of the transaction data. This stage helps identify basic patterns and frequencies in the data, such as the distribution of order times and the frequency of reorders.

Table 1 Table Attributes in transaction data

order_id	user_id	order_number	order_dow	order_hour_of_day	days_since_prior_order
1244093	104159	4	5	14	30.0
1244093	104159	4	5	14	30.0
1244093	104159	4	5	14	30.0
1244093	104159	4	5	14	30.0
1244093	104159	4	5	14	30.0

product_id	add_to_cart_order	reordered	department_id	department	product_name
57	1	0	14	breakfast	granola
37	2	0	1	frozen	ice cream ice
121	3	0	14	breakfast	cereal
121	4	0	14	breakfast	cereal
24	5	1	4	produce	fresh fruits

2.3 Data Preparation

1. Data Cleaning

Data Cleaning refers to cleaning data (removing inconsistent data noise) before mining. This stage is crucial to ensure the quality and accuracy of the data used for analysis. The data cleaning stage involves filling in missing or zero values in columns that may occur due to data consolidation. In this research, missing data entries are identified as wholly absent or marked as null, while zero values are identified in contexts where a positive value is expected. For numerical data, missing values are imputed using the mean or median to preserve the central tendency and robustness against outliers. The mode (most frequent category) is used for categorical data to maintain representativeness. These methods ensure data integrity and reliability, essential for accurate and meaningful data mining results..

2. Data Reduction

This phase is the process of reducing the dimensionality of the data and deactivating/removing fields that are considered not contributing to the final result. The following is the data reduction process for the original data to reduce the actual data or to remove unimportant characteristics, such as attributes: user_id, order_number, order_dow, order_hour_of_day, days_since_prior_order, product_id, add_to_cart_order, department_id, department. Below is an example of the results of the reduction in the data set that will be used in Table 2.

Table 2 Reduction Results

order_id	product_name
1244093	granola
1244093	ice cream ice
1244093	cereal
1244093	cereal
1244093	fresh fruits
1244093	soft drinks
1244093	soft drinks
1244093	packaged vegetables fruits
1244093	frozen appetizers sides

3. Data Transformation
At this stage, the data undergoes a series of changes and merges to ensure optimized quality. This involves converting the data into a format suitable for mining and performing operations with improved efficiency and effectiveness. As given in table 3, where the sales transaction data will be merged with every same order_id, put together the shopping data into one row.

Table 3 Transformation Results

Order_id	Product_name
1244093	granola,ice cream ice,cereal,cereal,fresh fruits,soft drinks,soft drinks,packaged vegetables fruits,frozen appetizers sides,frozen breakfast,eggs
1269990	refrigerated,crackers,packaged cheese,packaged cheese,fresh fruits,chips pretzels,canned meals beans,candy chocolate,chips pretzels,hot cereal pancake mixes
1846216	fresh fruits,fresh fruits,fresh fruits,packaged vegetables fruits,refrigerated,crackers,soy lactosefree,soy lactosefree

2.4 Apriori Algorithm
The Apriori algorithm is a method employed to identify frequent association patterns in order to generate related rules. It leverages prior knowledge about the attributes of previously known frequently occurring itemsets [17]. The steps taken in finding the frequency of items with the a priori algorithm are described below [2]:
1. Join
This procedure involves merging items with other items until no further combinations can be produced.
2. Prune
The results of the items obtained from the pruning process are combined items and then pruned using a predetermined minimum support value.
The principles in the apriori algorithm are as follows:
1) Collect one item, then find the item with the highest frequency;
2) Get candidate pairs, then count the principal pairs of each item;
3) Find the candidate of each principal pair of each item, and so on;
4) Each part of the frequent itemset must appear frequently.

In Figure 4, the algorithm counts the number of items in the dataset. Then, it creates itemsets to identify candidates and checks their support and confidence against the user-defined minimum thresholds. Finally, it uses frequent itemsets to uncover association rules between items in significant database transactions.

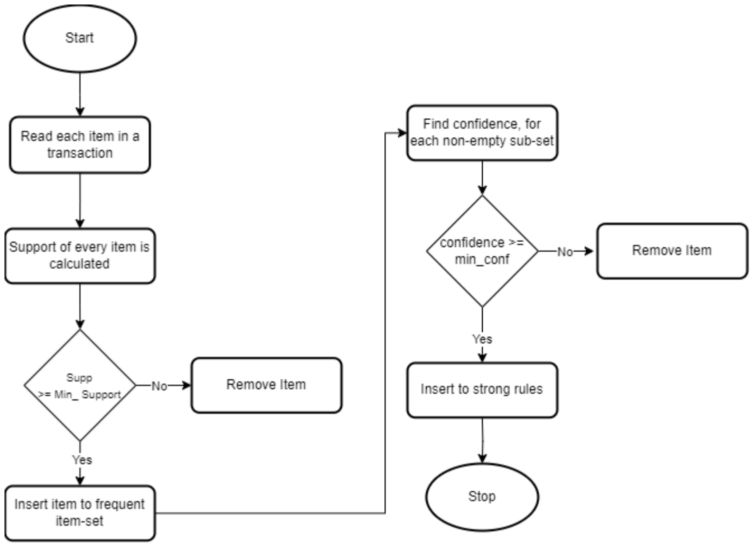

Figure 4 Flowchart Apriori[1]

2.5 Association Rules

Generating association rules entails analyzing pre-existing datasets to uncover relationships between items [18]. The basic rules in creating association rules require support values and confidence values that are greater than or equal to the support values and confidence values specified by the user. The first step in forming association rules is to identify product combinations that meet the predefined support values in the database. The second is that after all product combinations that meet the support value are met, the product combination obtained will also be calculated for the confidence value that has been determined. This is done to explain the confidence level of buying item X and then buying item Y. The support and confidence formula can be written in the equation below:

$$\text{Support X} = \frac{\sum \text{Transaction contains item X}}{\sum \text{Transactions}} \quad (1)$$

$$\text{Support X,Y} = \frac{\sum \text{Transaction contains item X \& Y}}{\sum \text{Transactions}} \quad (2)$$

$$\text{Confidence X -> Y} = \frac{\sum \text{Transaction contains item X \& Y}}{\sum \text{Transactions } contains\ item\ X} \quad (3)$$

After obtaining the final association rules from product combinations that have met the predetermined support and confidence values, the Final Association Rule will be validated using the Lift Ratio parameter. The lift ratio is an important measuring tool in generating association rules and measuring the precision and accuracy of measuring instruments (support and confidence). In this case, the lift ratio confirms whether item A is used together with item B. The Lift Ratio Correlation value can be determined through the following equation formula:

$$\text{Lift(X ,Y)} = \frac{P(X \cap Y)}{P(X) * P(Y)} \quad (4)$$

If the result of the lift ratio is below 1, it indicates a negative correlation between variables X and Y. Conversely, a result above 1 indicates a positive correlation. However, if the result equals 1, it indicates no correlation between variables X and Y[19].

This is to make it easier to understand the stages of forming the final association rule based on the formula that has been given. Here, a calculation using ten transaction data will be given. The explanation will be made by determining the support value of 40% and 100% confidence. In Table 4, the data is neat and ready to be calculated. Table 5 is the stage where each item item's support value is calculated. Then, each item that meets the support value will be combined with other itemsets that also meet the support value, while items that do not meet the support value will be deleted. Then, the calculation is carried out again from each combination of new itemsets until there are no more combinations left, as will be made in Table 6,7,8.

Table 4 lists the product names associated with each Order ID, providing a detailed view of the transaction data. This stage helps identify the items purchased in each transaction, forming the basis for subsequent calculations.

Table 4 List of 10 Order_id

Order_id	Product_name
998034	frozen produce,chips pretzels,fresh dips tapenades,hot cereal pancake mixes,fruit vegetable snacks,packaged vegetables fruits,tortillas flat bread,chips pretzels,popcorn jerky,condiments,fresh fruits,fresh fruits,packaged cheese,yogurt,fresh fruits,frozen appetizers sides,fresh
1244093	granola,ice cream ice,cereal,cereal,fresh fruits,soft drinks,soft drinks,packaged vegetables fruits,frozen appetizers sides,frozen breakfast,eggs
1269990	refrigerated,crackers,packaged cheese,packaged cheese,fresh fruits,chips pretzels,canned meals beans,candy chocolate,chips pretzels,hot cereal pancake mixes
1846216	fresh fruits,fresh fruits,fresh fruits,packaged vegetables fruits,refrigerated,crackers,soy lactosefree,soy lactosefree
……….	…………………….
2821997	cream,dry pasta,yogurt

Table 5 calculates the support value for each item (1-itemset). Items meeting the minimum support value (40%) are retained for further combination, while those that do not are discarded.

Table 5 Process C1 to L1

Item	Frequent	Support
buns rolls	1	10%
candy chocolate	1	10%
canned meals beans	**3**	**30%**
cereal	1	10%
chips pretzels	**3**	**30%**

In Table 6, the 2-itemsets are formed by combining the 1-itemsets that met the support value. The support for these new combinations is then calculated, and the process continues iteratively.

Table 6 Process C2 to L2

Item	Frequent	support
crackers ,packaged vegetables fruits	1	10%
crackers ,refrigerated	**3**	**30%**
crackers ,yogurt	1	10%
fresh fruit ,frozen appetizer sides	**3**	**30%**

fresh fruit ,packaged cheese	2	20%
…..	……	…..
refrigerated ,yogurt	1	10%

Table 7 continues the process with 3-itemsets, combining the 2-itemsets that meet the support threshold. This iterative process continues until no further combinations meet the support value.

Table 7 Process C3 to L3

Item	Frequent	Support
crackers refrigerated ,fresh fruit	**3**	**30%**
fresh fruit frozen appetizer sides ,packaged vegetables fruits	**3**	**30%**
fresh fruit frozen appetizer sides ,refrigerated	2	20%
….	….	….

Then, each combination of item sets that have met the support value will be tested for the confidence value to get the association rule, as shown in Table 8 below.

Table 8 Calculate All Confidence on Association Rule

Combination	Support A ∩ B	Support A	Confidence
crackers → fresh fruit	**3**	**3**	**100%**
crackers → refrigerated	**3**	**3**	**100%**
fresh fruit → frozen appetizer sides	3	6	50%
fresh fruit → packaged vegetables fruits	4	6	67%
fresh fruit → refrigerated	3	6	50%
crackers fresh fruit → refrigerated	**3**	**3**	**100%**
crackers refrigerated → fresh fruit	**3**	**3**	**100%**
fresh fruit frozen appetizer sides → packaged vegetables fruits	**3**	**3**	**100%**
fresh fruit crackers → refrigerated	**3**	**3**	**100%**
crackers → fresh fruit	**3**	**3**	**100%**

Table 9 validates the final association rules using the Lift Ratio. A lift ratio greater than 1 indicates that the rule is significant and helpful in making product recommendations.

Table 8 Validate Final Association Rule with Lift Ratio

Rule	**xy**	**x**	**y**	**Lift**
crackers → fresh fruit	**3**	**3**	6	1,67
crackers → refrigerated	**3**	**3**	3	3,33
…..	…..	…..	…..	…..
fresh fruit crackers → refrigerated	**3**	**3**	3	3,33

3. RESULTS AND DISCUSSION

In modeling and implementing this research, we will use a rapid miner to calculate the database obtained using the apriori algorithm to obtain association rules. Then, implement the apriori algorithm to make product recommendations on e-commerce websites on the product details page when users shop. In this study, due to the limitations of Microsoft Excel, the number of transaction IDs that can be processed is only 103,761 transaction ID. However, in its application, the researcher only limits it by taking the last 5000 transaction ID to prove whether the product recommendation system runs dynamically when 500 new transaction ID are added, which can be seen in table 10 where product ID 55,56,57 are products added to 5000 transactions so that if there is a new transaction entry, the a priori process will continue by taking the last 5000 transaction id. The results of the product recommendations have met the previously set support and confidence values of at least 5% support and at least 40% confidence. On the e-commerce website that is made, if the product selected when shopping does not meet the predetermined support and confidence values, it will display the same recommended product as the product category, which can be seen in Figure 5.

Table 9 Data Testing

Premises	Conclusion	Support	Confidence	Lift
115	83	0,085	0,451	1,044
24, 21	83, 123	0,068	0,453	1,989
21	24, 83	0,101	0,454	1,461
123, 84	120	0,050	0,473	1,914
24	123	0,267	0,476	1,332
56	**57, 55**	**0,095**	**0,919**	**9,633**
55	**56**	**0,095**	**0,973**	**9,371**
55	**57, 56**	**0,095**	**0,973**	**10,174**
57, 56	**55**	**0,095**	**0,998**	**10,174**
57, 55	**56**	**0,095**	**1,000**	**9,633**
56, 55	**57**	**0,095**	**1,000**	**8,165**

Table 10 presents the data testing results, including premises, conclusions, support values, confidence values, and lift ratios. These metrics were obtained through several steps. The premises represent specific combinations of items that frequently appear together in transactions, and the support values indicate the proportion of transactions that include these item combinations. High support values suggest that the items are commonly purchased together. Confidence values measure the likelihood that a particular item will be purchased given that another item has been purchased, helping to identify strong association rules where high confidence values indicate a substantial likelihood of co-purchase. The lift ratio is used to evaluate the strength and significance of the association rules; a lift ratio greater than one indicates that the occurrence of the items together is higher than expected if the items were independent, suggesting a meaningful association between the items.

Figure 5 illustrates the implementation results of the product recommendation system on the e-commerce product detail page displaying shaving needs. This figure shows the leading products and relevant product recommendations based on prior transaction data analysis using the apriori algorithm. The system provides dynamic product recommendations in two scenarios: one with various recommended products, including health care, personal care, and household items, and another with fewer but more specific recommendations. These results demonstrate the system's effectiveness in offering relevant products that are strongly associated with the leading products, thereby enhancing the user shopping experience. By displaying products with solid associations, the system not only helps increase customer satisfaction but also has the potential to boost sales, proving that transaction data analysis can be used to generate more accurate and beneficial product recommendations.

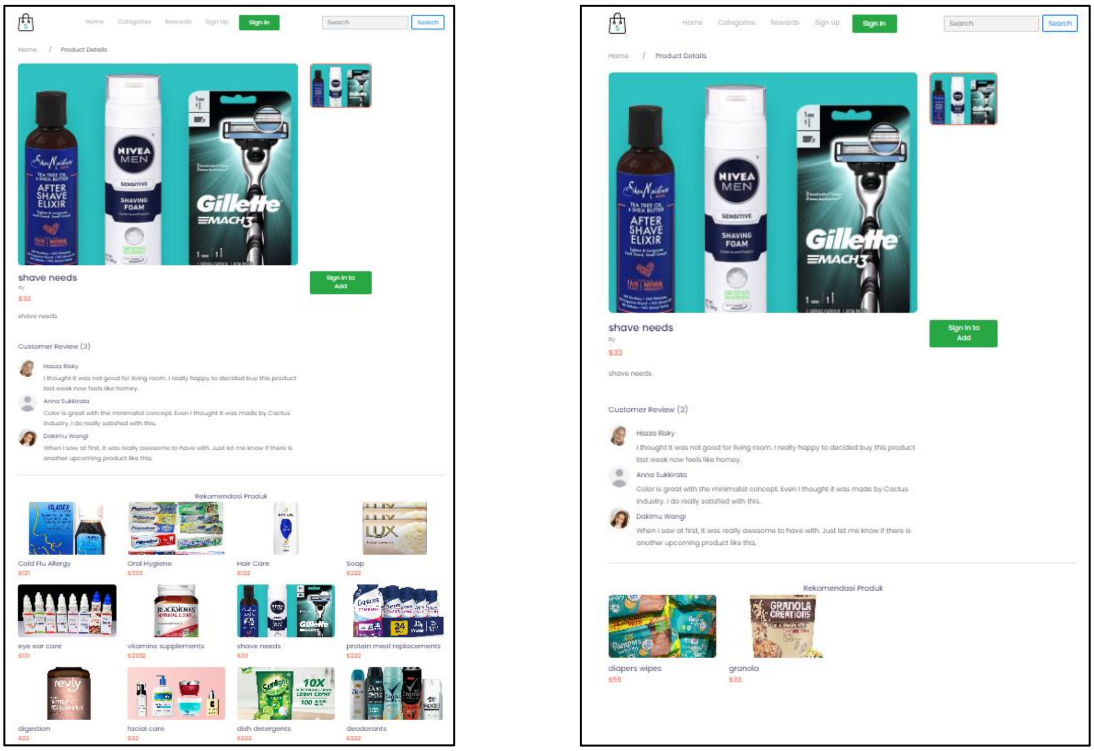

Figure 5 Recommendation Product If It Has Rules-Right, Has No Rules-Left

The testing phase revealed several key findings. The association rules generated by the apriori algorithm consistently showed high support and confidence values, indicating solid relationships between specific products. When implemented on the e-commerce platform, the system provided practical real-time product recommendations, enhancing the user shopping experience by suggesting complementary products based on recent sales trends. The lift ratio analysis demonstrated that most association rules had a lift ratio greater than 1, confirming the significance and usefulness of the generated rules. This positive correlation indicates that the recommended products are highly relevant to the customers' purchasing patterns. This result aligns with previous studies indicating that higher lift ratios correlate with improved recommendation accuracy, thereby validating the use of the apriori algorithm in e-commerce settings.

The iterative process of testing and refining the system led to several improvements. Improved data preprocessing techniques ensured that the data used for analysis was clean and relevant, enhancing the accuracy of the generated rules. Adjusting the minimum support and confidence thresholds helped generate more precise and meaningful association rules. Incorporating user feedback allowed for continuous improvement of the recommendation system, making it more adaptive to changing customer preferences. These iterative enhancements highlight the importance of a robust data pipeline and the need for adaptive algorithms that can evolve with user behavior.

The results from comprehensive testing, implementation, and evaluation demonstrate the effectiveness of utilizing market basket analysis with the apriori algorithm for product recommendation systems in e-commerce. The high support, confidence values, and positive lift ratios confirm the system's ability to provide relevant and valuable recommendations. These findings suggest that the developed system can significantly enhance the online shopping experience, improve customer satisfaction, and increase sales. Future research could explore integrating more sophisticated machine learning models to refine recommendation accuracy further and incorporate additional contextual data to personalize recommendations even further.

4. CONCLUSIONS

Through comprehensive testing, implementation, and evaluation, it has been ascertained that utilizing market basket analysis with the apriori algorithm is an incredibly effective method for assisting online shoppers in their product selection process. This cutting-edge system provides shoppers with dynamic product recommendations based on the latest sales data, enabling sellers to understand their customers' shopping behaviors better. The comprehensive testing phase involved evaluating the accuracy of the generated association rules, which consistently showed high support and confidence values. During the implementation phase, the system was integrated into an e-commerce platform that successfully provided real-time recommendations, enhancing the user shopping experience. The evaluation phase included the Lift Ratio correlation assessment, which consistently yielded values above 1, indicating high relevance and usefulness in the final association rules. These results demonstrate the system's robustness and effectiveness in improving product recommendations, potentially increasing customer satisfaction and sales.

REFERENCES

[1] A. Setiawan and R. Mulyanti, “Market Basket Analysis dengan Algoritma Apriori pada Ecommerce Toko Busana Muslim Trendy (Market Basket Analysis with Apriori Algorithms in Ecommerce Trendy Muslim Clothing Stores),” *JUITA : Jurnal Informatika*, vol. 8, no. 1, pp. 11–18, May 2020.

[2] C. S. Fatoni, E. Utami, and F. W. Wibowo, “Online Store Product Recommendation System Use Apriori Method,” in *Journal of Physics: Conference Series*, Institute of Physics Publishing, Dec. 2018. doi: 10.1088/1742-6596/1140/1/012034.

[3] BANK INDONESIA, “BANGKIT DAN OPTIMIS : SINERGI DAN INOVASI UNTUK PEMULIHAN EKONOMI,” *BANK INDONESIA*, JAKARTA, pp. 1–61, Nov. 24, 2021.

[4] M. Sholik and A. Salam, “Implementasi Algoritma Apriori untuk Mencari Asosiasi Barang yang Dijual di E-commerce OrderMas Implementation of Apriori Algorithm to Find Items Association That Sold at OrderMas E-commerce,” *Jurnal Teknologi Informasi* , vol. 17, no. 2, pp. 158–170, May 2018, doi: https://doi.org/10.33633/tc.v17i2.1656.

[5] D. I. Setiani, F. Nasafi, T. Mauritsius, E. R. Kaburuan, and R. Jayadi, “Data mining implementation with association method and apriori algorithm for store display design in home center indonesia,” *ICIC Express Letters*, vol. 15, no. 11, pp. 1221–1226, Nov. 2021, doi: 10.24507/icicel.15.11.1221.

[6] A. Maske and B. Joglekar, “Survey on Frequent Item-Set Mining Approaches in Market Basket Analysis,” in *2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA)*, India: 2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA), 2018.

[7] E. Ozgormus and A. E. Smith, “A data-driven approach to grocery store block layout,” *Comput Ind Eng*, vol. 139, Jan. 2020, doi: 10.1016/j.cie.2018.12.009.

[8] M. Marsono, “Penerapan Data Mining Pengaturan Pola Tata Letak Barang Pada Berkah Swalayan Untuk Strategi Penjualan Menggunakan Algoritma Apriori,” *InfoTekJar (Jurnal Nasional Informatika dan Teknologi Jaringan)*, vol. 3, no. 2, pp. 170–175, Feb. 2019, doi: 10.30743/infotekjar.v3i2.908.

[9] T. Joe, R. Sreejith, and K. Sekar, “Optimization of store layout using market basket analysis,” *International Journal of Recent Technology and Engineering*, vol. 8, no. 2, pp. 6459–6463, Jul. 2019, doi: 10.35940/ijrte.B2207.078219.

[10] R. Purwaningsih, L. Tri Utami, Y. Widharto, and N. Susanto, "REDESAIN TATA LETAK PRODUK DI SUPERMARKET BERDASARKAN PERILAKU PEMBELIAN DENGAN METODE MARKET BASKET ANALYSIS," *Jurnal Teknik Industri*, vol. 15, no. 3, 2020.

[11] D. I. Setiani, F. Nasafi, T. Mauritsius, E. R. Kaburuan, and R. Jayadi, "Data mining implementation with association method and apriori algorithm for store display design in home center indonesia," *ICIC Express Letters*, vol. 15, no. 11, pp. 1221–1226, Nov. 2021, doi: 10.24507/icicel.15.11.1221.

[12] S. Halim, T. Octavia, and C. Alianto, "Designing facility layout of an amusement arcade using market basket analysis," *Procedia Comput Sci*, vol. 161, pp. 623–629, 2019, doi: 10.1016/j.procs.2019.11.165.

[13] M. Loukili, F. Messaoudi, and M. El Ghazi, "Machine learning based recommender system for e-commerce," *IAES International Journal of Artificial Intelligence*, vol. 12, no. 4, pp. 1803–1811, Dec. 2023, doi: 10.11591/ijai.v12.i4.pp1803-1811.

[14] S. Jabbour, F. E. El Mazouri, and L. Sais, "Mining Negatives Association Rules Using Constraints," *Procedia Comput Sci*, vol. 127, pp. 481–488, Jan. 2018, doi: 10.1016/J.PROCS.2018.01.146.

[15] C. Chapman and E. M. Feit, "Association Rules for Market Basket Analysis," in *R For Marketing Research and Analytics*, C. Chapman and E. M. Feit, Eds., Cham: Springer International Publishing, 2019, pp. 341–362. doi: 10.1007/978-3-030-14316-9_12.

[16] S. Huber, H. Wiemer, D. Schneider, and S. Ihlenfeldt, "DMME: Data mining methodology for engineering applications – a holistic extension to the CRISP-DM model," *Procedia CIRP*, vol. 79, pp. 403–408, 2019, doi: https://doi.org/10.1016/j.procir.2019.02.106.

[17] Y. Apridonal M, W. Choiriah, and A. Akmal, "PENERAPAN DATA MINING MENGGUNAKAN METODE ASSICIATION RULE DENGAN ALGORITMA APRIORI UNTUK ANALISA POLA PENJUALAN BARANG," *JURTEKSI (Jurnal Teknologi dan Sistem Informasi)*, vol. 5, no. 2, pp. 193–198, Jul. 2019, doi: 10.33330/jurteksi.v5i2.362.

[18] H. K. Lin, C. H. Hsieh, N. C. Wei, and Y. C. Peng, "Association rules mining in R for product performance management in industry 4.0," *Procedia CIRP*, vol. 83, pp. 699–704, 2019, doi: 10.1016/j.procir.2019.04.099.

[19] Y. A. Ünvan, "Market basket analysis with association rules," *Commun Stat Theory Methods*, vol. 50, no. 7, pp. 1615–1628, 2021, doi: 10.1080/03610926.2020.1716255.