Journal of Applied Informatics and Computing (JAIC)
Vol.9, No.2, April 2025, pp. 556~563
e-ISSN: 2548-6861

Implementation of Apriori Algorithm in Identifying Purchase Relationships at Bluder Cokro Pakuwon Mall

Angeline Ivana [1]*, Indra Maryati [2]*
* Information Systems for Business, Universitas Ciputra Surabaya
angelineivana01@student.ciputra.ac.id [1], indra.maryati@ciputra.ac.id [2]

Article Info

Article history:

Received 2025-02-17
Revised 2025-03-12
Accepted 2025-03-13

Keyword:

Bluder Cokro,
Apriori Algorithm,
CRISP-DM,
Association Rules,
Data Mining.

ABSTRACT

Bluder Cokro Store, located at Pakuwon Mall, specializes in traditional bluder bread with a wide range of flavor variations. This study aims to identify consumer purchasing patterns at the store to enhance promotional strategies and optimize product placement. The research applies the Cross Industry Standard Process for Data Mining (CRISP-DM) methodology, which includes phases such as business understanding, data understanding, data preparation, modeling, evaluation, and deployment. The dataset used consists of 4,371 transactions from October to December 2024. This study uses the Apriori algorithm to find patterns of association between products, with the goal of determining the scope of correlation between products and frequently co- purchased items. The results reveal nine significant association rules, with the strongest relationship observed between coklat keju and keju, having a support value of 0.100394 and a lift of 1.31. These findings indicate that strategic product placement and bundling promotions can enhance sales performance. Optimizing the store layout by placing coklat keju near coklat can increase purchase likelihood, while targeted discounts, such as "Buy coklat keju, get 10% off keju," can drive transaction values. This study serves as a recommendation framework rather than an experimental validation, offering insights on how transaction data and association rule mining can inform business decisions. The findings offer actionable insights for improving store layouts and promotional effectiveness, making this research valuable for retailers.

I. Introduction

Bluder Cokro Store at Pakuwon Mall is a business specializing in traditional bluder bread with various flavor variations. The purpose of this study is to identify consumer purchasing patterns at Bluder Cokro Store, Pakuwon Mall, enabling the store to design more effective promotions and optimize product placement. The store often faces challenges in determining which products should be placed close together to attract customer interest [1].

Currently, the promotional strategies implemented have not been effective, as the promoted products often lack a strong correlation with each other [2]. This results in customers being less interested in purchasing more than one type of product during the promotion, impacting sales targets and inventory management. For instance, the chocolate-flavored bluder, which is highly favored by customers, is often promoted alongside a variant that does not share a similar purchasing pattern, making it less appealing for additional or bundled purchases [3].

To address this issue, Bluder Cokro Store needs to utilize existing sales transaction data to uncover customer purchasing patterns [4]. One approach that can be applied is Data Mining using the Association Rule method with the Apriori algorithm, first developed by Agrawal and Krishnan in 1994 [5].

Apriori systematically generates frequent itemsets and mines association rules by scanning the database multiple times and applying a pruning mechanism to eliminate unimportant itemsets. The algorithm is particularly effective in scenarios where datasets are relatively small and sparse, such as this study’s dataset of 4,371 transactions.

FP-Growth, introduced by Han Jiawei in 2000, improves efficiency by scanning the database only twice and avoiding candidate set generation. However, it does not explicitly mine association rules. Apriori, on the other hand, provides a clear step-by-step association rule extraction, making it more intuitive and actionable for business owners who may not have a technical background in data mining. A study demonstrated how Apriori helps e-commerce businesses analyze customer purchasing habits, enabling targeted promotions and personalized recommendations without requiring advanced data mining expertise [6].

While FP-Growth is more efficient for large datasets, studies have shown that Apriori produces stronger association rules despite requiring longer execution times. A study analyzing sales transactions at Indomaret Tanjung Anom found that the Apriori algorithm generated 26 lift ratio association rules, with higher confidence values compared to FP-Growth, proving its effectiveness in discovering stronger and more reliable purchasing patterns [5]. To mitigate Apriori's longer processing time, a study implements the algorithm using Python, a programming language widely used in data analysis and data science. The use of Python ensures that the algorithm runs efficiently while maintaining accuracy in generating association rules [7].

Despite its strengths, Apriori has limitations. Its computational complexity increases exponentially with larger datasets due to the need to generate and evaluate multiple candidate itemsets. However, since this study deals with a moderate dataset of 4,371 transactions, the algorithm remains computationally feasible [6]. Studies have shown that for moderate-sized datasets, the accuracy of association rules generated by Apriori and FP-Growth is comparable. In a study analyzing customer purchase patterns, both Apriori and FP-Growth produced the same frequent itemset combinations, with a support value of 27,7% and a confidence value of 98% [8]. This demonstrates that despite FP-Growth's efficiency advantages in large datasets, Apriori remains a viable and reliable choice when working with a dataset of this size. As a result, the analysis can help identify relationships between bluder variants that are frequently purchased together, allowing for more effective product combinations in promotions [9].

By understanding the relationships between frequently purchased flavor variants, the store can develop more relevant and engaging promotional strategies for customers [10]. This study aims to identify consumer purchasing patterns at Bluder Cokro Store, Pakuwon Mall, so that the store can design more effective promotions and optimize product placement.

II. METHODOLOGY

This study applies the Apriori algorithm to sales data, using a case study of a bakery that sells Bluder bread, specifically Bluder Cokro Pakuwon Mall. The store offers a total of 17 different Bluder flavors. Data collection techniques include interviews with the store manager and direct observation of the sales process.

This research will adopt the CRISP-DM (Cross-Industry Standard Process for Data Mining) methodology [11], which consists of six main stages:

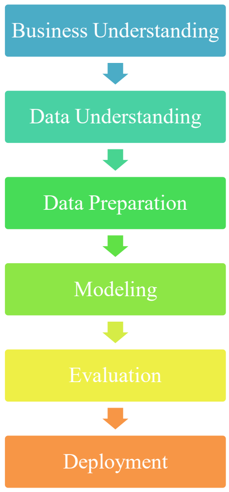

Figure 1. CRISP-DM Diagram

A. Business Understanding

The first step is to understand the business objectives of this research. This stage also involves identifying customer shopping patterns. By analyzing purchasing patterns, Bluder Cokro can plan product placement and promotional strategies more efficiently [8].

B. Data Understanding

The data used in this research consists of sales transaction records from October to December 2024. This data includes information on the types of bread purchased, purchase quantities, and transaction times. The data is collected through the cashier system and analyzed to understand its structure and characteristics.

C. Data Preparation

The data collected from the cashier system is then processed to ensure its consistency and quality. This process includes the following steps:

1) *Data Cleaning*: This step is to remove null values and correct formatting errors in transaction records.

2) *Data Transformation:* This step is to convert the data into a format suitable for analysis using the Apriori algorithm, ensuring the dataset consists only of transactions represented as item lists [12].

D. Modeling

At this stage, the Apriori algorithm is implemented using Python to discover association patterns within purchase data. To determine the optimal threshold values for support and confidence, a grid search approach is applied. This ensures that the chosen parameters are based on systematic evaluation rather than arbitrary selection. The grid search iterates through multiple combinations of minimum support and confidence values, selecting the combination that generates the most meaningful association rules. The following steps break down how these values will be applied in the Apriori algorithm:

1) *Calculating Support*: This step is to determine the support value for each itemset, which is calculated by dividing the number of transactions containing the item by the total transactions and then multiplying by 100%.

2) *Calculating Confidence*: This step is to compute the confidence for each generated association rule by dividing the number of transactions containing both U and T by the number of transactions containing T and then multiplying by 100%.

3) *Utilizing Python Libraries:* In this step, the Apriori algorithm is implemented using Python libraries such as pandas and mlxtend [13]. Python is chosen as the implementation language due to its efficiency in memory and CPU usage, as well as its extensive support for data analysis [7]. The pandas library is used for data manipulation, while mlxtend provides functions for applying the Apriori algorithm and generating association rules.

E. Evaluation

Once the model is built, the next step is to evaluate the generated association rules using multiple metrics:

1) *Lift Ratio*: Lift is calculated by dividing the confidence of a rule by the probability of the consequent occurring independently, which is equivalent to dividing the probability of both items appearing together by the product of their individual probabilities [6]. A lift value greater than 1 indicates a strong association, meaning two products frequently appear together, while a value equal to 1 signifies independence, and a value less than 1 suggests a negative correlation. A higher lift value suggests a meaningful purchasing pattern, useful for product recommendations and marketing strategies. For example, a study selected the association rule with the highest lift value of 1.25 (cigarettes and lighters) because a higher lift indicates a stronger purchasing correlation, ensuring that the identified item pairs have a significantly higher likelihood of being bought together than expected by chance [7]. It validates transactional processes and confirms whether a specific product is truly bought alongside another [8].

2) *Leverage and Conviction*: Leverage measures the difference between the observed frequency of an item pair and the expected frequency if the two items were independent. A positive leverage value indicates that the items co-occur more often than expected by chance. Conviction evaluates how much more likely the antecedent item leads to the consequent item than if they were independent. A conviction value greater than 1 suggests a meaningful relationship between the items [14].

F. Deployment

In the deployment phase, the insights derived from data analysis are translated into actionable business strategies. This involves compiling a report that outlines recommended approaches for product placement and marketing strategies based on the discovered association rules. These recommendations serve as a basis for decision-making by the management team.

Before implementation, it is essential to validate the proposed strategies by consulting the owner, to assess their feasibility in real-world operations. Additionally, businesses may conduct pilot testing, such as A/B testing and experimental marketing plans, to measure the effectiveness of these strategies before full-scale adoption [15].

III. RESULTS AND DISCUSSION

The results obtained include the most popular Bluder item combinations and the association rules derived from these combinations based on the collected transaction data.

A. Business Understanding

This study focuses on Bluder Cokro’s primary objective of optimizing product placement and promotional strategies. Currently, the store relies on a random approach for both product placement and promotions, without leveraging data-driven insights.

By utilizing past transaction data, the store can gain insights into customer purchasing patterns. These insights enable Bluder Cokro to refine its promotional strategies by offering targeted discounts and bundling frequently co-purchased products, rather than relying on random promotions. Moreover, optimizing product placement based on data will help improve customer shopping experiences and encourage impulse purchases. With this knowledge, Bluder Cokro can create more accurate product placement and implement effective promotional strategies.

B. Data Understanding

The data used in this study consists of sales transaction records from October to December 2024, totaling 4,371 transactions. These transaction records serve as the foundation for the data mining analysis. However, since the dataset covers only a three-month period, it may not fully represent year-round shopping patterns. While external factors could have influenced purchasing behaviour, their impact appears to be limited. Customer preferences for certain products remained stable throughout the period, even during the festive season. Although sales increased in December, the growth was not concentrated on specific products but rather spread across all items. Additionally, the highest-selling products remained consistent, suggesting that seasonal trends had minimal influence on product demand. Moreover, market trends during this period did not undergo significant changes, meaning broader consumer behaviour shifts, such as dietary habits or economic conditions, were unlikely to have affected purchasing patterns.

TABLE I
SALES TRANSACTION DATA

No	Date	Type	Payment
1	01-10-2024	ORIGINAL	Cash
2	01-10-2024	KEJU LUMER	Cash
3	01-10-2024	HAZELNUT	BCA QR
…			
4371	31-12-2024	ORIGINAL, ORIGINAL, ORIGINAL, KISMIS	Cash

The raw dataset is in CSV format, where each row represents a single transaction. The key columns in the dataset include No, Date, Type, and Payment.

TABLE II
VARIABLE DESCRIPTION

Variable	Description
No	Sequential number
Date	Sales transaction date
Type	Bluder flavor type
Payment	Payment method

Before conducting the analysis, the dataset underwent a data cleaning process using Power Query in Microsoft Excel to ensure accuracy and consistency.

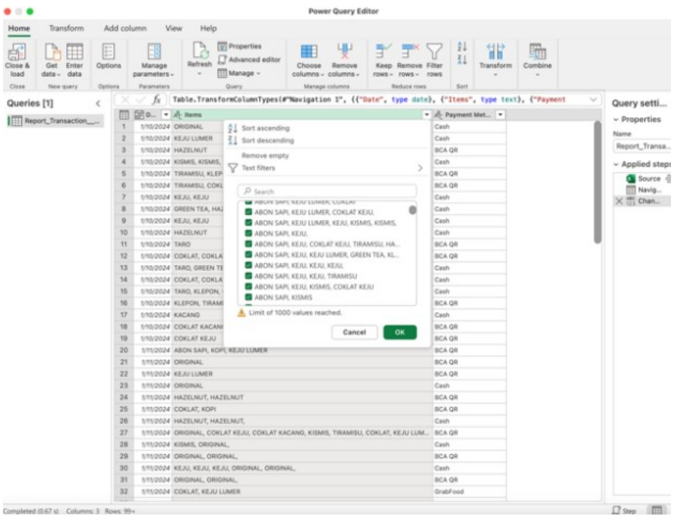

Figure 2. Data cleaning

The following cleaning steps were performed:

1) *Removing Null Values*: Transactions with missing data were removed to maintain data integrity.

2) *Correcting Column Data Types:* Columns such as "Date" were converted into the correct format (Date format), while categorical columns like "Type" and "Payment" were properly classified as text.

3) *Fixing Formatting Errors:* Any inconsistencies, such as incorrect spacing or misformatted entries, were corrected to standardize the dataset.

To gain a deeper understanding of the dataset, an initial analysis was conducted to determine key characteristics, such as the total number of unique items and the average number of items per transaction.

One important aspect of this analysis is identifying the total number of unique items available in the transactions. This helps in understanding product diversity and potential associations between items.

```
df['Items'] = df['Items'].apply(lambda x: x.split(',')
                                if isinstance(x, str) else [])
unique_items = pd.Series([item.strip() for sublist in df['Items']
                          for item in sublist]).nunique()

print(f"Total unique items in the dataset: {unique_items}")
Total unique items in the dataset: 18
```

Figure 3. Calculating total unique items

There are 18 of total unique items in this dataset, which are ABON SAPI, COKLAT, COKLAT KACANG, COKLAT KEJU, GREEN TEA, HAZELNUT, KACANG, KEJU, KEJU LUMER, KISMIS, KLEPON, KOPI, KRUMPUL, ORIGINAL, SOBEK, TARO, and TIRAMISU.

To further analyze purchasing behaviour, the average number of items per transaction was calculated. This metric provides insights into how many products customers typically purchase in a single transaction, which can help in understanding buying patterns and optimizing sales strategies.

```
avg_items_per_transaction = df['Items'].apply(len).mean()
print(f"Average number of items per transaction: {avg_items_per_transaction:.2f}")
Average number of items per transaction: 3.67
```

Figure 4. Calculating average number of items per transaction

With an average of 3.67 items per transaction, this suggests that customers often purchase multiple products at once, indicating potential opportunities for product bundling or promotional strategies.

C. Data Preparation

From all available transaction data, only transactions containing more than one type of Bluder will be used in this study. This is because the research focuses on identifying relationships between different items. Therefore, a mapping of purchased item types is conducted to analyze these associations effectively.

```
df_items = df['Items'].apply(lambda x: [item.strip()
for item in x.split(',') if item.strip() != ''])

df_items = df_items[df_items.apply(lambda x: len(x)
> 1)].reset_index(drop=True)
```

TABLE III
PREPROCESSED TRANSACTION DATA

No	Type
1	KISMIS, KISMIS, KOPI, ABON SAPI, KLEPON, KLEPON, HAZELNUT, HAZELNUT, KACANG, KEJU, KEJU, COKLAT KACANG , KISMIS, KISMIS, KEJU, GREEN TEA, KEJU LUMER, COKLAT KEJU, COKLAT KEJU, TARO, TIRAMISU

2	TIRAMISU, KLEPON, KEJU, COKLAT, KOPI, ABON SAPI, TARO
3	TIRAMISU, COKLAT, KEJU LUMER
…	
3047	ORIGINAL, ORIGINAL, ORIGINAL, KISMIS

From this data, each transaction is needed to be converted into a binary matrix form, where each column represents an item and each row represents a transaction. If an item is present in a transaction, its value will be 1, and if it is not present, the value will be 0.

```
from sklearn.preprocessing import
MultiLabelBinarizer
mlb = MultiLabelBinarizer()

binary_df =
pd.DataFrame(mlb.fit_transform(df_items),
columns=mlb.classes_)
```

TABLE IV
BINARY DATAFRAME

No	Abon Sapi	Coklat	Coklat Kacang	…	Taro	Tira misu
1	1	0	1		1	1
2	1	1	0		1	1
3	0	1	0		0	1
…						
3047	0	0	0		0	0

D. Modeling

First, the minimum support and minimum confidence values are set. This threshold was not chosen arbitrarily but determined through a grid search approach, which systematically evaluates different parameter combinations to identify the most meaningful association rules.

```
from itertools import product

support_values = [0.1, 0.2, 0.3]
confidence_values = [0.3, 0.5, 0.7]

best_rules = None
best_params = None
max_rules = 0

for min_sup, min_conf in product(support_values, confidence_values):
    frequent_itemsets = apriori(binary_df, min_support=min_sup, use_colnames=True)
    rules = association_rules(frequent_itemsets, metric="confidence", min_threshold=min_conf)

    if len(rules) > max_rules:
        best_rules = rules
        best_params = (min_sup, min_conf)
        max_rules = len(rules)

print(f"Best Parameters: min_support={best_params[0]}, min_confidence={best_params[1]}")
print(f"Total Rules Found: {max_rules}")
```
```
Best Parameters: min_support=0.1, min_confidence=0.3
Total Rules Found: 9
```

Figure 5. Grid Search for Apriori Optimization

Through this process, the best parameter values obtained are minimum support of 0.1 (10%) and minimum confidence of 0.3 (30%), which resulted in 9 association rules. Based on the grid search results, a minimum support of 0.1 (10%) was applied to ensure only frequently purchased items are considered. This threshold filters out infrequently purchased items while ensuring that only frequently bought products are considered [2]. A minimum confidence of 30% (0.3) is set based on the grid search, ensuring that the relationships between itemsets are statistically significant. This means that the likelihood of an item being purchased when another is bought is strong enough to be useful for decision-making, such as product recommendations or promotional strategies [16]. These values ensure that only frequently purchased products are included in the analysis while maintaining statistical significance in the relationships identified.

Next, the Apriori algorithm will be used to find association patterns. This algorithm will generate frequent itemsets and association rules with a minimum support of 0.1 (10%) to ensure that only frequently purchased items are displayed [2].

```
from mlxtend.frequent_patterns import apriori,
association_rules

binary_df = binary_df.astype(bool)

frequent_itemsets = apriori(binary_df,
min_support=0.1, use_colnames=True)

itemsets_table = frequent_itemsets[['itemsets',
'support']]
itemsets_table =
itemsets_table.sort_values(by='support',
ascending=False).reset_index(drop=True)
```

TABEL V
FREQUENT ITEMSETS

No	Itemset	Support
1	(ORIGINAL)	0.415026
2	(COKLAT)	0.394357
3	(COKLAT KEJU)	0.291995
4	(KEJU)	0.262467
5	(KISMIS)	0.258858
6	(KEJU LUMER)	0.233596
7	(HAZELNUT)	0.223097
8	(COKLAT, ORIGINAL)	0.155184
9	(TIRAMISU)	0.141076
10	(KLEPON)	0.131562
11	(COKLAT, COKLAT KEJU)	0.129625
12	(COKLAT, KEJU)	0.107612
13	(KISMIS, COKLAT)	0.106299
14	(KISMIS, ORIGINAL)	0.102034
15	(KEJU, COKLAT KEJU)	0.100394

After identifying the most frequent itemsets, the next step is to generate association rules using the Apriori algorithm. These rules are filtered based on a minimum confidence of 30% (0.3) [16]. Similar to support, the confidence threshold was determined using grid search optimization, not randomly assigned. This process guarantees that the extracted rules provide valuable insights for product recommendations and marketing strategies.

```
rules = association_rules(frequent_itemsets,
metric="confidence", min_threshold=0.3)
association_table = rules[['antecedents',
'consequents', 'support', 'confidence', 'lift',
'leverage', 'conviction']]
association_table['antecedents'] =
association_table['antecedents'].apply(lambda x: ',
'.join(list(x)))
association_table['consequents'] =
association_table['consequents'].apply(lambda x: ',
'.join(list(x)))
association_table =
association_table.sort_values(by='support',
ascending=False).reset_index(drop=True)
display_table =
association_table[['antecedents', 'consequents', 'su
pport', 'confidence']]
```

TABLE VI
ASSOCIATION RULES

No	Antecedent	Consequent	Supp	Conf
1	COKLAT	ORIGINAL	0.155184	0.393511
2	ORIGINAL	COKLAT	0.155184	0.373913
3	COKLAT KEJU	COKLAT	0.129265	0.442697
4	COKLAT	COKLAT KEJU	0.129265	0.327787
5	KEJU	COKLAT	0.107612	0.410000
6	KISMIS	COKLAT	0.106299	0.410646
7	KISMIS	ORIGINAL	0.102034	0.394170
8	COKLAT KEJU	KEJU	0.100394	0.343820
9	KEJU	COKLAT KEJU	0.100394	0.382500

Antecedents and consequents represent the association between products found in customer transaction data [5]. Confidence indicates the probability that the consequent (associated product) is purchased when the antecedent (initial product) is bought.

E. Evaluation

Next, the next step is to evaluate the generated association rules using the lift ratio. The heatmap is generated by pivoting the association table, with antecedents on the y-axis, consequents on the x-axis, and lift values as the color intensity. Higher lift values indicate stronger associations between item pairs.

```
import seaborn as sns
import matplotlib.pyplot as plt
heatmap_data =
association_table.pivot(index='antecedents',
columns='consequents', values='lift')
plt.figure(figsize=(8, 6))
sns.heatmap(heatmap_data, annot=True, cmap="YlGnBu",
linewidths=0.5, fmt=".2f")
plt.title("Heatmap of Association Rules (Lift
Values)")
plt.xlabel("Consequents")
plt.ylabel("Antecedents")
plt.show()
```

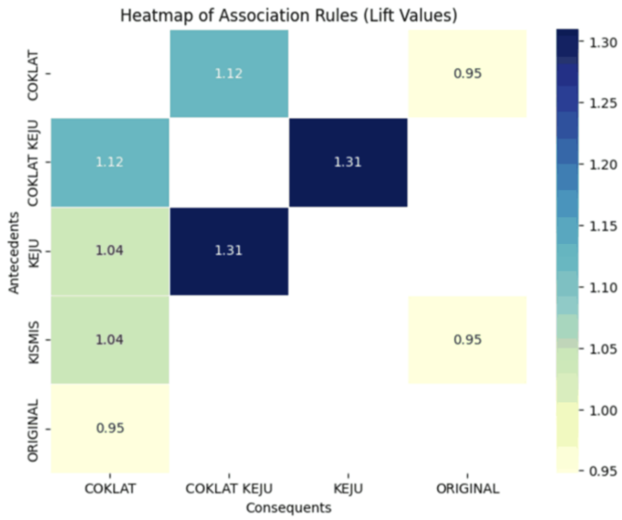

Figure 6. Heatmap of lift values for association rules

This heatmap visually represents the lift values of the generated association rules. Darker or more intense colors indicate higher lift values, meaning stronger correlations between products. By analyzing the heatmap, it is evident that the strongest relationship is found between COKLAT KEJU → KEJU and KEJU → COKLAT KEJU, with a lift value of 1.31. This indicates that the purchase of KEJU significantly increases the likelihood of purchasing COKLAT KEJU, and vice versa. The association between COKLAT KEJU and COKLAT was also considered strong, with a lift of 1.12 and a support of 0.129265. However, while a lift above 1 suggests a positive relationship, a value close to 1 signifies a relatively weak association [7]. In contrast, the higher lift value (1.31) observed in the heatmap for COKLAT KEJU and KEJU indicates a more significant purchasing pattern.

```
association_table_sorted =
association_table.sort_values(by='lift',
ascending=False)
top_3_rules =
association_table_sorted.head(3).reset_index(drop=True)

display_table = top_3_rules[['antecedents',
'consequents', 'lift', 'leverage', 'conviction']]
```

TABLE VII
TOP THREE HIGHEST LIFT VALUES

No	Antecedent	Consequent	Lift	Lever	Convi
1	KEJU	COKLAT KEJU	1.31	0.02	1.15
2	COKLAT KEJU	KEJU	1.31	0.02	1.12
3	COKLAT	COKLAT KEJU	1.12	0.01	1.05

The top three association rules with the highest lift values are summarized in Table VII, showing the strongest product relationships. The highest lift value (1.31) is observed in the rules KEJU → COKLAT KEJU and COKLAT KEJU → KEJU, indicating a strong two-way association. This means that customers who purchase KEJU are significantly more likely to buy COKLAT KEJU, and vice versa.

A positive leverage value suggests that the items appear together more often than expected by chance, indicating a meaningful relationship. In this case, the highest leverage values (0.02) for KEJU → COKLAT KEJU and COKLAT KEJU → KEJU confirm that these item pairs frequently co-occur beyond random chance. Conviction indicates how strongly the presence of an antecedent affects the occurrence of the consequent. A conviction value greater than 1 suggests that the rule is meaningful. In this case, KEJU → COKLAT KEJU has a conviction of 1.15, meaning the absence of KEJU reduces the likelihood of purchasing COKLAT KEJU by 15%, and vice versa.

These findings reinforce the importance of KEJU and COKLAT KEJU as highly co-purchased items, which can be leveraged for cross-selling strategies, product bundling, and targeted promotions. By focusing on these strong associations, businesses can optimize product placement and marketing efforts to drive higher sales and customer satisfaction.

F. Deployment

The deployment phase involves applying the insights gained from the association rule mining process to enhance business operations and decision-making. However, it is important to note that this study only provides recommendations based on data analysis and does not present experimental validation. To confirm the effectiveness of these recommendations, A/B testing and data-driven marketing experiments should be conducted before full-scale implementation. During discussions with the business owner, it was confirmed that the products identified in the top association rules match Bluder Cokro's current best-selling products. This verification by the owner indicates that the discovered patterns accurately reflect actual purchasing behaviour. As a result, the business recognizes the potential value of these insights for marketing and product placement strategies, reinforcing confidence in the recommendations. The proposed strategies focus on two key areas:

1) *Product Placement*: Rearranging the store layout based on the discovered patterns can further enhance sales performance. Placing frequently co-purchased items in proximity or within themed sections can improve navigation and shopping efficiency [8]. Research from the IAENG International Journal of Computer Science provides empirical evidence that placing highly associated products near each other impacts sales [17]. The study analyzed store layouts optimized using association rules (ARs) and found that grouping frequently co-purchased items together improved the shopping experience and increased purchase frequency. To verify this recommendation, an A/B testing plan is suggested to measure the impact of product placement on sales performance [15]. The experiment would involve two groups: a test group, where frequently associated products (e.g., COKLAT KEJU and KEJU) are placed closer together or within themed sections, and a control group, where the store layout remains unchanged. The experiment should run for four until eight weeks to capture variations in shopping behaviour and ensure reliable data collection. Key metrics to evaluate the effectiveness of this strategy include the sales volume of affected products, revenue changes before and after the layout modification, and customer engagement levels, such as increased interaction with repositioned products. By analyzing these factors, the store can determine whether strategic product placement enhances visibility and encourages bundled purchases [18]. However, since this study only provides recommendations based on data analysis, further testing is essential to validate these findings before implementing the strategy on a broader scale.

2) *Marketing:* The Apriori algorithm provides valuable insights into customer purchasing patterns, allowing businesses to create data-driven promotional strategies that maximize sales [10]. Research has shown that data-driven bundling strategies are more effective than random promotions, as they enhance the customer shopping experience and increase the likelihood of purchase. For instance, a study on retail sales demonstrated that bundling frequently co-purchased items using association rule mining increased overall transaction value by up to 20% [19]. A study by Yan Fang et al. demonstrated that using optimized bundling based on frequent itemset analysis and genetic algorithms increased retailer income by 8.3%, from 39,364,142 to 42,645,327 [20]. To evaluate this approach, an experimental marketing plan is recommended, where one group of customers receives a targeted bundling promotion, such as "Buy COKLAT KEJU, get 10% off KEJU," based on the association rule analysis, while another group is offered a general discount on all bakery items without item-specific bundling. This promotion should run for a period of four to eight weeks to ensure that data is collected across different shopping patterns. The effectiveness of the strategy can be measured by analyzing key metrics, including the redemption rate of targeted bundling compared to general discounts, the increase in sales for associated products, and the impact on customer retention and purchase frequency. A study on retail sales demonstrated that bundling frequently co-purchased items using association rule mining increased overall transaction value by up to 20% [19]. Similarly, a study by Yan Fang et al. found that using optimized bundling based on frequent itemset analysis and genetic algorithms increased retailer income by 8.3%, from 39,364,142 to 42,645,327 [20]. These strategies appeal to different customer segments, including those who prefer convenience and those seeking cost savings [18]. For Bluder Cokro, these recommendations suggest that implementing Apriori-based bundling strategies could drive higher sales and improve customer retention. However, since this study does not present direct experimental results, further testing is necessary before adopting these strategies on a larger scale.

IV. CONCLUSION

This study examines the use of data mining techniques, specifically the Apriori algorithm, to analyze sales transaction records and uncover customer purchasing patterns. The data used consists of 4,371 transaction records from October to December 2024, which were processed and transformed to enable Apriori calculations through Python programming. The goal of the analysis was to identify frequently purchased product combinations, resulting in nine association rules. The strongest association was observed between COKLAT KEJU and KEJU, with a support value of 0.100394 and a lift of 1.31, indicating a positive relationship. Based on these findings, it is recommended that the store owner consider implementing strategic product placement and targeted bundling promotions to enhance sales performance. Optimizing store layouts by positioning frequently co-purchased items together can improve customer convenience, increase visibility, and encourage impulse purchases. The revised placement of COKLAT KEJU near KEJU aligns with established retail strategies that enhance product discoverability, potentially leading to higher transaction volumes. Additionally, targeted promotions, such as "Buy COKLAT KEJU, get 10% off KEJU," may incentivize purchases and increase overall sales revenue. Research on data-driven bundling strategies has suggested that targeted promotions based on frequent itemsets can increase transaction value by up to 20% and boost retailer income by 8.3%. However, it is important to note that these recommendations are derived solely from historical transaction data analysis and not from direct experimentation. This study does not provide experimentally validated results, and the effectiveness of these strategies may vary based on real-world implementation. Therefore, it is advised that the store owner conduct A/B testing or controlled marketing experiments to assess the actual impact of product placement and promotional strategies before fully adopting them. Additionally, to gain a more comprehensive understanding of year-round shopping patterns and minimize the influence of seasonal variations, future studies should analyze a full year's worth of transaction data. This approach would allow for more accurate identification of consistent purchasing behaviours and long-term market trends, ensuring that business strategies remain effective beyond the observed three-month period.

REFERENCES

[1] H. N. Utami, S. N. Wiyono and A. Nugraha, "Pengambilan Keputusan Pengadaan Produk Segar Dan Kinerja Layanan Suplier Di Ritel Supermarket: Sebuah Perspektif Business-To-Business," *Agricore: Jurnal Agribisnis Dan Sosial Ekonomi Pertanian UNPAD,* vol. 9, no. 1, pp. 48-61, 2024.

[2] Y. A. Alhillah, W. Priatna and A. Fitriyani, "Implementation of Apriori Algorithm for Determining Spare Parts Product Recommendation Packages," *Journal of Applied Informatics and Computing (JAIC),* vol. 7, no. 2, pp. 212-217, 2023.

[3] K. Martowinangun, D. J. S. Lestari and K. Karyadi, "Pengaruh Strategi Promosi Terhadap Peningkatan Penjualan Di Cv. Jaya Perkasa Motor Rancaekek Kabupaten Bandung," *Jurnal Co Management,* vol. 1, no. 1, pp. 139-152, 2023.

[4] S. Tualeka, F. Alameka and N. W. W. Sari, "Implementasi Data Mining Untuk Memprediksi Penjualan Dan Penempatan Stok Barang Pada Cv Pasti Jaya Houseware Dengan Menggunakan Algoritma Apriori," *SEMINASTIKA,* vol. 3, no. 1, pp. 115-123, 2021.

[5] M. H. Santoso, "Application of Association Rule Method Using Apriori Algorithm to Find Sales Patterns Case Study of Indomaret Tanjung Anom," *Brilliance: Research of Artificial Intelligence,* vol. 1, no. 2, pp. 54-65, 2021.

[6] H. Xie, "Research and Case Analysis of Apriori Algorithm Based on Mining Frequent Item-Sets," *Open Journal of Social Sciences,* vol. 9, no. 4, pp. 458-468, 2021.

[7] R. Takdirillah, "Penerapan Data Mining Menggunakan Algoritma Apriori Terhadap Data Transaksi Sebagai Pendukung Informasi Strategi Penjualan," *Edumatic: Jurnal Pendidikan Informatika,* vol. 4, no. 1, pp. 37-46, 2020.

[8] N. Purwati, Y. Pedliyansah, H. Kurniawan, S. Karnila and R. Herwanto, "Komparasi Metode Apriori dan FP-Growth Data Mining Untuk Mengetahui Pola Penjualan," *Jurnal Informatika: Jurnal pengembangan IT (JPIT),* vol. 8, no. 2, pp. 155-161, 2023.

[9] K. Dion, I. P. Satwika and W. N. Utami, "Analisis Transaksi Penjualan Barang Menggunakan Metode Apriori pada UD. Ayu Tirta Manis," *Jurnal Krisnadana,* vol. 1, no. 2, pp. 11-20, 2022.

[10] R. Anugrah, T. Widiharih and S. Sugito, "GUI R Untuk Analisis Keranjang Belanja Dengan Algoritma Apriori Pada Suatu Perusahaan E-Commerce," *Jurnal Gaussian,* vol. 11, no. 2, pp. 278-289, 2022.

[11] M. Fitriani, G. F. Nama and M. Mardiana, "Implementasi Association Rule Dengan Algoritma Apriori Pada Data Peminjaman Buku Upt Perpustakaan Universitas Lampung Menggunakan Metodologi CRISP-DM," *Jurnal Informatika dan Teknik Elektro Terapan (JITET),* vol. 10, no. 1, pp. 41-49, 2022.

[12] A. D. Kuswanto, A. R. Blessar, A. Goni, A. N. N. Sidiki, O. R. A. Haryu and H. A. Hamiki, "Penerapan Algoritma Apriori Dalam Analisis Keranjang Belanja Retail Di Wilayah Jawa Barat," *Saturnus : Jurnal Teknologi dan Sistem Informasi,* vol. 2, no. 3, pp. 139-150, 2024.

[13] E. Kaban, I. G. M. Darmawiguna and M. W. A. Kesiman, "Optimizing Customer Purchase Insights: Apriori Algorithm for Effective Product Bundle Recommendations," *Brilliance: Research of Artificial Intelligence,* vol. 4, no. 2, pp. 747-756, 2024.

[14] A. Silvanie, "Pencarian Frequent Itemset Dengan Algoritma Apriori Dan Python. Studi Kasus: Data Transaksi Penjualan Eceran Online di UK," *JUNIF Jurnal Nasional Informatika,* vol. 1, no. 2, pp. 103-113, 2020.

[15] M. Mandić, I. Gregurec and U. Vujović, "Measuring The Effectiveness Of Online Sales By Conducting A/B Testing," *Market-Tržište,* vol. 35, no. 2, pp. 223-249, 2023.

[16] S. Styawati, A. Nurkholis and K. N. Anjumi, "Analisis Pola Transaksi Pelanggan Menggunakan Algoritme Apriori," *Jurnal Sains Komputer & Informatika (J-SAKTI),* vol. 5, no. 2, pp. 619-626, 2021.

[17] F. J. M. Arboleda, G. Garani and A. F. A. Correa, "Supermarket Product Placement Strategies Based on Association Rules," *IAENG International Journal of Computer Science,* vol. 51, no. 6, pp. 650-662, 2024.

[18] A. Irfan and . I. G. A. K. G. Suasana, "The Effect of Bundling Strategy, Price Perception, and Brand Image on Purchase Decisions (Study on Local Fast Food Retailers in Bali Indonesia)," *American Journal of Humanities and Social Sciences Research (AJHSSR),* vol. 5, no. 6, pp. 295-299, 2021.

[19] K. A. Carroll, A. Samek and L. Zepeda, "Consumer Preference for Food Bundles under Cognitive Load: A Grocery Shopping Experiment," *Foods 2022,* vol. 11, no. 7, p. 973, 2022.

[20] Y. Fang, R. Wang, M. Guo and Y. Hou, "Product bundling for online supermarkets by frequent itemset mining and optimization approach mining and optimization approach," *Procedia Computer Science,* vol. 207, pp. 4434-4441, 2022.