JAIA – Journal Of Artificial Intelligence And Applications
Journal Homepage : http://jurnal.sar.ac.id/index.php/JAIA/

Optimizing Sales Strategies to Address Excessive Stock Accumulation: A Data Mining Approach

Susandri Susandri [1], Muhammad Arief Solihin [2], Hamdani [3], Asparizal [4]
[1] STMIK Amik Riau, Jl Purwodadi Indah Km 10 Panam, Pekanbaru Riau, Indonesia
[2] STMIK Amik Riau, Jl Purwodadi Indah Km 10 Panam, Pekanbaru Riau, Indonesia
[3] STMIK Amik Riau, Jl Purwodadi Indah Km 10 Panam, Pekanbaru Riau, Indonesia
[4] Universitas Dumai, Jl. Utama Karya, Bukit Batrem, Kota Dumai Riau, Indonesia

Article Information

Received : March 14, 2024
Revised : March 14, 2024
Accepted : April 1, 2024

Keywords :

Data mining
Aassociation-rule-apriori
Tenun

Abstract

The Two Pelita Weaving Business has recorded significant sales in the weaving industry, despite facing challenges in managing product stock due to the accumulation of excess stock caused by a lack of customer interest. This study employs data mining techniques, specifically the Association Rule and Apriori algorithms, to analyze sales patterns. The analysis results using Python and Orange Data Mining showed consistency in the relationship between Siku Keluang Weaving and Pucuk Rebung Weaving products, with high occurrence rates of purchase patterns (11.74% and 10%, respectively). High confidence levels with Python at 96.36% and Orange Data Mining at 99.1% indicate that customers who purchase Siku Keluang Weaving are also likely to purchase Pucuk Rebung Weaving products.

Susandri,
Email: susandri@sar.ac.id.

1. Introduction

Indonesia is rich in culture, including the traditional woven fabrics that spread throughout the archipelago. Each region has its own unique woven fabric with distinct motifs and meanings, reflecting local beliefs and traditions. Woven fabric goes beyond mere clothing; it holds high cultural value, symbolizing heritage and beauty crafted from the skillful arrangement of its motifs [1].

One of the weaving industry hubs in Indonesia is the Bengkalis Riau. There, numerous weaving businesses operate, including two Pelita weavings, which have achieved significant sales in recent years. However, the company faces challenges in managing its product stock. Some products have excessive stock because a lack of customer interest, leading to a decreased value.

To address this issue, extensive sales data analysis is necessary. This analysis aims to determine which products need more stock and which need to be reduced, allowing the company to avoid substantial losses.

This study proposes a solution using data mining for sales data analysis using the association rule a priori algorithm. This algorithm analyzes product recommendations based on several purchasing patterns made by customers, considering variables such as customer categories, transaction quantities, and transaction periods [2][3]. The analysis results provide recommendations on which products should have more stock and which products should have less stock, thereby enabling the company to maintain profitability [4].

The use of an association rule a priori algorithm in product recommendation data analysis has been conducted in previous studies. [5] demonstrated how sales transaction data for products can be analyzed to assist in developing marketing strategies. This study uses sales transaction data for Internet quota SIM card products from several mobile telecommunication operators in the Ciamis, Tasikmalaya, and Garut regions. This study showed that sales transaction data for products can be used to identify products that are frequently purchased together.

[6] Implemented an association rule algorithm for website product promotions. They used the significant potential of available data to develop effective business strategies, including determining the most preferred products by consumers, developing appropriate marketing strategies, and managing product stock availability. The application of the association rule algorithm successfully discovered 11 association rules, with a minimum support parameter of 13% and a minimum confidence parameter of 30%. These findings provide valuable information for management, assisting them in directing business strategies, including arranging the products most preferred by customers, determining effective marketing strategies, and managing product stock availability more efficiently.

[7] used this algorithm to analyze book-borrowing patterns and provide recommendations to library authorities for arranging book layouts. The results show that this algorithm is excellent and optimal, with a minimum support value of 3% and a minimum confidence value of 50%.

Based on existing issues and previous studies, this research aims to generate product recommendation data analysis at the Tenun Dua Pelita Company based on user interaction patterns using the association rule a priori algorithm.

The results of this study are expected to benefit: Tenun Dua Pelita Company: Enhancing the efficiency of product stock management and increasing company profits. Provide solutions to address product stock issues and enhance the competitiveness of the traditional weaving industry. Provide references and contributions to research in

the fields of data mining and product recommendation data analysis.

2. Research Methods

2.1 Methods

The research methodology employed encompasses the following steps: data collection of sales and product stock, data processing to clean and prepare the data, data mining analysis using association rules and a priori algorithms to identify consumer purchasing patterns, evaluation of analysis results to identify relevant rules or patterns, and determination of sales strategy recommendations based on analysis results. The research stages are illustrated in Figure 1.

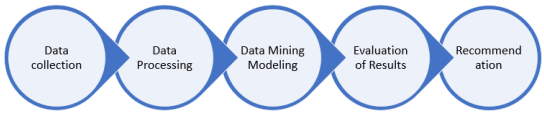


Figure 1. Research Stages

2.2 Data collection and processing

The research data were collected from Tenun Dua Pelita Company, located in Jl. Utama, Desa Sebauk, Bengkalis Regency, Riau. The transaction data consist of columns for month, gender, age, occupation, purchase quantity, product, totaling 904 records.

The application of the Association Rule algorithm to these data involves pre-processing to transform the data into a suitable format for analysis. At this stage, data cleaning processes are conducted to eliminate duplicate data and select the attributes required for data mining.

The transformation of numerical data into categorical data was performed with age data grouped into categories of 0–9, 10–15, 16–30, 31–40, 41–50, and >51 years. Month data were grouped into January–April, May–August, and September–December. Purchase quantity data were grouped into 1-30, 31-50, 51-100, and > 100 pcs.

Table 1. The result of the transformation

Month	Age	Quantity pcs	Product
January–April	16-30	1- 30	Tenun Siku Keluang
January–April	16-30	> 100	Tenun Siku Keluang
May–August	31-40	31-50	Tenun Siku Awan
Sept–Dec	10-15	51-100	Tenun Pucuk Paku

2.3 Data Mining Modeling

This stage involves data mining analysis using the Apriori Algorithm with Python language, which includes the installation of the required libraries to run the program. The installed libraries comprised the following:

A. Pandas: used for efficient data processing.
B. Numpy: employed for scientific computing purposes.
C. Apyori: used to implement the apriori algorithm in association rule mining.
D. Wordcloud: employed for visualization in the form of a word cloud.
E. Plotly: used for creating interactive visualizations.

For further comparison, an analysis was conducted using the Orange Data Mining application. Orange Data Mining is used for data processing and prediction [8]. The analysis process in Orange Data Mining is carried out through the associate widget, which can be downloaded via the Orange Data Mining application [4][9]. Data mining modeling involves the use of techniques such as classification, clustering, association rules, or predictive algorithms to analyze sales and inventory data [10]. The main objective is to build a model to identify patterns, rules, or relationships that can be used to optimize sales strategies [11].

3. Results and Discussion

In this study, the data obtained were stored in CSV format and used as input. Analysis and evaluation were conducted

using Python and Orange applications. The results were compared and used as recommendations for sales strategies.

3.1. Evaluation Results using Python

The apriori algorithm was employed to discover associations among items in the dataset. The parameters used included min_support, min_confidence, and min_lift to determine the minimum thresholds for support, confidence, and lift in the association discovery.

The min_support parameter determines the minimum support threshold that an association rule must achieve to be considered statistically significant. A value of 0.1 was used in the program code to indicate that a rule must have at least 0.1 (10%) support from the total transactions to be considered significant.

The min_confidence parameter sets the minimum confidence threshold that the association rules must achieve to be considered significant. A value of 0.2 used in the code indicates that a rule must have at least 0.2 (20%) confidence to be considered significant.

The min_lift parameter determines the minimum threshold value. A value of 1 used in the code indicates that a rule must have at least one lift to be considered significant.

The result of calling the a priori function returned an association rule object containing the generated association rules based on the provided transaction data.

The association results were displayed in the form of a dataframe with empty columns and 'Rule,' 'Support,' and 'Confidence' columns to store the association results. The program iterated over each association rule generated from the association rule object. For each association rule, item pairs involved in the rule were extracted. The names of these item pairs were converted to strings and added to the result dataframe, along with the support and confidence values rounded to percentages, as shown in Table 2.

Table 2. Association Results using the Program

Rule	Support %	Conf %
['1 Tenun siku bunga'] -> ['2 Tenun pucuk rebung']	11.18	93.52
['2 Tenun pucuk rebung'] -> ['1 Tenun siku bunga']	11.18	44.3
['1 Tenun siku keluang'] -> ['2 Tenun pucuk rebung']	11.74	96.36
['2 Tenun pucuk rebung'] -> ['1 Tenun siku keluang']	11.74	46.49

The association results are shown in Table 2, with indicated the associations between frequently purchased items in the analyzed dataset. Association rules comprised two groups of items, namely, items frequently purchased together in a single transaction, while others showed relationships between various groups of items. The support results indicated how often purchase combinations occurred, whereas confidence indicated the likelihood of purchasing one item followed by another [3].

The association rules obtained indicated that the purchase of 'Tenun siku bunga' was likely followed by the purchase of 'Tenun pucuk rebung.' This rule had support of 11.18%, meaning that this purchase combination occurred in approximately 11.18% of total transactions. The confidence of this rule reached 93.52%, indicating that approximately 93.52% of purchases of 'Tenun siku bunga' were followed by purchases of 'Tenun pucuk rebung. ' '.

Data analysis results were visualized to understand The patterns and relationships between the observed entities. WordCloud visualization was used in this study. A word cloud is a visual representation of word frequencies or items that appear in a text or dataset [12]. The word cloud visualization was created from the transaction data, and the results are displayed in Figure 2.

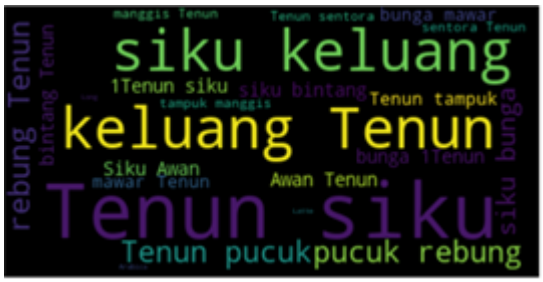


Figure 2. Transaction Data Visualization Results

Figure 2 illustrates the word cloud visualization results, showing words that appear frequently in the dataset of item purchases visually. The size of the words in the word cloud indicates their frequency of occurrence, with larger words appearing more frequently in the dataset.

3.2. Results of the Evaluation Using Orange

After conducting data mining analysis using the Association Rule with the Apriori algorithm in Python, the analysis process was further carried out using the Orange application. Orange data mining applications used for data processing, prediction, and digital image processing [4]. The process in Orange is performed using the Associate widget.

The Associate widget is necessary for the data mining process for association rules; by default, Orange data mining does not add this widget, therefor it needs to be added for association rule purposes in the widget-addons menu [9]. The data used in the Apriori analysis process are displayed using a data table widget. The evaluation framework using Orange is shown in Figure 3.

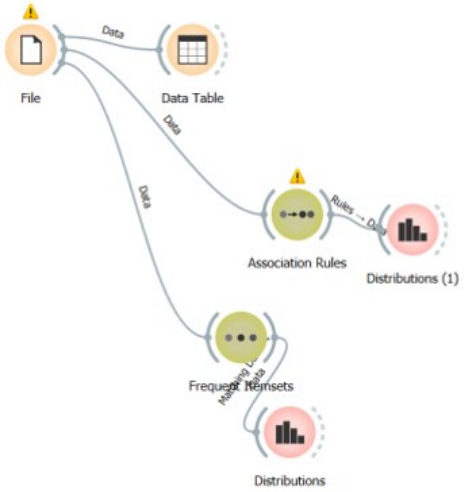
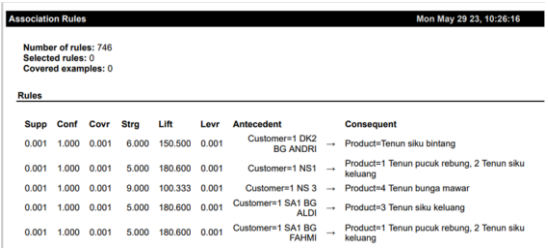


Figure 3. Framework Association Rule Orange

Figure 3 illustrates the process of the apriori association rule, displaying tools such as the data table for the association rule process, itemset frequency, and graph visualization using the distributions widget. The results of the evaluation using Orange are shown in Figure 4.



Association Rules — Mon May 29 23, 10:26:16

Number of rules: 746
Selected rules: 0
Covered examples: 0

Rules

Supp	Conf	Covr	Strg	Lift	Levr	Antecedent		Consequent
0.001	1.000	0.001	6.000	150.500	0.001	Customer=1 DK2 BG ANDRI	→	Product=Tenun siku bintang
0.001	1.000	0.001	5.000	180.600	0.001	Customer=1 NS1	→	Product=1 Tenun pucuk rebung, 2 Tenun siku keluang
0.001	1.000	0.001	9.000	100.333	0.001	Customer=1 NS 3	→	Product=4 Tenun bunga mawar
0.001	1.000	0.001	5.000	180.600	0.001	Customer=1 SA1 BG ALDI	→	Product=3 Tenun siku keluang
0.001	1.000	0.001	5.000	180.600	0.001	Customer=1 SA1 BG FAHMI	→	Product=1 Tenun pucuk rebung, 2 Tenun siku keluang

Figure 4. Orange Association Results

Figure 4 displays the results of product purchase rules frequently performed by customers obtained from Orange data mining. Support and Confidence: Each association rule has high support and confidence, namely 0.001 and 1.000, respectively. This indicates that the combination of these items often appears in product purchase transactions.

The generated lift values are also noteworthy; it can be seen that some rules have significantly high lift values, reaching above 100. This indicates a strong relationship between items in product

purchase transactions. The purchase patterns obtained are as follows:

A. The first pattern shows that customers who purchase '1 Tenun siku bunga' tend to also purchase '2 Tenun pucuk rebung' with a support level of 11.18% and a confidence level of 93.52%.
B. The second pattern shows that customers who purchase '2 Tenun pucuk rebung' tend to also purchase '1 Tenun siku bunga' with a support level of 11.18% and a confidence level of 44.3%.
C. The third pattern shows that customers who purchase '1 Tenun siku keluang' tend to also purchase '2 Tenun pucuk rebung' with a support level of 11.74% and a confidence level of 96.36%.
D. The fourth pattern shows that customers who purchase '2 Tenun pucuk rebung' tend to also purchase '1 Tenun siku keluang' with a support level of 11.74% and a confidence level of 46.49%.

The results of the testing from both modes conducted using Orange data mining and Python. The highest data from both models are extracted, as shown in Table 3.

Table 3. Comparison Table of the Model Results

Model	Rule	Supp %	Conf %
Python	['1 Tenun siku keluang'] -> ['2 Tenun pucuk rebung']	11.74	96.36
orange	Product=1 Tenun pucuk rebung, 2 Tenun siku keluang	10	99.1

Based on Table 3, the best pattern found by Python utilizes the association rule ['1 Tenun siku keluang'] -> ['2 Tenun pucuk rebung'] with a support level of 11.74% and a confidence level of 96.36%. The orange pattern utilizes the association rule Product=1 Tenun pucuk rebung, 2 Tenun siku keluang with a support level of 10% and a confidence level of 99.1%.

These results indicate consistency between the two models in identifying the relationship between "1 Tenun siku keluang" and "2 Tenun pucuk rebung". The high support levels, 11.74% and 10%, suggest that the purchase pattern of these two products frequently occurs in the analyzed dataset. Furthermore, the high confidence levels, 96.36% and 99.1%, indicate that when customers purchase "1 Tenun siku keluang", they are highly likely to also purchase "2 Tenun pucuk rebung".

4. Conclusion

The analysis results from both programs and tools, namely Python and Orange Data Mining, revealed consistency in identifying the correlation between Tenun Siku Keluang and Tenun Pucuk Rebung products. The occurrence rate of purchasing patterns for both products is quite high, namely 11.74% and 10%. This means that these two products are frequently bought together by customers. The high confidence level, with Python showing 96.36% and Orange Data Mining showing 99.1%, indicates that when customers purchase Tenun Siku Keluang, there is a high likelihood of also purchasing Tenun Pucuk Rebung.

5. Reference

[1] B. E. Pratiwi, “Motif Pucuk Rebung pada Kain Tenun Songket Melayu Riau,” 2021.
[2] E. T. N. & Andri, “Penerapan Data Mining Untuk Analisis Daftar Pembelian Konsumen Dengan Menggunakan Algoritma Apriori Pada Transaksi Penjualan Toko Bangunan MDN,” *J. Nas. Ilmu Komput.*, vol. 2, no. 2, pp. 89–101, 2021.
[3] J. Choi, Y. Won, and J. Kim, “Association Rule Mining with Apriori Algorithm for Pediatric Foot Disorders,” *WSEAS Trans. Comput.*, vol. 21, pp. 66–70, 2022, doi: 10.37394/23205.2022.21.9.
[4] M. H. Santoso, “Application of Association Rule Method Using Apriori Algorithm to Find Sales Patterns Case Study of Indomaret Tanjung Anom,” *Brill. Res. Artif. Intell.*, vol. 1, no. 2, pp. 54–66, 2021.
[5] U. Baetulloh *et al.*, “Penerapan Metode

Association Rule Mining pada Data Transaksi Penjualan Produk Kartu Perdana Kuota Internet," *J. SIMETRIS*, vol. 10, no. 1, pp. 173–188, 2019.

[6] F. Afif, R. Swedia, and M. Cahyanti, "Implementasi Algoritma Association Rule untuk Promosi Produk Berbasis Website pada Bengkel Delta Jaya Motor," *J. Ilm. Teknol. dan Rekayasa Vol.*, vol. 24, no. 2, pp. 152–160, 2019.

[7] D. Rusdianto *et al.*, "Implementasi Data Mining Menggunakan Algoritma Apriori untuk Mengetahui Pola Peminjaman Buku di Perpustakaan Universitas," *J. Sist. Informasi, J-SIKA*, vol. 02, no. 2, pp. 1–10, 2020.

[8] N. Ghafoor and M. Ahmad, "Prioritizing Effectiveness of Algorithms of Association Rule Mining," *J. Comput. Learn. Strateg. Pract.*, vol. 1, no. I, pp. 18–30, 2021.

[9] R. Abd and A. Shekan, "Data Mining and Knowledge Discovery for Big Data in Cloud Environment," *Webology*, vol. 18, pp. 1118–1131, 2021, doi: 10.14704/WEB/V18SI04/WEB18186.

[10] H. B. Sabila, "Implementation of Apriori Algorithm for Data Mining on Sales Transaction Data," *Int. J. Electr. Energy Power Syst. Eng.*, vol. 6, no. 3, pp. 189–193, 2023.

[11] J. Gao, Y. Zhao, L. Li, and S. Deebhijarn, "Feasibility Analysis Model of Transformation from Real Rule Algorithm," *Hindawi Mob. Inf. Syst.*, vol. 2021, pp. 1–8, 2021.

[12] S. Defit, F. Riandari, and B. Sinaga, "Timeline of Exploration on the Best Response Time in WhatsApp Group " Our Jokes at STMIK Amik "," *Matrik J. Manajemen, Tek. Inform. dan Rekayasa*, vol. 20, no. 2, p. 317~324, 2021, doi: 10.30812/matrik.v20i2.1149.