Development of a Stock Recommendation System Through Purchase Pattern Analysis Using the FP-Growth Algorithm

Ari Saepudin[1], Yan Puspitarani[2]

[1]ari.saepudin@widyatama.ac.id

[2]yan.puspitarani@widytama.ac.id

FakultasTeknik Informatika,Universitas Widyatama, Indonesia

Jl. Cikutra No.204A, Sukapada, Kec. Cibeunying Kidul, Kota Bandung

Abstract — Efficient stock management is a critical challenge in logistics and inventory management systems. Errors in stock demand prediction can lead to overstock or stockouts, impacting operational costs. This study aims to develop a stock recommendation system based on purchasing pattern analysis using the FP-Growth algorithm. The algorithm was chosen for its ability to efficiently extract high-frequency patterns from large transactional datasets.

The developed system utilizes historical transaction data to identify consumer purchasing patterns, such as frequently purchased item combinations. This information is then used to provide recommendations for optimal stock quantities. Testing was conducted on real transaction datasets, focusing on specific time periods to ensure the accuracy and relevance of the recommendations. This system is expected to serve as a practical solution for companies to support data-driven decision-making.

Keywords: Data Mining, FP-Growth Algorithm, Data Transaction, Purchasing Patterns

I. INTRODUCTION

Inventory is a crucial component of a company, as effective inventory management can prevent stockouts and overstocking. According to [1], inventory management is essential to maintain stock balance. Another source explains that overstock can lead to losses due to increased storage costs and tied-up capital, while stockouts risk disrupting operational continuity [2].

Inventory management plays a critical role in ensuring smooth operations and customer satisfaction. Imbalances in inventory, whether in the form of overstock or stockouts, often cause issues such as wasted storage costs, reduced profits, and missed sales opportunities.

With the advancement of information technology, sales transaction data recorded in logistics management systems can be utilized as a valuable source of information. This data can be analysed to uncover hidden patterns that provide insights into consumer purchasing behaviour. One widely used method in transaction pattern analysis is Market Basket Analysis. This method employs the FP-Growth algorithm to identify items that are frequently purchased together [3].

Based on this premise, this study aims to develop an Inventory Recommendation System capable of analysing consumer purchasing patterns using the FP-Growth algorithm.

II. RESEARCH METHODOLOGY

This chapter provides a more in-depth analysis of issues related to inventory management and how the FP-Growth algorithm can be applied to provide stock recommendations based on consumer purchasing patterns. The problem analysis stage includes identifying the main issues, determining the methods to be used, and designing the steps for problem resolution.

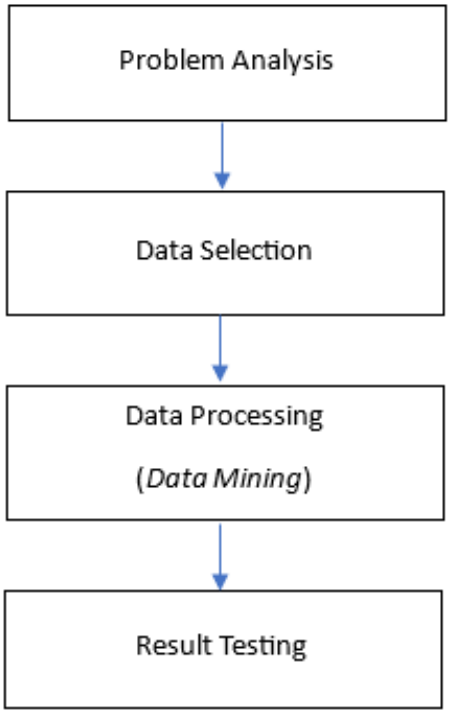


Figure 2.1 Research Design

A. Problem Analysis
The main issue faced is how to provide optimal inventory levels based on consumer purchasing patterns.

B. Data Selection
Data selection is a crucial step in the research process, aimed at selecting relevant data from the sales transaction dataset to be further analysed using the FP-Growth algorithm. Data selection ensures that only relevant, valid, and high-quality data are used in the analysis process.

C. Data Processing
After data selection, the chosen data is processed to generate useful information. In this research, data processing includes transforming transaction data into a format suitable for the FP-Growth algorithm and analysing purchasing patterns.

D. Result Testing
Once the data processing is complete and association rules are formed, the results are tested to ensure their accuracy. The testing is conducted using historical transaction data that has not yet been analysed, to assess how well the system can predict future purchasing patterns and stock requirements.

III. THEORETICAL FOUNDATION

3.1 Literature Review

3.1.1 Definition of Transaction Data

Transaction data refers to data that records events that have occurred within a business or company, including sales, shipments, debts, and other events. Transaction data can typically be categorized into three groups based on the verbs used: finance for orders and payments, labour for schedules and work records, and logistics for shipments. The general definition of a transaction includes any activity that impacts the assets or finances of an organization or individual. Some examples of transactional activities include sales, purchases, salary payments, and so on. In transaction data, transaction management is responsible for accurately recording financial changes using specific methods. Sales transaction data holds significant value in business decision-making [4].

3.1.2 KDD (*Knowledge Discovery Database*)

KDD (Knowledge Discovery in Database) focuses on the data exploration phase to uncover patterns or new knowledge from large and complex datasets. This aligns with the objective of the research, which is to discover consumer purchasing patterns through transaction data analysis using the FP-Growth algorithm. KDD emphasizes the data analysis process until it results in patterns or useful information (e.g., frequent itemset in FP-Growth). The discovery of patterns, such as association rules for inventory recommendations, is highly relevant to the principles of KDD.

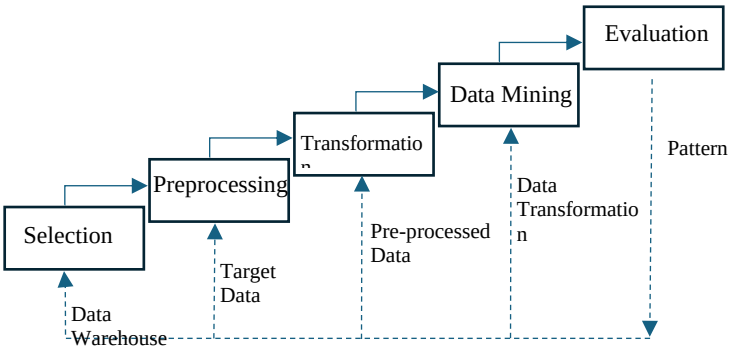


Figure 3.1 KDD Process Stages

3.1.3 Microsoft Visual Studio

Microsoft Visual Studio is a comprehensive software suite that can be used for application development, whether for business applications, personal applications, or application components, in the form of console applications, Windows applications, or web applications.

Microsoft Visual Studio can be used to develop applications in native code (in the form of machine language running on Windows) or managed code (in the form of Microsoft Intermediate Language running on the .NET Framework).

3.1.4 Unifield Modelling Language (UML)

Unified Modelling Language (UML), according to Fowler, "is a family of graphical notations supported by a single meta-model, which helps in the description and design of software systems, particularly systems built using object-oriented (OO) programming" [5].
Unified Modelling Language (UML) is a standard specification language used to document, specify, visualize, and construct software systems, as well as business modelling and other systems. UML is not based on any specific programming language. The UML specification standard became the de facto standard by OMG (Object Management Group) in 1995. Object-oriented UML includes several standard notations.

a) Activity Diagram
b) Use Case Diagram
c) Class Diagram

IV. RESULTS AND DISCUSSION

4.1 Inventory Recommendation System

Based on the analysis of consumer purchasing patterns, an inventory recommendation system was developed to assist companies in managing their inventory. The system aims to provide recommendations on which items should be prioritized for restocking based on the frequent itemset and association rules that have been identified.

4.1.1 Use Case Diagram

Use Case Diagram for the Inventory Recommendation System

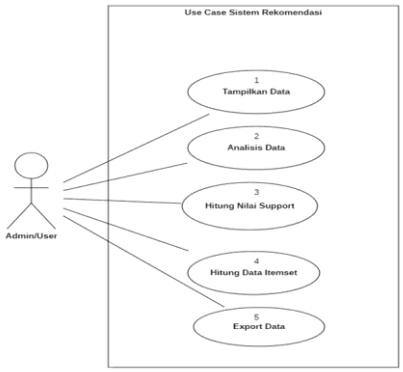
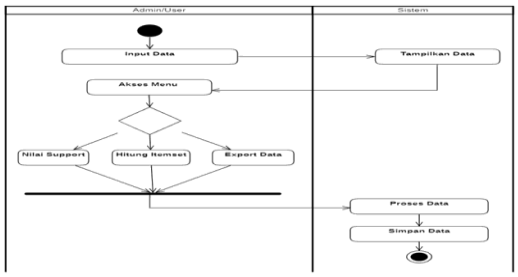


Figure 4.1 Inventory Recommendation System

Description: This use case involves a single actor, the admin/user. The system is entirely managed by the admin/user, who can perform tasks such as managing or analysing sales data, inputting raw sales data, and processing the data to generate recommendations on which items are frequently purchased together. This allows the system to provide stock requirement recommendations.

4.1.2 Activity Diagram

An activity diagram illustrates the workflow or activities of a system or business process. For more clarity, refer to the flow of the activity diagram below.



Gambar 4.2 Activity Diagram

4.1.3 *Class Diagram*

A class diagram is one of the UML diagrams that describes the structure, explanation, and relationships between each class, method, and attribute. The representation of a class diagram is relatively easy to understand and use, making it commonly employed in object-oriented projects. The class diagram of the system can be seen in Figure 4.3.

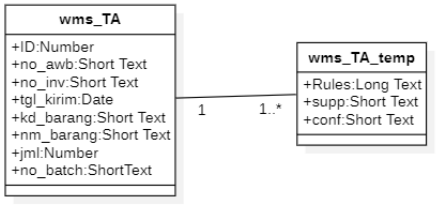


Figure 4.3 Class Diagram

4.2 Form Interface

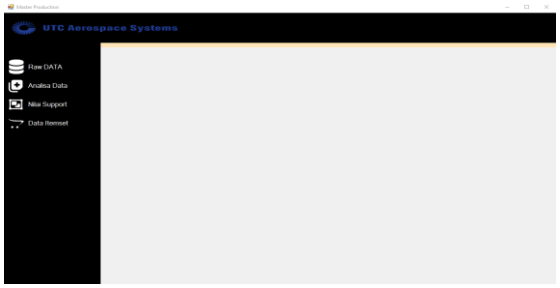


Figure 4.4 Main Page

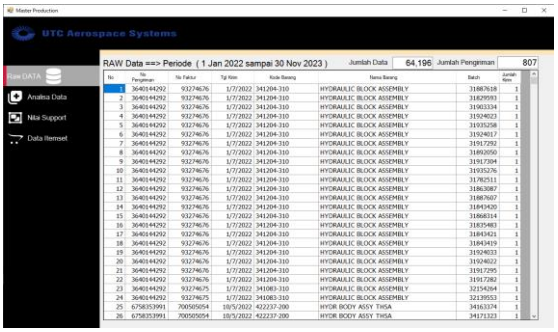


Figure 4.5 Raw Data Page

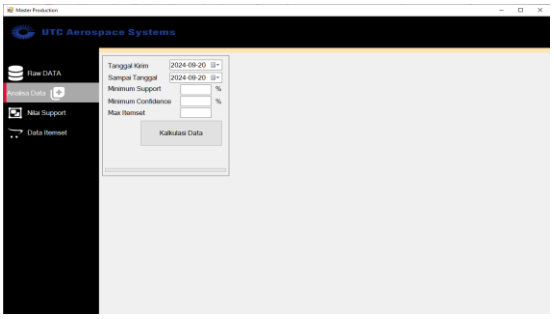


Figure 4.6 Analysis Page

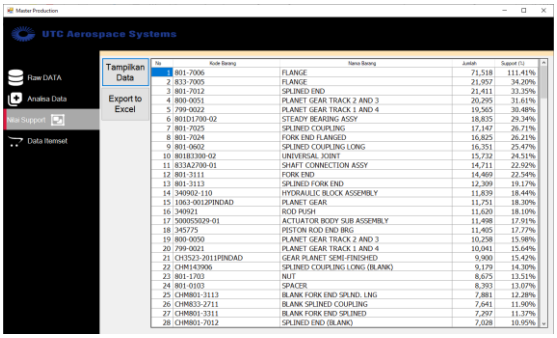


Figure 4.7 Support Value Page

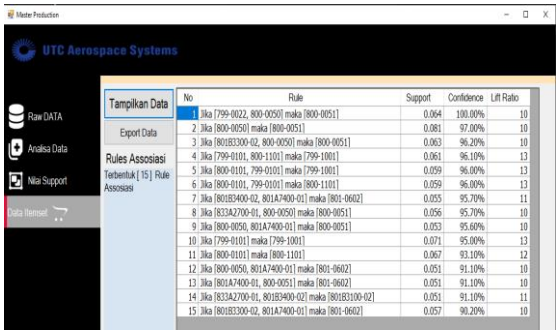


Figure 4.9 Itemset Page

4.3 Results of Purchasing Pattern Analysis

After the analysis process, several relevant purchasing patterns were identified based on Table 4.1, listing the top 10 rules generated by the FP-Growth algorithm based on the highest confidence values, as follows:

Table 4.1 Association Rules Results

Rule	Support	Confidence	Lift Ratio
Jika [799-0022, 800-0050] maka [800-0051]	6.4	100.00%	10
Jika [800-0050] maka [800-0051]	8.1	97.00%	10
Jika [801B3300-02, 800-0050] maka [800-0051]	6.3	96.20%	10
Jika [799-0101, 800-1101] maka [799-1001]	6.1	96.10%	13
Jika [800-0101, 799-0101] maka [799-1001]	5.9	96.00%	13
Jika [800-0101, 799-0101] maka [800-1101]	5.9	96.00%	13
Jika [801B3400-02, 801A7400-01] maka [801-0602]	5.5	95.70%	11
Jika [833A2700-01, 800-0050] maka [800-0051]	5.6	95.70%	10
Jika [800-0050, 801A7400-01] maka [800-0051]	5.3	95.60%	10
Jika [799-0101] maka [799-1001]	7.1	95.00%	13

4.3.1 Analysis of Association Rules Data

1. Support Analysis

   Support represents how frequently an itemset combination appears in the dataset. A higher support indicates that the combination occurs more frequently in transactions.

   a. Highest Support (0.081):
      If [800-0050], then [800-0051] is recommended.
      Interpretation: Items 800-0050 and 800-0051 are frequently purchased together, making this combination important to consider in inventory management.

   b. Lowest Support (0.053):
      If [800-0050, 801A7400-01], then [800-0051] is recommended.
      Interpretation: This combination occurs infrequently but still has a strong relationship (high confidence).

4.3.2 Confidence Analysis

Confidence indicates the likelihood that item B will be purchased when item A has already been purchased. Confidence approaching 100% indicates a very strong relationship.

a. Rule with the Highest Confidence:
   If [799-0022, 800-0050], then [800-0051] → 100.00%
   Interpretation: Every time customers buy 799-0022 and 800-0050 together, they will definitely buy 800-0051. This relationship is very strong.
   Recommendation: Ensure that 800-0051 is always in stock when 799-0022 and 800-0050 are frequently purchased together.

b. Rules with High Confidence (>95%):
   Other rules have confidence between 95.00% and 97.00%, which are still very strong:

   If [800-0050] → [800-0051] (97.00%)
   If [799-0101, 800-1101] → [799-1001] (96.10%)
   If [801B3300-02, 800-0050] → [800-0051] (96.20%)
   Conclusion: Items 800-0051 and 799-1001 have strong relationships with other combinations and should be a focus in the recommendation strategy.

V. CONCLUSION

This research successfully implemented the FP-Growth algorithm in the data mining process to analyze consumer purchasing patterns based on sales transaction data from PT UTC Aerospace Systems Bandung. The analysis results revealed significant purchasing patterns, providing valuable insights into consumer behaviour. These insights assist the company in making more accurate, data-driven decisions related to inventory management, thereby improving the efficiency and effectiveness of stock management.

Data validation showed that the purchasing patterns discovered through the FP-Growth algorithm have high accuracy, with significant support and confidence values. Based on the results:

a. The support of the generated rules ranges from 5.3% to 8.1%, which falls within the moderate support category (5%-20%). This indicates that the identified transaction patterns involve products that are frequently purchased

together, reflecting relevant product combinations for inventory

b. The confidence for all the rules is above 95%, which falls into the high confidence category (>80%), indicating a very strong relationship between items in the pattern. With high confidence, these patterns can serve as a reliable basis for stock recommendations.

DAFTAR PUSTAKA

[1] Bahagia, Sistem Inventory, Penerbit ITB, 2006.

[2] S. L.D., Pengendalian Persediaan, Forum Pemuda Aswaja, 2021.

[3] T. W. U. d. R. W. N. D. Andriai, "Implementasi Algoritma FP-Growth Menganalisis Pola Belanja Konsumen Pada Data Transaksi Penjualan," *Jurnal Ilmiah,* 2019.

[4] B. T. M. &. C. B. Triloka, "Analisis Pola Penjualan dan Prediksi Permintaan Produk Parfum Toko Kayyash Menggunkan Model FP-Growth dan Arima," *Jurnal Sistem Informasi,* vol. XIV, no. ISSN 2962-5300, p. 1, 2024.

[5] F. Martin, UML Distilled Edisi 3, Yogyakarta: Andi, 2005.