



Application of Data Mining to Predict Medical Device Sales Using Apriori Algorithm

Dinda Gustari¹, Sumarna²

¹Information Systems Study Program, Universitas Nusa Mandiri, Indonesia

²Informatics Study Program, Universitas Nusa Mandiri, Indonesia

ARTICLE INFO

Article history:

Received Nov 03, 2022
Revised Nov 16, 2022
Accepted Nov 30, 2022

Keywords:

Application,
Data Mining,
Predict Medica

ABSTRACT

It is difficult to obtain information regarding the sale of the company and the level of sales of products. In data processing, especially when selling merchandise, manual sales reports are still used. The research method used is a quantitative research method with Primary Data obtained by Interview Techniques using a priori algorithm data analysis. Application of a priori algorithm in 305 drug sales transaction data in PT. Tiga Cahaya Nagarimas using RapidMiner software with a minimum specified support of 25% and confidence of 50% resulted in 24 rules of association, thus PT. The Three Lights of Nagarimas can find out the relationship between the frequency pattern of selling the most medical devices purchased by consumers simultaneously and the medical devices that are most sold. With the knowledge of medical devices that are often purchased, PT. Three Lights Nagarimas can develop marketing strategies, regulate the stock of medical devices so that there is no accumulation of medical devices or empty medical device stocks and arrange the layout of storage of medical devices that are often purchased simultaneously to make it easier for employees to find medical devices that are needed to optimize services to buyers.

This is an open access article under the [CC BY-NC](https://creativecommons.org/licenses/by-nc/4.0/) license.



Corresponding Author:

Sumarna,
Informatics Study Program,
Universitas Nusa Mandiri,
Jl. Jatiwaringin No.2, Cipinang Melayu, East Jakarta City, 13620, Indonesia.
Email: sumarna@nusamandiri.ac.id

1. INTRODUCTION

As information technology develops rapidly, it becomes important for businesses to keep up. This has forced many people to compete with each other in order to stay competitive. In today's competitive business world, success depends on constant adaptation and development. As businesses grow, they need to find ways to compete with their competitors, who use newer and more innovative technologies. To do this, they may need to start using new technology to support their day-to-day operations. Having these operational systems in place is important, but they don't produce the strategic information we expect. Information technology is a combination of computer technology, telecommunications and other technologies (Aminuddin, 2013).

PT. Tiga Cahaya Nagarimas is a company engaged in the sale of medical devices. There are several problems that can arise when trying to sell a product, often including

selling issues. The difficulty of obtaining information about company sales and product sales levels. In data processing, especially when selling merchandise, sales reports are still used manually. In (Mohammad Ahmadar et al., 2021) An increase in total assets or a decrease in total liabilities of a business entity arising from the delivery of goods, services or other activities within a period is part of sales. And Marketing is a social process in which individuals and groups get what they need and want by creating and offering freely exchanging products of value with other parties according to Kothler and Keller in (Jasmalinda, 2021)

So using this the author uses data mining techniques with the a priori algorithm method. By knowing the value of the final association that meets support and confidence, you can find the best-selling products, as a result, companies can develop good marketing tactics to market products using other brands by researching what are the advantages of the best-selling products compared to other products and can increase inventory. best selling product. Because of that, the writer is interested in raising the title of the thesis "Application of Data Mining to Predict Sales of Medical Devices Using the Apriori Algorithm".

Data Mining has actually been known since 1990, when the use of data was one of the crucial things in various academic, business, medical and other fields. Data Mining is a process of receiving useful information according to a database warehouse in the form of knowledge (Ulfa & Amin, 2020). Data Mining is the process of looking for patterns and interesting information on selected data using specific techniques or methods. Techniques, methods, or solving procedures in data mining vary greatly in choosing the perfect solving procedure method, depending on the goals and the KDD process as a whole (Suliman, 2021).

Data mining is a process for manipulating data by extracting previously unknown information according to large datasets (Oktanisa & Supianto, 2018). From the definitions that have been cited, it can be concluded that data mining is the process of looking for patterns using proprietary methods to obtain information from a large dataset in the Data Mining Stages. In quote (Supoyo & Prasetyaningrum, 2022) Data Mining is an integrated part of knowledge discovery in the database which is in the following order:

Data Selection, Relevant data will be retrieved from a database, then a relationship analysis will be carried out to obtain data characteristics. Data Cleaning, Elimination of incomplete, error, and asynchronous data according to the dataset. Data Integration, Data is retrieved from various data sources & stored in a coherent data store. Data Transformation, Data that is transformed into a synchronous format so that it can be obtained in the next process.

Data Mining, Applying algorithms and extracting data patterns. Pattern Evaluation

An interesting pattern identification for presenting knowledge according to several measurements that have been made. Knowledge Presentation Visualization techniques to represent the knowledge acquired. In the book (Jollyta et al., 2020) the data mining method is a method that is applied and needs to be adjusted to the goals of its users. There are several divisions of data mining methods as follows:

Association, It is useful to recognize the behavior according to specific events where the association link process occurs in each event and try to set up the values that occur simultaneously in each row and make the result inferred to a rule using the Support or Confidence parameter.

Classification, Classification is a single example model that is most often used as a learning example in data mining with the aim of forming an example to predict consumer behavior in the future using a structured division of records in the database into several class groups based on certain criteria. Clustering can also be used to automatically detect clusters based on adjacent records that have a certain meaning in the holistic variables. Forecasting. This model is carried out using a way of estimating future values based on record patterns using mutually continuous values obtained.

Regression, Is a statistical assumption technique that uses predictive mapping. Sequence Discovery Is an identification by association or pattern whose goal is to model the state according to the process of forming a sequence to extract, report deviations and trends. Visualization, This model is a presentation according to the data as a result the user can see the complete pattern. Based on (Rerung, 2018) Data mining is divided into several groups based on the task or work performed, namely, Description Researchers and analysts simply try to find ways to describe the patterns and similarities that still exist in the data.

Estimation, The estimated target variable is more numerical than categorical. The model is built using complete data lines (records) using and providing values according to the target variable to become predictive values.

Predictions, Almost the same as using translation and forecasting, except for predicting the value of the output that will occur in the future. Classification Contains target categorical variables. Clustering Grouping records, observing, or paying attention and building several classes of objects that are similar.

Association, Assigned to data mining to find attributes that arise at one time. The apriori algorithm is a classical solving procedure that is used by personal computers to be able to view association rules and find patterns of interaction between dataset items. For example, for example, looking for and observing the pattern of goods that consumers buy the most at the same time using a list of goods for promotion as a result, consumers are required to buy goods in greater quantities (Prasetya et al., 2021 In (Jollyta et al., 2020) Association rule is a close analysis of the study of "what uses what" or "something has continuous use of something" using association rules that allow the purchase of goods that are still in the database.

According to the regulation of the Minister of Health No. 1189 of 2010 in (Nazmi, 2018) Medical devices are instruments, apparatus, machines or implants that do not contain drugs that are used to prevent, diagnose, relieve disease and treat sick people and build structures to improve body functions.

2. RESEARCH METHOD

The research method is a systematic and useful research stage that helps research and is well directed. The following is the research method used.

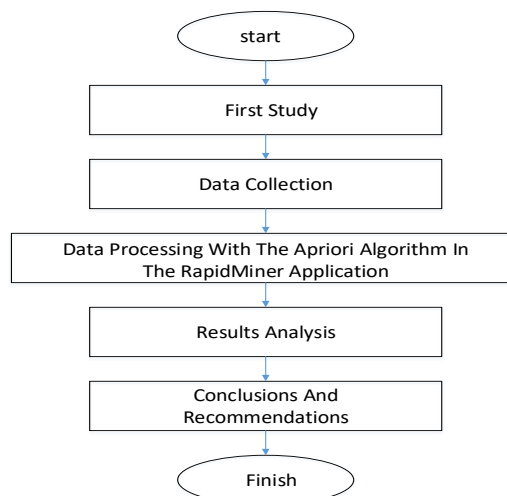


Figure 1. Research Stages

Based on the picture of the research stages above, it can be explained as follows:

- a. Early Studies
The first step based on this research is to find and investigate cases that occurred in PT. Tiga Cahaya Nagarimas. Then choose the scope, background of the problem, and investigate several methods related to the problem and how to find solutions to cases that occur within the company.
- b. Data Collection
In this stage it is carried out by coming directly to PT. Tiga Cahaya Nagarimas. Then conduct an interview with one of the Company's employees to find out the information needed. As well as collecting medical device sales data for the last 3 months at PT. Tiga Cahaya Nagarimas.
- c. Processing data with Data Mining
The data obtained is then processed using the a priori algorithm method. The first stage is creating a tabular format using Microsoft office excel, then using RapidMiner for high frequency pattern analysis, forming association rules and final association rules so that the final output is obtained.
- d. Results Analysis
At this stage, output analysis is carried out according to the calculation of the a priori algorithm procedure and output analysis based on the system built as a result of making decisions that can be used for decision making in determining Medical Device sales patterns.
- e. Conclusions and Suggestions
The final result is obtained from the final association rule process that has been carried out using the Apriori Algorithm

2.1 Data analysis method

The research was conducted at PT. Tiga Cahaya Nagarimas in the field of selling medical devices. In analyzing the data, namely by calculating the a priori algorithm as follows:

- a. High frequency pattern analysis
In this stage, look for item combinations that meet the minimum requirements of the support value in the database. The following formula can be used to obtain a support value:

$$\text{support}(A) = \frac{\text{transaction amount contains } A}{\text{transaction totals}}$$

And finding the support value of 2 items can be obtained by using the following formula :

$$\text{support}(A,B) = \frac{\sum \text{transaction contains } A \text{ and } B}{\sum \text{transaction}}$$

- b. Formation of association rules
After all high-frequency patterns are found, then look for association rules that meet the minimum requirements for confidence by calculating the confidence of the associative rules A B
The formula for the confidence value of the A B rule is obtained as follows :

$$\text{Confidence} = P(B|A) = \frac{\sum \text{transaction contains } A \text{ and } B}{\sum \text{transaction contains } A}$$

- c. Application of the Apriori Algorithm on RapidMiner
After the calculation results of the a priori algorithm are obtained using manual calculations, it will be more accurate by using the RapidMiner software by

inputting tabular tables into Ms.Excel which is then entered into the RapidMiner Software.

3. RESULTS AND DISCUSSIONS

3.1 Data Mining Analysis

Before processing 305 medical device sales transaction data using the RapidMiner application, it is first explained how to manually calculate the a priori algorithm. Here we take data as an example, namely as many as 12 medical device sales transaction data as in the table below.

Table 1. Data on Sales of Medical Devices Transactions for January-March 2022

Transaction	Medical Device Name
1	NOUSMED Suction Connecting Tube Yankeur,Masker Oksigen Dewasa,ABADA nebulizer Adult,SAFELOCK Urine Bag With Hanger,ABADA Folley Catheter No.14'
2	SAFELOCK Mask Earloop,SAFELOCK Urine Bag Non Hanger,ABADA nebulizer Adult,SAFELOCK Urine Bag With Hanger,OTO underpad,Medice Kassa 40cm x 80 yard,ABADA Folley Catheter No.14
3	SAFELOCK Urine Bag With Hanger,SAFELOCK Nassal Cannula Adult,NOUSMED Infusion Set Y,Formalin 37% 5L
4	Masker oksigen Dewasa,ABADA Folley Catheter No.14,Abada Nebulizer Mask Child,OTO underpad,SAFELOCK Urine Bag non Hanger
5	NOUSMED Infusion Set Y,1NOUSMED Suction Connecting Tube Yankeur,SAFELOCK Nassal Cannula Adult
6	KIMSED Syringe 3ml,SAFELOCK Nassal Cannula Adult,SAFELOCK Urine Bag Non Hanger,ABADA Folley Catheter No.14,Medice Kassa 40cm x 80 yard
7	SAFELOCK Nassal Cannula Adult,ABADA Folley Catheter No.14, OTO underpad,Formalin 37% 5L,SAFELOCK Urine Bag Non Hanger
8	SAFELOCK Urine Bag non Hanger,NOUSMED Infusion Set Y,SAFELOCK Mask Earloop,OTO underpad
9	ABADA Folley Catheter No.14,SAFELOCK Urine bag with hanger,Formalin 37% 5L,ABADA nebulizer Adult,
10	SAFELOCK Nassal Cannula Adult,6Medice Kassa 40cm x 80 yard,Abada Nebulizer Mask Child, SAFELOCK Urine Bag Non Hanger
11	STERINTECH Flat Sterilization Pouch 20cm,STERINTECH Flat Sterilization Pouch 40cm
12	NOUSMED Infusion Set Y,NOUSMED Suction Connecting Tube Yankeur,SAFELOCK Nassal Cannula Adult

3.1 Formation of Itemsets

The process of forming C1 or called 1 Itemset with a minimum amount of support = 25% using equation 1. From the above process, there is the formation of items with a minimum support of 25% so that a minimum standard of support is obtained, namely 8 types, from the 8 types, these are then formed into a combination of 2 items.

Table 2. Minimum Support From 1 Itemset 25%

Medical Device Name	AMOUNT	PROCESS SUPPORT	SUPPORT
ABADA nebulizer Adult	3	(3/12)	25%
OTO underpad	4	(4/12)	33%
SAFELOCK Urine Bag With Hanger	5	(5/12)	42%
SAFELOCK Nassal Cannula Adult,	4	(4/12)	33%
NOUSMED Infusion Set Y	3	(3/12)	25%
ABADA Folley Catheter No.14	6	(6/12)	50%
Formalin 37% 5L	3	(3/12)	25%
SAFELOCK Urine Bag non Hanger	6	(6/12)	50%

Formation of C2 or 2 items with an amount that contains a minimum support of 25% can be solved with equation 2. From the combination of 2 itemsets with a minimum

support of 25%, it is known that the items that meet the minimum standards are as follows :

Table 3. Minimum Support From 2 Itemset 25%

Medical Device Name	AMOUNT	PROCESS SUPPORT	SUPPORT
ABADA nebulizer Adult,SAFELOCK Urine Bag With Hanger	3	(3/12)	25%
ABADA nebulizer Adult,ABADA Folley Catheter No.14	3	(3/12)	25%
OTO underpad,ABADA Folley Catheter No.14	3	(3/12)	25%
OTO underpad,SAFELOCK Urine Bag non Hanger	4	(4/12)	33%
SAFELOCK Urine Bag With Hanger,ABADA Folley Catheter No.14	3	(3/12)	25%
ABADA Folley Catheter No.14,SAFELOCK Urine Bag non Hanger	3	(3/12)	25%

Formation of C3 or 3 items with an amount that contains a minimum support of 25% can be solved with equation 3. From the combination of 3 itemsets with a minimum support of 25%, it is known that the items that meet the minimum standards are as follows :

Table 4. Minimum Support From 3 Itemset 25%

Medical Device Name	AMOUNT	PROCESS SUPPORT	SUPPORT
ABADA nebulizer Adult,SAFELOCK Urine Bag With Hanger, ABADA Folley Catheter No.14	3	12	25%
OTO underpad,ABADA Folley Catheter No.14, SAFELOCK Urine Bag non Hanger	3	12	25%

3.3 Formation of Association Rules

After all high-frequency patterns are found, then we look for association rules that meet the minimum requirements for confidence by calculating the confidence of the associative rules $A \rightarrow B$. Minimum Confidence = 50% and the Confidence value of the rules

$A \rightarrow B$ is obtained :

$$\text{support (A,B,C)} = \frac{\sum \text{transaction contains A,B and C}}{\sum \text{transaction}}$$

Table 5. . Association Rules 3 Itemset

RULE	$\sum A \text{ DAN } B$	$\sum A$	Confidence
if you sell ABADA nebulizer Adult then sell ABADA Folley Catheter No.14 and SAFELOCK Urine Bag With Hanger	3	3	100%
if you sell OTO underpad then sell ABADA Folley Catheter No.14 and SAFELOCK Urine Bag non Hanger	3	4	75%
if selling ABADA Folley Catheter No.14 then selling SAFELOCK Urine Bag non Hanger and Oto underpad	3	6	50%
if selling ABADA Folley Catheter No.14 then selling SAFELOCK Urine Bag With Hanger and ABADA nebulizer Adult	3	6	50%
if selling ABADA Folley Catheter No.14 and SAFELOCK Urine Bag non Hanger then selling OTO Underpad	3	6	50%
If selling ABADA Folley Catheter No.14 and OTO underpad then selling SAFELOCK Urine Bag non Hanger	3	6	50%
if selling ABADA Folley Catheter No.14 and SAFELOCK Urine Bag With Hanger then selling ABADA nebulizer Adult	3	6	50%
if you sell ABADA Folley Catheter No.14 and ABADA nebulizer Adult then sell SAFELOCK Urine Bag With Hanger	3	6	50%
if you sell SAFELOCK Urine Bag With Hanger, then sell ABADA Folley Catheter No.14 and ABADA nebulizer Adult	3	3	100%
if you sell SAFELOCK Urine Bag With Hanger and ABADA Adult nebulizer then sell ABADA Folley Catheter No.14	3	3	100%

RULE	$\sum A$ B	\sum A	Confidence
if you sell a non-hanger SAFELock Urine Bag, then sell ABADA Folley Catheter No.14 and OTO underpad	3	6	50%
if selling non-hanger SAFELock Urine Bag and OTO underpad then selling ABADA Folley Catheter No.14	3	6	50%

3.4 Implementation of the Apriori Algorithm in the Rapidminer application

To implement the a priori algorithm in rapidminer, the method is as follows: The first step is to create a Tabular Format on the Ms.Excel worksheet from 305 transaction data. Then loaded into a tabular format in Ms.Excel, to facilitate calculations in the implementation of the Apriori Algorithm. the data above is then converted into the form of binary numbers 1 and 0. The number 1 indicates the product is sold and the number 0 indicates the product is not sold and save the database that has been created in csv or xls format. Then select the operator on the left of the display that will be used to move the operator.

3.5 Final Association Rules

The final association rule is obtained from the percentage value that meets the two predetermined minimum support and confidence requirements.

Table 6. The final association rule

No	Aturan	Support	Confidence
1	if selling ABADA Folley Catheter No.14 then selling SAFELock Urine Bag With Hanger	25%	50%
2	if selling ABADA Folley Catheter No.14 then selling Oto Underpad	25%	50%
3	if selling ABADA Folley Catheter No.14 then selling ABADA Nebulizer Adult	25%	50%
4	if selling ABADA Folley Catheter No.14 then selling SAFELock Urine Bag non Hanger and Oto underpad	25%	50%
5	if you sell a non-hanger SAFELock Urine Bag, then sell ABADA Folley Catheter No.14 and OTO underpad	25%	50%
6	if selling ABADA Folley Catheter No.14 then selling AFELOCK Urine Bag With Hanger and ABADA nebulizer Adult	25%	50%
7	if you sell SAFELock Urine Bag With Hanger then sell BADA Folley Catheter No.14	25%	60%
8	if you sell SAFELock Urine Bag With Hanger then sell ABADA nebulizer Adult	25%	60%
9	if you sell SAFELock Urine Bag With Hanger, then sell ABADA Folley Catheter No.14 and ABADA nebulizer Adult	25%	60%
10	if selling ABADA Folley Catheter No.14 then selling SAFELock Urine Bag non Hanger	33%	67%
11	if you sell non-hanger SAFELock Urine Bag, then you sell Abada Folley Catheter No.14	33%	67%
12	if you sell a non-hanger SAFELock Urine Bag, then you sell the Oto underpad	33%	67%
13	if you sell OTO Underpad then sell ABADA Folley Catheter No.14	25%	75%
14	if selling ABADA Folley Catheter No.14 and SAFELock Urine Bag non Hanger then selling OTO Underpad	25%	75%
15	if you sell OTO underpad then sell ABADA Folley Catheter No.14 and SAFELock Urine Bag non Hanger	25%	75%
16	if selling non-hanger SAFELock Urine Bag and OTO underpad then selling ABADA Folley Catheter No.14	25%	75%
17	if you sell ABADA Adult nebulizer then sell ABADA Folley Catheter No.14	25%	100%
18	if you sell OTO underpad then sell SAFELock Urine Bag non Hanger	33%	100%
19	if selling ABADA nebulizer Adult, then selling SAFELock Urine Bag With Hanger	25%	100%
20	If selling ABADA Folley Catheter No.14 and OTO underpad then selling SAFELock Urine Bag non Hanger	25%	100%
21	if selling ABADA Folley Catheter No.14 and SAFELock Urine Bag With Hanger then selling ABADA nebulizer Adult	25%	100%
22	if you sell ABADA nebulizer Adult then sell ABADA Folley Catheter No.14 and SAFELock Urine Bag With Hanger	25%	100%
23	if you sell ABADA Folley Catheter No.14 and ABADA nebulizer Adult then sell SAFELock Urine Bag With Hanger	25%	100%

No	Aturan	Support	Confidence
24	if you sell SAFELOCK Urine Bag With Hanger and ABADA Adult nebulizer then sell ABADA Folley Catheter No.14	25%	100%

4. CONCLUSION

Based on the discussion in all previous chapters, it can be concluded that: Application of the a priori algorithm to 305 drug sales transaction data at PT. Tiga Cahaya Nagarimas using RapidMiner software with a specified minimum support of 25% and 50% confidence resulting in 24 association rules, thus PT. Tiga Cahaya Nagarimas can find out the relationship between the frequency pattern of selling medical devices the most purchased by consumers together as well as the most sold medical devices. By knowing the medical devices that are often purchased, PT. Tiga Cahaya Nagarimas can develop a marketing strategy, manage medical device stocks so that there is no accumulation of medical devices or empty medical device stocks and arrange the layout for storing medical devices that are often purchased together to facilitate employees in finding the necessary medical devices to optimize service to buyers..

REFERENCES

- Aminuddin. (2013). *Ensiklopedia Mini Teknologi Informasi*. CV.Angkasa.
- Jasmalinda. (2021). PENGARUH CITRA MEREK DAN KUALITAS PRODUK TERHADAP KEPUTUSAN PEMBELIAN KONSUMEN MOTOR YAMAHA DI KABUPATEN PADANG PARIAMAN. 1(10), 2199–2206.
- Jollyta, D., Ramdhan, W., & Zarlis, M. (2020). *Konsep Data Mining dan Penerapan*.
- Mohammad Ahmadar, Perwito, & Taufik, C. (2021). PERANCANGAN SISTEM INFORMASI PENJUALAN BERBASIS WEB PADA RAHAYU PHOTO COPY DENGAN DATABASE MySQL. 10(4), 284–289. <https://doi.org/10.24198/dharmakarya.v10i4.35873>
- Nazmi. (2018). IMPLEMENTASI KEBIJAKAN PENGEMBANGAN INDUSTRI ALAT KESEHATAN DALAM NEGERI. 07(01), 42–48.
- Oktanisa, I., & Supianto, A. A. (2018). PERBANDINGAN TEKNIK KLASIFIKASI DALAM DATA MINING UNTUK BANK DIRECT MARKETING. 5(5), 567–576. <https://doi.org/10.25126/jtiik20185958>
- Prasetya, T., Yanti, J. E., Purnamasari, A. I., Dikananda, A. R., & Anwar, S. (2021). Analisis Data Transaksi Terhadap Pola Pembelian Konsumen Menggunakan Metode Algoritma Apriori. 6(1), 43–52.
- Rerung, R. R. (2018). Penerapan Data Mining dengan Memanfaatkan Metode Association Rule untuk Promosi Produk. 3(1), 89–98. <https://doi.org/10.31544/jtera.v3.i1.2018.89-98>
- Suliman. (2021). IMPLEMENTASI DATA MINING TERHADAP PRESTASI BELAJAR MAHASISWA BERDASARKAN PERGAULAN DAN SOSIAL EKONOMI DENGAN ALGORITMA K-MEANS CLUSTERING. 6(1), 1–11.
- Supoyo, A., & Prasetyaningrum, P. T. (2022). Analisis Data Mining Untuk Memprediksi Lama Perawatan Pasien Covid-19 Di DIY Agung. 10(1), 21–29.
- Ulfa, N. F., & Amin, R. (2020). IMPLEMENTASI DATA MINING UNTUK MENGETAHUI POLA PEMBELIAN OBAT MENGGUNAKAN ALGORITMA APRIORI. 17(2), 396–402.