IMPLEMENTATION OF THE ASSOCIATION METHOD IN THE ANALYSIS OF SALES DATA FROM MANUFACTURING COMPANIES

Fachri Amsury[1], Nanang Ruhyana[2], Andry Agung Riyadi[3], Ihsan Aulia Rahman[4]

Information System / Faculty of Technology and Information
University of Nusa Mandiri
fachri.fcy@nusamandiri.ac.id

Data Science / Faculty of Technology and Information
University of Nusa Mandiri
fachri.fcy@nusamandiri.ac.id[1], nanang.ngy@nusamandiri.ac.id[2], andriagu1603@nusamandiri.ac.id[3], ihsanaulia24@gmail.com[4]
(*) Corresponding Author

Abstract

The company produces sales data every day. Over time, the data increases, and the amount becomes very large, and the data is only stored without understanding the benefits that exist from these data due to limitations in proper knowledge in analyzing the data, especially transaction data. Sale. In order to overcome these problems, a study focused on reprocessing sales transaction data in 2018 with a data mining technique approach using the Knowledge Discovery in Database (KDD) concept using the association method and apriori algorithm and a supporting application, namely RapidMiner. This study aims to help companies find customer buying habits or patterns based on 2018 sales transaction data. The results of this study produce 316 association rules where the best rules are generated on record 309 with PRO 889 & PRO 868 PRO 869 rules.

Keywords: Data Mining; KDD; Asosiasi; Algoritma Apriori

Abstrak

Perusahaan setiap harinya menghasilkan data penjualan, semakin berjalannya waktu data tersebut meningkat dan jumlahnya menjadi sangat besar, data tersebut hanya disimpan tanpa mengerti manfaat yang terdapat dari data-data tersebut, karena keterbatasan dalam pengetahuan yang tepat dalam melakukan analisa pada data tersebut khususnya pada data transaksi penjualan. Demi mengatasi permasalahan tersebut dilakukan sebuah penelitian yang berfokus dalam mengolah kembali data transaksi penjualan tahun 2018 dengan pendekatan teknik data mining menggunakan konsep Knowledge Discovery in Database (KDD) menggunakan metode asosiasi dan algoritma apriori, serta menggunakan aplikasi pendukung yaitu RapidMiner. Tujuan penelitian ini berusaha membantu perusahaan dalam mengetahui kebiasaaan atau pola pembelian pelanggan berdasarkan data transaksi penjualan 2018. Hasil penelitian ini menghasilkan 316 aturan asosiasi dimana aturan terbaik dihasilkan pada record 309 dengan aturan PRO 889 & PRO 868 ⮕ PRO 869.

Kata kunci: Data Mining; KDD; Asosiasi; Algoritma Apriori

INTRODUCTION

Industries in the property sector, such as housing, shops, apartments, and so on, are currently experiencing significant progress due to increasing population growth (Fadilah, Purnama, & Zamrudy, 2019). The development of the property industry has also increased the need for building materials. One example of building materials that have an essential role is instant cement. Instant cement is a ready-to-use building material with the easy application just by mixing water, and it already contains sand, so there is no need to add other mixed ingredients (Susilowati, 2012). The development and competition in the instant cement industry are getting stricter with many competitors and high demand for instant cement products. Business people must be able to find solutions and create a marketing strategy so that the business develops and survives (Listriani, Setyaningrum, & Eka, 2018).

Companies can use technology to find a solution in formulating an appropriate marketing strategy, such as utilizing their sales transaction data (Wijayanti, 2017). Companies usually only use transaction data to view stock, calculate the number of sales and calculate monthly profit (Rahmawati & Merlina, 2018). Sales transaction data is mostly only collected and archived without analysis on sales transaction data because the data that is growing increasingly from time to time makes it difficult for companies to utilize the data and not know what to do with it (Samuel, Sani, Budiyantara, Ivone, & Frieyadie, 2022). Use sales transaction data by analyzing the behavior and habits of customers who often buy instant cement products. An approach used is market basket analysis, which t

Utilization of sales transaction data by analyzing customer behavior and habits regarding what instant cement products are frequently purchased. Market basket analysis approach to analyze consumer spending patterns (Fitrina, Kustanto, & Vulandari, 2018). Millions of data stored in databases must be applied to the data change process to become helpful information for companies by applying data mining techniques (Silvanie, 2020).

Data mining is a technique that aims to extract information from large data sets to analyze and extract data to gain knowledge (Adha, Sianturi, & Siagian, 201 C.E.). One approach to data mining in various fields is the a priori algorithm (Putra, Raharjo, Sandi, Ridwan, & Prasetyo, 2019).

The apriori algorithm can find all items of association rules in transaction data that meet the minimum and minimum requirements because it is easy to understand, and some research literature has been proposed (Lestari & Hafiz, 2020). One of the applications of the a priori algorithm in the research conducted by (Gumilang, 2021) is The problem faced with providing product data that many customers buy as data and implementing the a priori algorithm on web-based applications. The research results are applications built based on the application of the a priori algorithm consisting of item set selection and association rules.

The following research was conducted by (Junaidi, 2019). The problem in the company is that it has several data on sales of goods, but using this data is not to develop future sales plans. The solution is to use the frequent pattern growth method, and the company can make decisions in determining products that require more inventory than other products, with a reference threshold of 60% and 90% confidence and paying attention to the relationship between support and confidence.

Based on the results of observations and interviews with the company's sales management, Utilization of sales transaction data is not optimal for determining unique buying habits and patterns of customers when buying more than one type of goods in one transaction by looking at transaction records on the company's internal applications and excel data.

The application of data mining and an a priori algorithm approach can be a solution for companies to provide an overview in seeing customer purchasing patterns in order to be able to recommend products that are suitable for customers and as a reference to determine the amount of product production that is most sought after by customers. The transaction data used in this study is sales data for installed cement in 2018 in its application using the RapidMiner application in helping to find association rules. The expected results in this study are so that the company can recommend the right product to customers and make a more effective sales strategy in the future, which will come.

RESEARCH METHODS

Research using the concept of Knowledge Discovery in a Database to create association rules for instant cement sales transactions (Takdirillah, 2020).

Types of research

This study applies a qualitative method. Qualitative research has the advantage of gaining a deep and fundamental understanding of the object being observed in a systematic scientific manner.
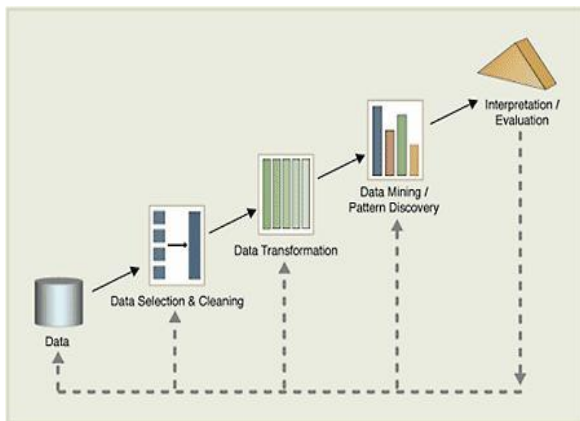
Time and Place of Research

The author researched instant cement manufacturing companies to obtain historical instant cement sales transaction data in 2018. Research time for two months by conducting research.

Research Target / Subject

The study aims to apply an a priori algorithm approach to find association rules and generate customer purchasing patterns at instant cement manufacturing companies based on sales transaction data that occurred in the 2018 period.

Procedure

The framework applied is based on the steps and procedures of this research process using Knowledge Discovery in Database.

Accredited rank 3 (SINTA 3), excerpts from the decision of the Minister of RISTEK-BRIN No. 200/M/KPT/2020

Source : (Takdirillah, 2020)
Figure 1. Knowledge Discovery in Database

Figure 1 describes the stages of the KDD process used to find and identify a pattern in the database. The data source of this research comes from instant cement sales transaction data in 2018. The data is selected and cleaned from noise, and the data is selected and transformed using the data mining process. Data mining involves looking for patterns or information from the a priori algorithm approach. The final stage is evaluation, namely translating the patterns generated from data mining and presented in a form that interested parties easily understand.

Data, Instruments, and Data Collection Techniques

This study uses primary data from interviews and instant cement sales transaction data from January 2018 – December 2018. Secondary data comes from research books and journals related to the research conducted. This study also conducted direct observations of the company to find out information related to existing sales information systems to assist in analyzing customer purchasing patterns and conducted interviews with sales managers regarding the instant cement sales transaction process to find out the workflow of related parties and obtain the required data.

Data analysis technique

Data analysis in Figure 2 describes the stages of the research carried out as follows:

1. Problem Identification
    The research begins by identifying the problems that occur in the company, namely having a lot of sales transaction data but only storing it in a database and not knowing how to use it, even though sales transaction data can be analyzed and to obtain important information contained therein.

2. Preprocessing
    Sales transaction data must be processed before data mining techniques can be applied by removing noise from the data, selecting data, and transforming data so it can be processed using data mining techniques.
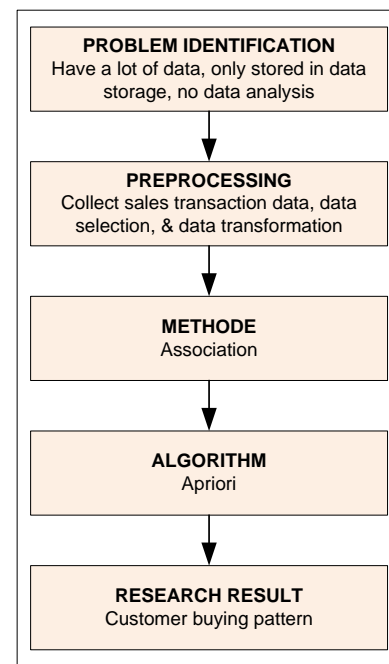

Figure 2. Research Stages

3. Method
    This study applies the association method to determine the relationship and interrelationships between products purchased in a sales transaction based on instant cement sales data for 2018.

4. Algorithm
    The algorithm approach applied in this research is the a priori algorithm. The primary process is to combine each item until no more combinations are formed using the minimum support parameter. (Anggraini, Putri, & Utami, 2020).
    At this point, the search for a combination of items that satisfies the minimum requirements of the support value is described in equation 1, and the confidence level is expressed in equation 2.

$$support(A \cap B) = \frac{\sum trans\ contain\ A\ and\ B}{Total\ transaction} \ldots\ldots\ldots\ldots(1)$$

Equation 1 is a formula for finding the support value of two items.

$$confidence\ P(B|A) = \frac{\sum trans\ contain\ A\ and\ B}{Total\ transaction\ A} \ldots\ldots\ldots(2)$$

Equation 2 is a formula for finding the confidence value by calculating the association rules A – B.

5. Research results

The results of this study are to obtain a pattern of instant cement purchase transactions so that companies know the best-selling products on the market, recommend the right products to customers, and increase sales of instant cement in the future.

RESULTS AND DISCUSSION

The observation and interview process results carried out at manufacturing companies producing instant cement obtained transaction data on sales of instant cement products in January - December 2018. Based on these data, the implementation of data mining techniques uses the association method. Association method to determine customer purchasing patterns through the a priori algorithm approach. The table below is a list of item codes and item names.

Table 1. List of Item Code and Description Item

No	Item Code	Description Item
1	PRO 686	Perekat Keramik
2	PRO 688	Perekat keramik area dalam
3	PRO 689	Perekat keramik area luar
4	PRO 787R	Topping screed
5	PRO 788R	Perata lantai
6	PRO 858	Plester finish
7	PRO 866	Acian profilan
8	PRO 867	Acian putih
9	PRO 868	Acian plesteran & beton
10	PRO 869	Acian plesteran
11	PRO 878	Plesteran premium
12	PRO 879R	Pasangan bata merah
13	PRO 888	Perekat pasangan beton ringan
14	PRO 889	Perekat pasangan beton ringan
15	PRO 889R	Perekat pasangan beton ringan

Table 1 explains the code for instant cement products and detailed descriptions of the existing products. The testing phase uses data from the transaction date and item code.

January		February		March	
03-01-18	PRO 868	01-02-18	PRO 858	04-03-18	PRO 858
	PRO 858		PRO 869		PRO 878
	PRO 868		PRO 879R		PRO 888
	PRO 869		PRO 888		PRO 889
	PRO 888		PRO 889	13-03-18	PRO 788R
	PRO 889		PRO 889R		PRO 858
12-01-18	PRO 869	08-02-18	PRO 688		PRO 867
	PRO 879R		PRO 689		PRO 868
	PRO 888		PRO 858		PRO 869
	PRO 889		PRO 869		PRO 878
	PRO 889R		PRO 879R		PRO 879R
20-01-18	PRO 878R		PRO 888		PRO 888
	PRO 868		PRO 889		PRO 889
	PRO 869		PRO 889R		PRO 889R
	PRO 888	15-02-18	PRO 686	17-03-18	PRO 879R
	PRO 889		PRO 688		PRO 889
26-01-18	PRO 869		PRO 689		PRO 889R
	PRO 868		PRO 788R	19-03-18	PRO 868
	PRO 869		PRO 868		PRO 869
	PRO 888		PRO 869		PRO 879R
	PRO 889		PRO 878		PRO 889
	PRO 889R		PRO 879R		PRO 889R
	PRO 889R		PRO 879R		PRO 889R
30-01-18	PRO688		PRO 888	29-03-22	PRO 688
	PRO 788R		PRO 889		PRO 689
	PRO 869		PRO 889R		PRO 788R
	PRO 879R	18-02-18	PRO 869		PRO 869
	PRO 888		PRO 878		PRO 879R
	PRO 889		PRO 879R		PRO 899
	PRO 889R		PRO 889		

Figure 3. Example of Sales Data 2018

The next stage is changing the data in tabulation form and saving it in an excel file. Then the process of analysis using the association method.

	B	C	D	E	F	G	H	I
1	PRO 686	PRO 688	PRO 689	PRO 787R	PRO 788R	PRO 858	PRO 866	PRO 867
2	1	0	0	0	0	1	0	0
3	0	0	0	0	0	1	0	0
4	1	0	1	0	0	0	0	0
5	0	1	0	0	0	1	0	0
6	0	0	0	0	0	0	0	0
7	0	1	0	0	1	0	0	0
8	1	1	0	0	0	0	0	0
9	0	0	1	0	0	1	0	0
10	0	1	0	0	0	0	0	0
11	0	0	0	0	0	0	0	0
12	0	0	0	0	0	0	0	0
13	0	0	0	0	0	0	0	0
14	0	0	0	0	0	0	0	0
15	0	1	0	0	1	0	0	0
16	0	1	1	0	0	0	0	0
17	0	1	0	0	1	1	0	0
18	0	0	0	0	0	1	0	0
19	0	0	0	1	0	0	0	0
20	0	0	0	0	1	0	0	0
21	0	1	0	0	0	1	0	0
22	1	1	1	0	0	1	0	0
23	0	0	1	0	0	0	0	0

Figure 4. Example of Data Tabulation

Figure 4 explains the results of changing data into tabulations in the form of binary numbers 0 and 1. The next step is importing the data that has become tabulated using the read excel operator in the RapidMiner application. The following is an analysis model created using the RapidMiner application.
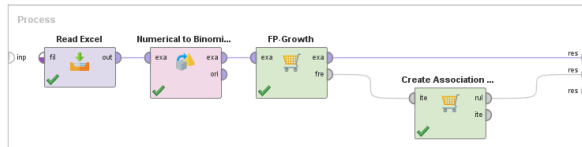

Figure 5. RapidMiner Model

Figure 5 is a model created in the rapid miner application, starting through the data import process using the read excel operator. The data is transformed using numerical to binomial operators, followed by an algorithmic approach using the FP-Growth operator and creating association rules with the create association rules operator.

The study carried out observations and several data tests to determine boundary parameters, with a minimum support value of 20% and a minimum confidence of 60%. The limit value was adjusted to obtain the most optimal association rule results.

Association Rule Result

The data processing results using the RapidMiner application obtain several association rules that provide many regulations for instant cement sales transaction data, and these results provide a reference to assist company decision-making. Based on the rules of the association, that has the highest level of trust, support, and confidence.

No.	Premises	Conclusion	Support	Confidence ↓	Lift
309	PRO 889, PRO 868	PRO 869	0.384	1	1.070
310	PRO 889, PRO 689	PRO 869	0.336	1	1.070
311	PRO 888, PRO 688	PRO 869	0.402	1	1.070
312	PRO 888, PRO 868	PRO 869	0.354	1	1.070
313	PRO 888, PRO 689	PRO 869	0.324	1	1.070
314	PRO 688, PRO 868	PRO 869	0.318	1	1.070
315	PRO 868, PRO 689	PRO 869	0.259	1	1.070
316	PRO 868, PRO 878	PRO 869	0.205	1	1.070
308	PRO 879R, PRO 888	PRO 869	0.577	0.995	1.065
306	PRO 868	PRO 869	0.497	0.994	1.064
307	PRO 889R, PRO 888	PRO 869	0.497	0.994	1.064
305	PRO 879R, PRO 868	PRO 869	0.473	0.994	1.063
304	PRO 879R, PRO 689	PRO 869	0.440	0.993	1.063
302	PRO 889R, PRO 868	PRO 869	0.408	0.993	1.062
303	PRO 889, PRO 688	PRO 869	0.408	0.993	1.062

Figure 6. Association Rules in RapidMiner

Figure 6 shows the association rules formed using the rapid miner application. Based on these results, rules number 309 – 316 get the highest confidence value of 100% and produce a lift value of 1,070, with a total of 316 rules formed.

Based on the results of these rules, it can be explained that.

- Customers buy PRO 889 & PRO 868 instant cement products, and the chance of customers also buying PRO 869 products is 100%.
- Customers buy PRO 889 & PRO 689 instant cement products, and the chance of customers also buying PRO 869 products is 100%.
- Customer buys PRO 888 & PRO 688 instant cement products, and the chance of customers also buying PRO 869 products is 100%.
- Customers buy PRO 888 & PRO 689 instant cement products, and the chance of customers also buying PRO 869 products is 100%.
- Customers buy PRO 688 & PRO 868 instant cement products, and the chances of customers also buying PRO 869 products are 100%.
- Customers buy PRO 868 & PRO 689 instant cement products, and the chance of customers also buying PRO 869 products is 100%.
- Customers buy PRO 868 & PRO 689 instant cement products, and the chance of customers also buying PRO 869 products is 100%.
- Customers buy PRO 868 & PRO 878 instant cement products, and the chance of customers also buying PRO 869 products is 100%.

RapidMiner application with a 100% confidence value recommends that PRO 869 is a solution for plastering plaster, while the highest demand is dominated by PRO 888 & PRO 889 types for adhesive solutions. Lightweight concrete pairs, so the PRO 869 type is the most suitable. It is recommended because after the process of installing lightweight brick materials for foundation walls, etc. The next stage is the wall plastering process, where the right product solution is PRO 869. This may occur due to the increasing housing, shops, office, or hotel property industry.

CONCLUSIONS AND SUGGESTIONS

Conclusion

The results of this study are recommended to provide consideration for company management to implement data mining methods in assisting decision-making by applying the association method to instant cement sales transaction data. The company can direct sales and marketing by providing customers with product recommendations and attractive promos based on the association rules. Based on the results of making association rules using instant cement sales data, all

regulations formed from processing the results of this study are recommended to provide consideration for company management to implement data mining methods in assisting decision-making by applying the association method to instant cement sales transaction data. The company can direct sales and marketing by providing customers with product recommendations and attractive promos based on the association rules.

Suggestion

The association method is a suitable recommendation to find association rules for the company's products to assist companies in formulating appropriate and accurate marketing strategies and production planning. This analysis can also be developed using transaction data in the following year, such as 2019-2021, to obtain more critical information from the resulting association rules. Research development can also be done by combining or comparing several other algorithms, such as linear regression algorithms, neural network support vector machines, etc.

REFERENCES

Adha, N., Sianturi, L. T., & Siagian, E. R. (201 C.E.). Implementasi Data Mining Penjualan Sabun Dengan Menggunakan Metode Apriori (Studi Kasus : PT. Unilever). *Informasi Dan Teknologi Ilmiah (INTI)*, *4*(3), 219–223. Retrieved from http://www.ejurnal.stmik-budidarma.ac.id/index.php/inti/article/view/426

Anggraini, D., Putri, S. A., & Utami, L. A. (2020). Implementasi Algoritma Apriori Dalam Menentukan Penjualan Mobil Yang Paling Diminati Pada Honda Permata Serpong. *Jurnal Media Informatika Budidarma*, *4*(2), 302. https://doi.org/10.30865/mib.v4i2.1496

Fadilah, M. F., Purnama, D., & Zamrudy, W. (2019). Pengaruh Waktu Hidrasi dan Methyl Cellulose terhadap Kuat Tekan dan Water Retension dalam Pembuatan Semen Instan. *Distilat: Jurnal Teknologi Separasi*, *5*(2), 222–227. https://doi.org/10.33795/distilat.v5i2.43

Fitrina, N., Kustanto, K., & Vulandari, R. T. (2018). Penerapan Algoritma Apriori Pada Sistem Rekomendasi Barang Di Minimarket Batox. *Jurnal Teknologi Informasi Dan Komunikasi (TIKomSiN)*, *6*(2), 21–27. https://doi.org/10.30646/tikomsin.v6i2.376

Gumilang, J. R. (2021). Implementasi Algoritma Apriori Untuk Analisis Penjualan Konter Berbasis Web. *Jurnal Informatika Dan Rekayasa Perangkat Lunak*, *1*(2), 226–233. https://doi.org/10.33365/jatika.v1i2.612

Junaidi, A. (2019). Implementasi Algoritma Apriori dan FP-Growth Untuk Menentukan Persediaan Barang. *Jurnal Sisfokom (Sistem Informasi Dan Komputer)*, *8*(1). https://doi.org/10.32736/sisfokom.v8i1.604

Lestari, A. F., & Hafiz, M. (2020). Penerapan Algoritma Apriori Pada Data Penjualan Barbar Warehouse. *INOVTEK Polbeng - Seri Informatika*, *5*(1), 96. https://doi.org/10.35314/isi.v5i1.1317

Listriani, D., Setyaningrum, A. H., & Eka, F. (2018). Penerapan Metode Asosiasi Menggunakan Algoritma Apriori Pada Aplikasi Analisa Pola Belanja Konsumen. *Jurnal Teknik Informatika*, *9*(2), 120–127. https://doi.org/10.15408/jti.v9i2.5602

Putra, J. L., Raharjo, M., Sandi, T. A. A., Ridwan, R., & Prasetyo, R. (2019). Implementasi Algoritma Apriori Terhadap Data Penjualan Pada Perusahaan Retail. *Jurnal Pilar Nusa Mandiri*, *15*(1), 85–90. https://doi.org/10.33480/pilar.v15i1.113

Rahmawati, F., & Merlina, N. (2018). Metode Data Mining Terhadap Data Penjualan Sparepart Mesin Fotocopy Menggunakan Algoritma Apriori. *PIKSEL : Penelitian Ilmu Komputer Sistem Embedded and Logic*, *6*(1), 9–20. https://doi.org/10.33558/piksel.v6i1.1390

Samuel, S., Sani, A., Budiyantara, A., Ivone, M., & Frieyadie, F. (2022). Sales Level Analysis Using the Association Method With the Apriori Algorithm. *Jurnal Riset Informatika*, *4*(4), 331–340. https://doi.org/10.34288/jri.v4i4.422

Silvanie, A. (2020). Pencarian Frequent Itemset dengan Algoritma Apriori dan Python. Studi kasus: Data Transaksi Penjualan Eceran Online di UK. *Jurnal Nasional Informatikan*, *1*(2), 103–113. Retrieved from http://ejournal-ibik57.ac.id/index.php/junif/article/view/124

Susilowati, A. (2012). Pemanfaatan Sampah Daun Sebagai Bahan Campuran Papan Panel Semen Dengan Bahan Ikat Mu Plesteran. *Jurnal Poli-Teknolog*, *11*(1), 31–38. Retrieved from https://jurnal.pnj.ac.id/index.php/politeknologi/article/view/20

Takdirillah, R. (2020). Penerapan Data Mining Menggunakan Algoritma Apriori Terhadap Data Transaksi Sebagai Pendukung Informasi Strategi Penjualan. *Edumatic : Jurnal Pendidikan Informatika*, *4*(1), 37–46. https://doi.org/10.29408/edumatic.v4i1.2081

Wijayanti, A. W. (2017). Analisis Hasil Implementasi Data Mining Menggunakan Algoritma Apriori pada Apotek. *Jurnal Edukasi Dan Penelitian Informatika (JEPIN)*, *3*(1), 60. https://doi.org/10.26418/jp.v3i1.19534