Analysis of Sales Transactions in Increasing Sales Promotion Using the WEKA Tool

Harma Oktafia Lingga Wijaya
Satrianansyah
Elmayati
Shelly Nadia
Abdurahman Sanjaya

Faculty of technic, Information Systems Study Program, Universitas Bina Insan, Lubuklinggau, South Sumatra,Indonesia
harmaoktafialingga@univbinainsan.ac.id, satrianansyah@univbinainsan.ac.id, elmayati@univbinainsan.ac.id, shellynadia3006@gmail.com, sanjayagc47@gmail.com

Abstract

Supermarket xyz is one of the growing basic food shops in Lubuklinggau City that provides daily staples. More than 50 types of food and non-food products are available at competitive prices, meeting almost all the daily needs of consumers. With a fairly large number of transactions, xyz supermarkets need analytical tools to provide useful information for xyz supermarkets in marketing strategies, what goods are most in demand by consumers and others [1]. The problem that exists is the sales transaction data of XYZ supermarkets. has not been able to provide information about the pattern or relationship of a set of items purchased by customers. To help the algorithm process using a priori where a priori will help solve the existing stages, for mining a test is carried out with the Weka application, WEKA is a practical tool. This study processes sales transaction data as many as 1339 transactions and produces 6 best rules that can be used by XYZ Supermarkets as a marketing strategy in the future.

Keywords Analisis, Data mining, apriori,Weka

1. Introduction

In modern times like today, developments in various business fields occur very quickly. Not only in the field of information technology which has recently experienced very rapid development but also other fields that aim to fulfill and satisfy the needs and lifestyles of consumers [1]. The development of information technology is needed by all companies, especially for companies that are developing. As at this time all sectors enter people's lives, we can't deny it all spreads to all sectors without exception to the company. companies, so companies must manage their companies well so that the goals can be achieved.

Marketing is an important part of running a business. There needs to be good management so that the business can grow and develop better so that it can compete with its competitors. One of the important elements in marketing is the marketing strategy and marketing mix. A marketing strategy is needed so that market segments, target market determination, and market positioning can be properly selected [2]. Utilization of existing data in the information system to support decision-making activities, it is not enough to just rely on operational data, a data analysis is needed to explore the potential of existing information. Decision makers try to take advantage of existing data warehouses to explore useful information to help make decisions, this encourages the emergence of new branches of science to overcome the problem of extracting important or interesting information or patterns from large amounts of data, which is called data mining [3].

This XYZ supermarket has recorded all sales transactions into the information system, only the existing data is not used as new information for the next marketing strategy. Data Analysis This sales transaction will combine with a computer, namely data mining techniques to find the rules of a combination of items. One of the stages of association analysis that has attracted the attention of many researchers to produce efficient algorithms is high-frequency pattern mining. The output of data mining can be used to improve future decision making. The purpose of using data mining with the Apriori Algorithm is to find out the results of product sales data processing, so that it can

make a marketing strategy to place products that are often purchased, making it easier for consumers to find the location of products to save time. Association analysis or association rule mining is a data mining technique to find the rules of a combination of items. One of the stages of association analysis that has attracted the attention of many researchers to produce efficient algorithms is high-frequency pattern mining. Support is the value of support or percentage combination of an item in the database. While confidence is the value of certainty, namely the strength of the relationship between items in an a priori [3]. The purpose of using data mining with the Apriori Algorithm is the result of processing product sales data, so that it can make marketing strategies to place products that are often purchased, making it easier for consumers to find product location to save time[4].

The method used in this research is the a priori algorithm with data processing using WEKA (Waikato Environment for Knowledge Analysis) tools. Research with the title Data Mining Analysis of Tire Sales Using the C4.5 Algorithm. Some studies that become reference material in this study are by [11]. In this study, a method is proposed in digging up sales transaction data information using the C4.5 algorithm to determine and analyze the amount of inventory in sales transactions. The existing data or information will then be classified to find a model that describes the data class which will later be converted into a dataset. Results Based on the results of testing and implementation of Data Mining using the C4.5 algorithm in classifying tire sales data at CV. RODA INTI MAS, it can be concluded that the testing process with an accuracy rate between 74% - 76%, with the root attribute being a category, and the attributes being branches are units, while the attributes being branches are many and few. The virtue of this research is the creation of a consumer buying pattern so that it can be used to assist in determining the marketing strategy of XYZ Supermarkets in the future.

1.1 Objectives

The importance of association rules can be seen from two parameters, namely minimum support (percentage of item combinations from all transactions) and minimum confidence (strong relationship between items in associative rules). In contrast to the application of the a priori algorithm in the above system, this study aims to examine the application of this algorithm and association rule to the sales transaction dataset to provide a sales strategy for XYZ supermarkets. which was established in early 2008 and sells various kinds of community needs/groceries. The dataset is in the form of sales transaction data for the months of January – December 2021 and January – May 2022. Analysis of the highest frequency pattern produces a combination of items that meet the specified minimum support requirements. The formation of the association rule pattern is carried out to find association rules that meet the minimum confidence of the high frequencies that have been found. From the references used and the existing problems, the researchers are interested in making the topic in this study, the novelty of this research is to make consumer buying patterns so that they can help supermarkets in making marketing strategies for the future. Flowchart is a chart with certain symbols that describe the sequence, the process in detail with the relationship between a process (instructions) with other processes in a program. The program flowchart is a more detailed description of how the actual procedure is carried out by a program. This flowchart describes the logical sequence of a problem-solving procedure with a flowchart as follows.

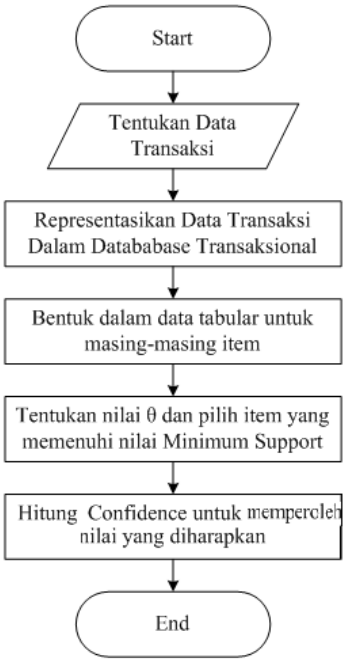


Figure 1 Solution flow

2. Literature Review

The author uses scientific journals, proceedings and e-books including Data Mining using Apriori Algorithm for Product Recommendation for Customers by Ariefana [5]. Implementation of Apriori Algorithm to Find Associations of Goods Sold in E-commerce Order Mas [6]. Implementation of Apriori Algorithm to Data Sales at Retail Companies [7] , "Prediction of Goods Sales Patterns in XYZ MSMEs with the Apriori Algorithm Method" [1], "Implementation of Rule Mining Associations on Sales Transaction Data Using Apriori Algorithms" [3] and the ISSN scientific journal entitled "Application of Apriori Algorithms On Cosmetic Sales Data at Sharly Stores, Lubuklinggau City"[8], "Implementation of Apriori Algorithms on Sales Data in Retail Companies" [9], as well as an e-book entitled "Data Mining: Practical Machine Learning Tools and Techniques 3rd Edition" [6] , "Discovering Knowledge in Data, an Introduction to Data Maning" [5], "Data Mining: Concepts and Techniques Third Edition" [4], as well as several in international and national journals regarding data mining and the availability and interrelationships between goods using other a priori algorithms.

3. Methods

This study applies the a priori algorithm to the dataset in the form of sales transaction history data. The stages of the research are data collection, data pre-processing, analysis of the highest frequency pattern using a priori algorithms, formation of association rule patterns, and testing of experimental results. The application of association rules with the a priori algorithm has advantages in simplicity and the ability to handle large data so that it is easier to use practically by XYZ supermarkets with limited data processing capabilities. This study uses the Weka 3.8 application.

Testing the results of the analysis is very important to determine and ensure whether the results of the analysis are correct or not. The software that the author uses in this test is Weka. The stages in testing on Weka are as follows:

Activate the Weka software so that the following image appears.

Figure 2 Weka Application

After the Weka window opens, click the Explorer menu, click the Preprocess menu, Open File, search for the database as shown below.

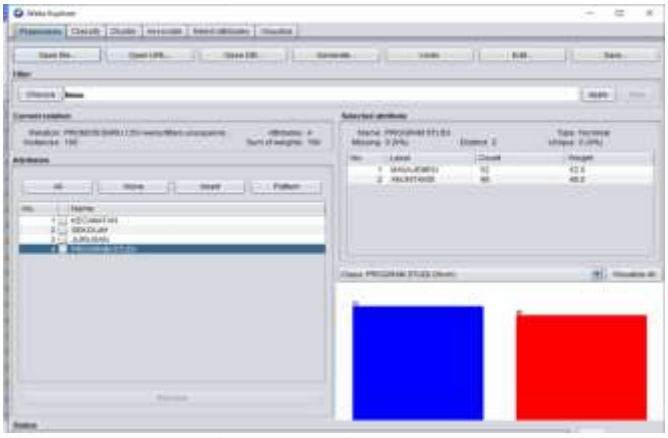


Figure 3. Explorer menu

The next step is to click the Associate menu, select Apriori and then click Start to bring up the formed rule, as shown below;

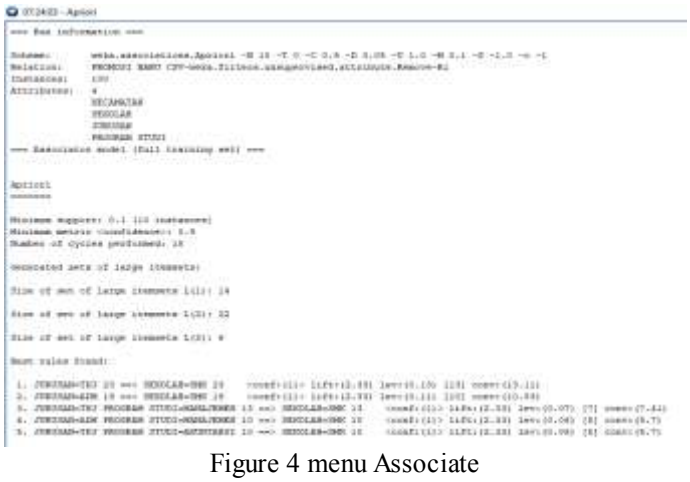


Figure 4 menu Associate

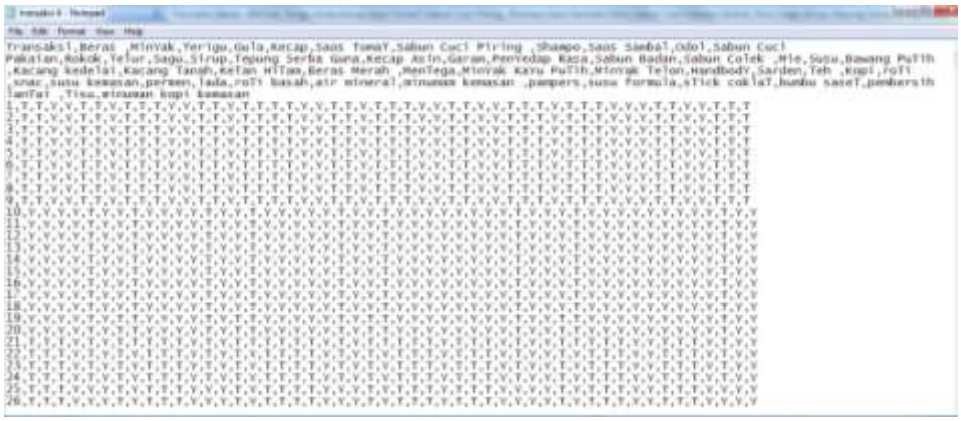
4. Data Collection

In this study, the data processing is carried out with data that has been tabulated into Microsoft Excel tables. Data processing uses transaction data for 7 months starting from November 2021 to May 2022. From this data obtained a total of 1339 sales transaction data, including the sales transaction table.

Table 1. Tabulation of sales transactions

Transaksi	Beras	Minyak	Terigu	Gula	Kecap	Saos Tomat	Sabun Cuci Piring	Shampo	Saos Sambal	Odol	Sabun Cuci Pakaian	Bokok	Telur	Sagu	Sirup	Tepung Serba Guna	Kecap Asin	Garam	Penyedap Rasa	Sabu Badu
1	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T
2	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T
3	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T
4	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T
5	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T
6	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T
7	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T
8	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T
9	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T
10	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T
11	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T
12	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T
13	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T
14	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T
15	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T
16	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T
17	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T
18	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T
19	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T
20	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T

Table 2. Sales Transactions in Csv. format



5. Results and Discussion

From the sales transaction data, it was obtained a total of 1339 sales transaction data including can see in table 2 sales transactions, as for the types of products sold including: rice, oil, flour, sugar, soy sauce, tomato sauce, dish soap, shampoo, chili sauce , toothpaste, laundry soap, cigarettes, egg, sago, syrup, all purpose flour, soy sauce, salt, flavoring, body soap, dab soap, noodles, milk, garlic , soybeans, peanuts, black sticky rice, rice red, butter, eucalyptus oil, telon oil, handbody, sardines, tea, coffee, bread, snack, bottled milk, sweets, pepper, wet bread, mineral water, bottled drinks, pampers, formula milk, chocolate sticks, flavoring sachets, floor cleaner, tissue, packaged coffee drinks.

All the data above is made in the form of an excel file, then saved in the form of a .csv file extension. after being saved in .csv form, the file is opened with the WEKA 3.8.3 tool, to be converted into a file with .arff extension. as for the form of data with extensions.

5.1 **Data Visualization Results of Each Attribute With Weka**

The following is the result of visualizing the entire product using the WEKA tools, as shown in Figure 1.

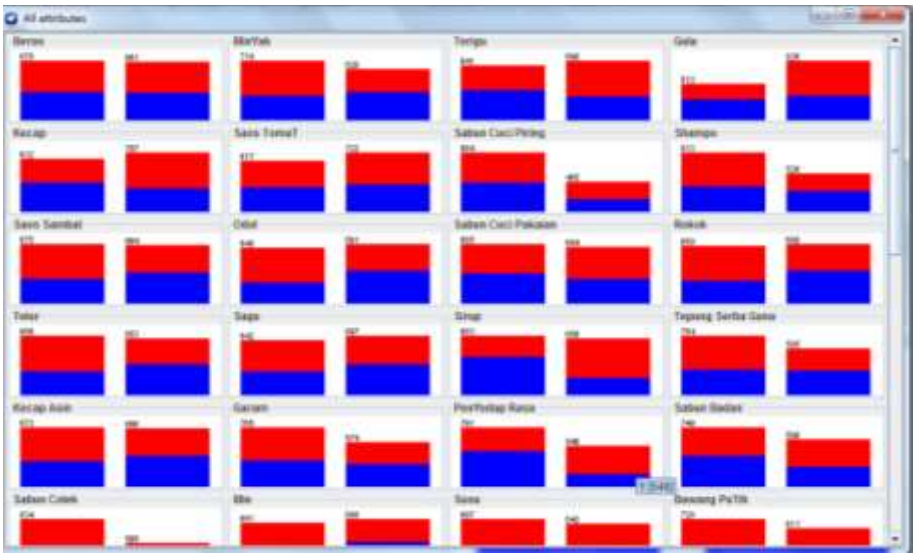


Figure 5 visualization of the whole product

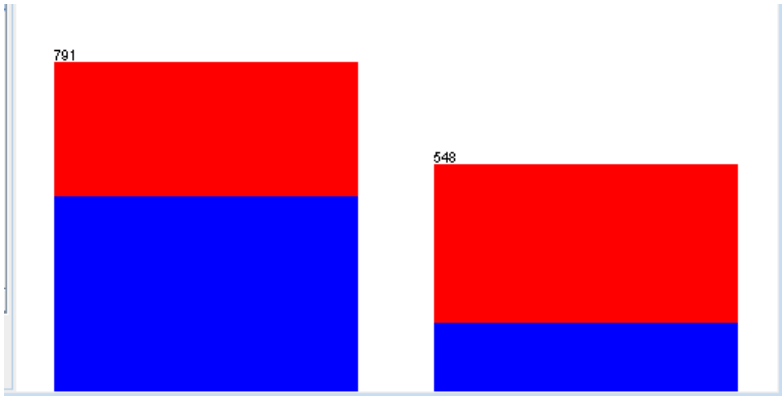


Figure 6 Visualization of flavoring

Based on Figure 2 above, it is known that from 1339 sales transactions, 791 did not make a purchase and 548 made a purchase on flavoring products.

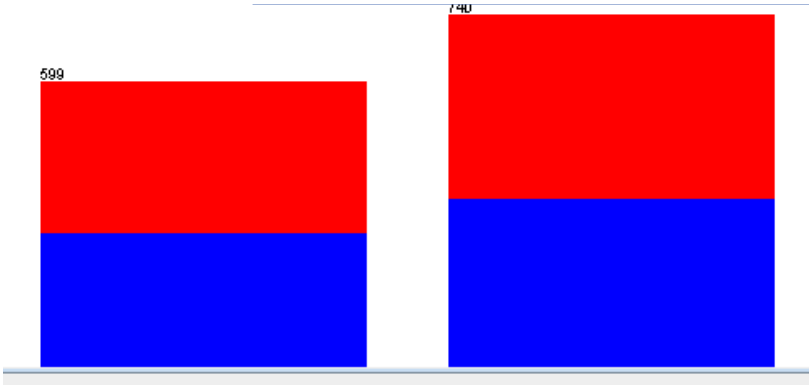


Figure 7 Visualization of black sticky rice products

Based on Figure 3 above, it is known that from 1339 sales transactions, 599 did not make purchases and 740 made purchases on black sticky rice products.

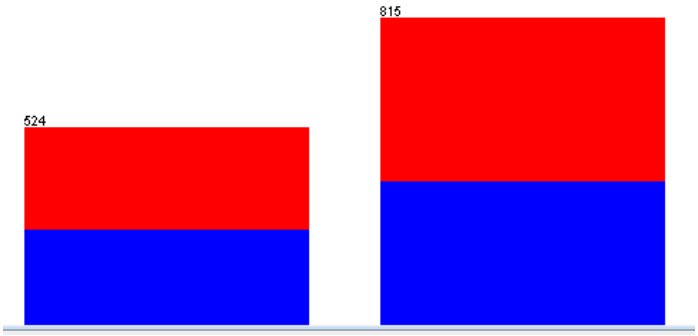


Figure 8 Snac product visualization

Based on Figure 3 above, it is known that from 1339 sales transactions, 524 did not make purchases and 815 made purchases on snack products.

5.2 **Association Analysis Results With Weka Tools 3.8**

Figure 9 Weka Analysis Results

From the results above, it can be seen that consumers' purchasing patterns if they buy coffee will buy a sachet of seasoning, if they buy packaged coffee drinks, they will buy packaged milk, if they buy black sticky rice, they will buy dish soap. Some of the rules formed above can be used as a reference in product recommendations as in [1]. Marketing strategies can be carried out, for example product bundling (packaged milk and packaged coffee will be sold as one saving package), buy 2 get 1 (if you buy coffee 2 you get 1 packet of seasoning sachets), and product discounts (get a 5% discount if you buy packaged coffee). above 10 pcs). Thus, the company has a reference for handling marketing strategies, not just estimates. The results of this study can be implemented by the company to support the marketing strategy.

5.3 Proposed Improvements

Efforts to improve results can still be increased, for example by modifying the minimum support and confidence values to get better or applying other association algorithms, such as the FP-Growth algorithm [15], Generalized Rules Induction, or hash-based algorithms and can also perform comparisons. of 2 algorithms, so you can get better results.

6. Conclusion

A priori algorithm that is tested on a flexible sales transaction dataset to be used as a basis for making company decisions in the marketing area. The association rules formed can be used as a reference for product recommendations that meet the minimum confidence and support values, and the Apriori Algorithm tested with the WEKA application produces a combination of itemset patterns and rules as useful knowledge and information from sales or transaction data in a store/ company, the results of consumer buying patterns can be used as a reference for marketing strategies that will be carried out in the future.

References

(Anas, 2020; Ini, Nofriansyah, & Syaifuddin, 2020; Nugraha, Aribawa, Priyana, & Indrawan, 2016; Nurjoko - & Kurniawan, 2016; Oktafia, Wijaya, Yuliansyah, et al., 2022; Oktafia, Wijaya, S, & Sari, 2022; Putra, Raharjo, Sandi, Ridwan, & Prasetyo, 2019; Riszky & Sadikin, 2019; S, Oktafia, Wijaya, & Sari, 2022; Sholik & Salam, 2018; Srikanti, Yansi, Norvahina, Permana, & Salisah, 2018; Syahdan & Sindar, 2018; Winasis, Program, Magister, Komputer, & Luhur, 2019)Anas, A. (2020). Implementasi Algoritma Apriori Untuk Menentukan Strategi Promosi STIE-Graha Karya Muara Bulian. *Jurnal Ilmiah Media Sisfo*, *14*(1), 64. https://doi.org/10.33998/mediasisfo.2020.14.1.790

Ini, C. A., Nofriansyah, D., & Syaifuddin, M. (2020). Market Based Analysis Menggunakan Algoritma Apriori Pada Data Transaksi Penjualan Parfume APF Barokah, *3*(8), 1366–1373.

Nugraha, P. G. S. C., Aribawa, I. W., Priyana, I. P. O., & Indrawan, G. (2016). Penerapan Metode Decision Tree(Data Mining) Untuk Memprediksi Tingkat Kelulusan Siswa Smpn1 Kintamani. *Seminar Nasional Vokasi Dan Teknologi (SEMNASVOKTEK)*, 35–44.

Nurjoko -, & Kurniawan, H. (2016). Aplikasi Data Mining Untuk Memprediksi Tingkat Kelulusan Mahasisa Menggunakan Algoritma Apriori Di IBI Darmajaya Bandar Lampung| - | Jurnal Teknologi Informasi Magister. *Jurnal Teknologi Informasi Magister*, *2*(1), 79–93. Retrieved from https://jurnal.darmajaya.ac.id/index.php/jtim/article/view/650

Oktafia, H., Wijaya, L., S, A. A. T., & Sari, W. M. (2022). Prediksi Pola Penjualan Barang pada UMKM XYZ dengan Metode Algoritma Apriori, *3*, 432–437. https://doi.org/10.30865/json.v3i4.4200

Oktafia, H., Wijaya, L., Yuliansyah, C., Sistem, P., Universitas, I., Insan, B., … Apriori, A. (2022). IMPLEMENTASI ASOSIASI RULE MINING PADA DATA TRANSAKSI, *1*(1), 30–35.

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

Riszky, A. R., & Sadikin, M. (2019). Data Mining Menggunakan Algoritma Apriori untuk Rekomendasi Produk bagi Pelanggan. *Jurnal Teknologi Dan Sistem Komputer*, *7*(3), 103–108. https://doi.org/10.14710/jtsiskom.7.3.2019.103-108

S, A. A. T., Oktafia, H., Wijaya, L., & Sari, W. M. (2022). Sistem Prediksi Pertumbuhan Ekonomi Kabupaten Musi Rawas , Kabupaten Musi Rawas Utara Dan Kota Lubuklinggau Dengan Metode Regresi Linier, *3*, 422–431. https://doi.org/10.30865/json.v3i4.4198

Sholik, M., & Salam, A. (2018). Implementasi Algoritma Apriori untuk Mencari Asosiasi Barang yang Dijual di E-commerce OrderMas. *Techno.COM*, *17*(2), 158–170.

Srikanti, E., Yansi, R. F., Norvahina, Permana, I., & Salisah, F. N. (2018). Penerapan Data Mining Untuk Menganalisis Penjualan Barang dengan Menggunakan Metode Apriori pada Supermarket Sejahtera Lhoksumawe. *Jurnal Ilmiah Rekayasa Dan Manajemen Sistem Informasi*, *4*, 77–80.

Syahdan, S. Al, & Sindar, A. (2018). Data Mining Penjualan Produk Dengan Metode Apriori Pada Indomaret Galang Kota. *Data Mining Penjualan Produk Dengan Metode Apriori Pada Indomaret Galang Kota*, *1*.

Winasis, P. H., Program, M., Magister, P., Komputer, I., & Luhur, U. B. (2019). Penerapan Data Mining Untuk Analisis Pola Apriori Pada Mall Cpm Jakarta, *2*(2).