

Analysis of Goods Stock Using the Apriori Algorithm to Aid Goods Purchase Decision Making

Nur Azis^{1)*}, Purwo Agus Sucipto²⁾, Agus Herwanto³⁾, Era Sari Munthe⁴⁾, Dola Irwanto⁵⁾

^{1)*}Sekolah Tinggi Teknologi Duta Bangsa, Bekasi, Indonesia, ^{2,4)}Universitas Jayabaya, Jakarta, Indonesia, ³⁾Universitas Esa Unggul, Jakarta, Indonesia, ⁵⁾Universitas Pamulang, Banten, Indonesia

^{1)*} azis.nur@gmail.com, ²⁾ purwoagussucipto@gmail.com, ³⁾ agus.herwanto@esaunggul.ac.id,

⁴⁾ sari_mun@yahoo.com, ⁵⁾ dosen01115@unpam.ac.id

Submitted : Dec 31, 2023 | **Accepted** : Jan 9, 2023 | **Published** : Jan 10, 2024

Abstract: After the covid-19 pandemic outbreak and the high uncertainty index during the covid-19 pandemic. The business world is experiencing a huge impact in addition to the sluggish interest of buyers is also limited in its movement. On this occasion, the researcher intends to provide an overview that can help business people, especially in purchasing goods that are useful for filling the stock of goods in the warehouse. To get maximum results and minimum error rate. Researchers use the Apriori Algorithm in analyzing stock items and use the Tanagra version 1.4 application. Research data used the sales history of the past 1 year here the data used is between May 2022 and April 2023. With a total itemset of 375. But after applying the Golden Rule (threshold), there are only 10 products with sales reaching 1623 items. This research produces a final ordered association based on the minimum support and minimum confidence that has been determined, namely 12 rules with a combination of 2 itemsets with a confidence value of 100%.

Keywords: Covid-19; Business Analysis ; Apriori Algorithm; Tanagra

INTRODUCTION

The business world is currently still affected by the Covid-19 pandemic that has occurred in recent years, so it requires an appropriate strategy so that the business being run can continue to run. (Azis, 2023). Apart from strategy, a high and strong work ethic is also required (Halim, 2021). On this occasion, the researcher intends to conduct research and it is hoped that the results of this research will become a reference for business actors in running a business, especially in terms of purchasing goods.

Purchasing goods here means filling the stock of goods in the warehouse which will later be sold. How important is stock management, According to (Sika, 2021) Stock management in a company is very important because it makes the sales process smooth. Whereas (Saputra & Qoiriah, 2022) that good stock management can function to avoid goods piling up in warehouses. This research uses an a priori algorithm. The Apriori algorithm is very efficient and can speed up the process of forming trends in itemset combination patterns from sales results (Thariq, 2023) Besides that, this algorithm is very popular in searching for frequent itemsets as well as determining or finding rules from large databases (Hernando, 2019).

In this study, the method used is to collect valid data on sales results for the last 1 year. The importance of using sales data is useful for digging up information and being able to predict how much stock must be prepared to be placed in the store so that there is no accumulation of excess stock (Asward & Brotosaputro, 2021). And from the results of this sales pattern can be used as a consideration in developing sales strategies, namely recommendations for making menu packages and recommendations for offering additional products or commonly referred to as cross-selling (Kurniana et al., 2023). Besides

*Nur Azis



that, sales data is obtained from activities carried out in the life of daily sales activities (Akbar et al., 2020).

Furthermore, the data obtained will be processed. Data itself is a collection of unprocessed facts or information such as numbers, words, images, or symbols obtained through a search or observation process from certain sources (Kwintiana et al., 2023). The processed data will later be useful for decision making. Because to produce good policies and decisions, it is necessary to process data into information that is relevant to the problems the company is facing (Sudjiman & Sudjiman, 2018).

LITERATURE REVIEWS

Additional research on stock analysis has been conducted by (Asward & Brotosaputro, 2021; Saptadi et al., 2023), which explains that predicting the quantity of stock is useful for determining the need for inventory of goods to prevent the accumulation of stock items in the warehouse. (Asana et al., 2022) analyze information on purchasing decision activities from the customer's perspective in order to determine consumer shopping patterns with regard to purchasing goods. Subsequent investigations into this sales trend may inform the formulation of a sales strategy, specifically suggestions for menu package development and the promotion of supplementary products, also known as cross-selling (Kurniana et al., 2023). Sales data obtained from daily sales activities is crucial for the analysis of stock items, as it can serve as the foundation for the procurement of products (Akbar et al., 2020; Panjaitan et al., 2019). (Dewantara & Giovanni, 2023) conducted research that elucidates the process of optimizing stock analysis for the purpose of analyzing customer sales forecasting. The research gap at hand pertains to the utilization of the a priori algorithm in stock analysis, which hinders the investigation of research implications by merely confirming the availability of stock items and serving as a resource for customers' knowledge of product purchases.

METHOD

Apriori Algorithm

The a priori algorithm aims to determine the sales of goods that are most in demand by consumers (Sriyanto et al., 2022). The Apriori algorithm is an algorithm that searches for frequent itemsets using association rule techniques (Damanik et al., 2022). The a priori algorithm also uses previous knowledge of an itemset that appears frequently, which is called a frequent itemset (Prasetya et al., 2022).

Confidence Value

The confidence value is the value of trust or certainty of the itemset-itemset relationship (Valeska et al., 2022). Confidence (certainty value) is the strength of the relationship between items in the associative rule (Herlawati et al., 2022). The calculation of the confidence value can be seen in the formula below.

$$\text{Confidence (A} \Rightarrow \text{B)} = \frac{\text{Support (A dan B)}}{\text{support (A)}} \times 100\% \quad (1)$$

Decision Support Systems

Decision Support Systems or often called Decision Support Systems (DSS) are model-based systems consisting of procedures for processing data and considerations to assist managers in making decisions. (Akmaludin et al., 2023; Sudipa, Wardoyo, et al., 2023). In Computer Science, there is a way to find the feasibility of a desired alternative according to the criteria determined by the company (Angela et al., 2022; Kaunan et al., 2023).

Research Flow

The research flow is presented in Figure 1. So that later it can help and provide an overview of the stages that need to be carried out in carrying out a research.

*Nur Azis



This is an Creative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

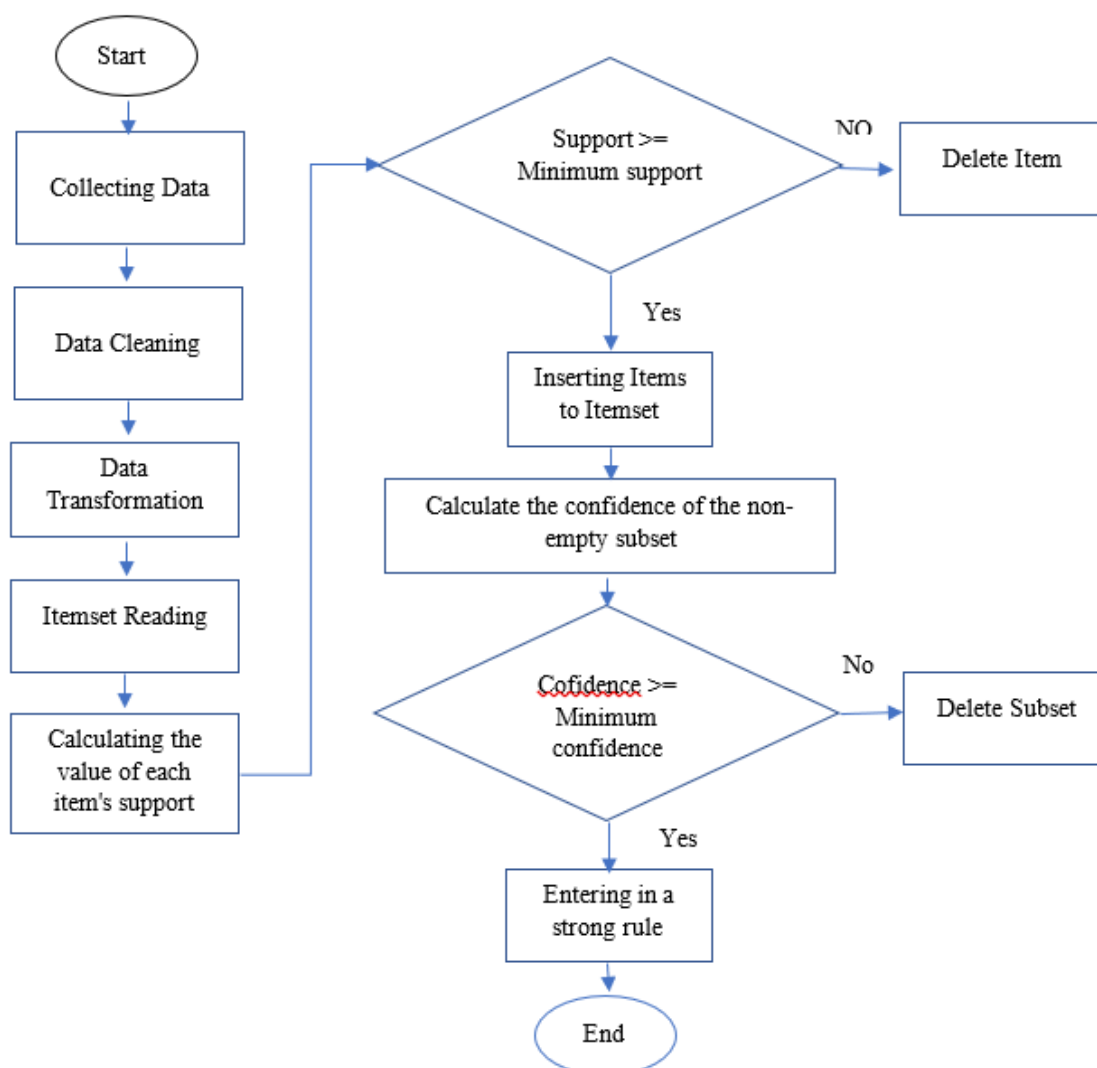


Fig.1 Research Flow Diagram

Based on Figure 1, the stages of the research conducted can be explained and completed with the following explanation.

- a) Collecting data
Data collection is an integral part of data analysis activities (Azis & Handoko, 2021). Data collection can be said to be the procedure of collecting, measuring, and analyzing accurate insights for research using standardized, validated techniques(Sileyew, 2019; Sudipa, Udayana, et al., 2023).
- b) Cleaning Data
The process of preparing data for analysis by deleting or modifying incorrect, irrelevant, duplicate and unformatted data (Ridzuan & Zainon, 2019).
- c) Data Transformation
To provide optimal results in data modeling (Modelling), some data requires a special format. Therefore, Data Transformation is carried out so that the data can be used during modeling. This process can also influence the results of the modeling presented at the evaluation stage(Rosmayati et al., 2023).
- d) Reading Itemset
Itemset is the need to develop a data structure or algorithm, every time a user wants to extract a different type of itemset(Fujioka & Shirahama, 2022).
- e) Calculate the value of each support for each item

*Nur Azis



In this process, C1 is formed by calculating the number of transactions for each item. Before the calculation process is carried out, the thing that must be done is to determine the Golden rule (threshold). To determine the Golden rule or the minimum number of transactions that will be carried out in the calculation process, each researcher can determine it by looking at the pattern of the data used. For this research, the Golden rule (threshold) used is: 20% or a minimum number of sales of 75. The type of sales item that will be calculated is a minimum number of 75.

- f) Inserting items into itemset
In this process, calculations are carried out by calculating 1 or more itemsets
- g) Calculates the confidence of a non-empty subset
This process is calculated based on the minimum support and minimum confidence that have been determined
- h) Include strong rules
In this process is the determination of the final ordered association rule based on the minimum support and minimum confidence that has been determined.

RESULTS

Apriori Implementation with the Tanagra Application

The first step is collecting data that will be tested to connect to the database on Tanagra. Tanagra application has various modeling algorithms including a priori algorithm in data processing and is equipped with data visualization and big data processing capabilities.

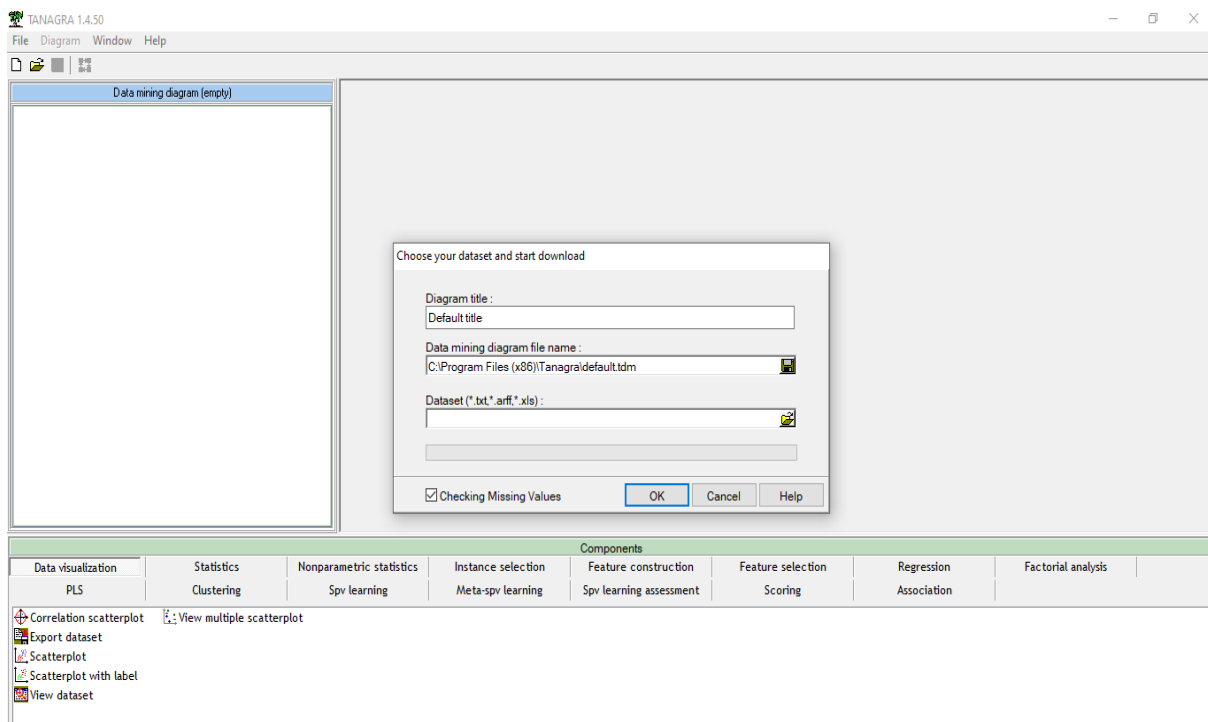


Fig.2 Initial display of data collection

Figure 2 shows the initial appearance of the Tanagra application which will retrieve data and then connect. After the data is connected, the next step is to select the items in the attributes. The next step is to perform frequent itemsets. Or in other terms, determining min support is in accordance with Figure 3.

*Nur Azis



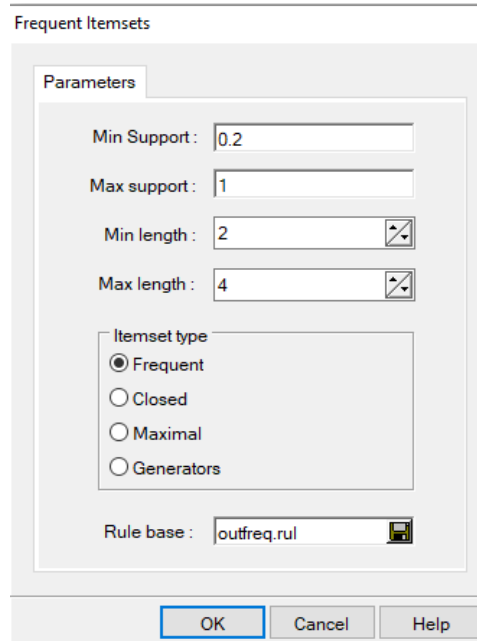


Fig.3 Frequent Itemsets display

The final step is to display the final association results, namely right click on Apriori then click View, then the final association results will appear. As seen in figure 4. Minimum Support is determined based on the lowest value used to determine frequent itemsets to be considered significant while maximum support is the top value that is acceptable to extract association rules.

Number of rules : 160					
N ^o	Antecedent	Consequent	Lift	Support (%)	Confidence (%)
1	"Refil Catridge Hp 680 Black=true" - "Catridge Original Canon 810 Black=true" - "Tinta Original Epson 664 Black=true"	"Toner Original Hp Laserjet 85A=true"	1.11111	80.000	88.889
2	"Refil Catridge Hp 680 Black=true" - "Catridge Original Canon 810 Black=true" - "Toner Original Hp Laserjet 85A=true"	"Tinta Original Epson 664 Black=true"	1.11111	80.000	100.000
3	"Refil Catridge Hp 680 Black=true" - "Refil Catridge Canon 745 Black=true" - "Toner Original Hp Laserjet 85A=true"	"Tinta Original Epson 664 Black=true"	1.11111	80.000	100.000
4	"Toner Original Hp Laserjet 85A=true"	"Refil Catridge Hp 680 Black=true" - "Tinta Canon Black 100ml=true" - "Tinta Original Epson 664 Black=true"	1.11111	80.000	100.000
5	"Refil Catridge Hp 680 Black=true" - "Tinta Original Epson 664 Black=true"	"Refil Catridge Canon 745 Black=true" - "Toner Original Hp Laserjet 85A=true"	1.11111	80.000	88.889
6	"Refil Catridge Hp 680 Black=true" - "Tinta Original Epson 664 Black=true"	"Catridge Original Canon 810 Black=true" - "Toner Original Hp Laserjet 85A=true"	1.11111	80.000	88.889
7	"Catridge Original Canon 810 Black=true" - "Toner Original Hp Laserjet 85A=true"	"Refil Catridge Hp 680 Black=true" - "Tinta Original Epson 664 Black=true"	1.11111	80.000	100.000
8	"Tinta Original Epson 664 Black=true"	"Refil Catridge Hp 680 Black=true" - "Catridge Original Canon 810 Black=true" - "Toner Original Hp Laserjet 85A=true"	1.11111	80.000	88.889
9	"Catridge Original Canon 810 Black=true" - "Tinta Original Epson 664 Black=true"	"Refil Catridge Hp 680 Black=true" - "Toner Original Hp Laserjet 85A=true"	1.11111	80.000	88.889
10	"Refil Catridge Hp 680 Black=true" - "Refil Catridge Canon 745 Black=true" - "Tinta Original Epson 664 Black=true"	"Toner Original Hp Laserjet 85A=true"	1.11111	80.000	88.889
11	"Refil Catridge Hp 680 Black=true" - "Toner Original Hp Laserjet 85A=true"	"Catridge Original Canon 810 Black=true" - "Tinta Original Epson 664 Black=true"	1.11111	80.000	100.000
12	"Toner Original Hp Laserjet 85A=true"	"Tinta Epson Black 100ml=true" - "Tinta Original Epson 664 Black=true"	1.11111	80.000	100.000
13	"Tinta Original Epson 664 Black=true"	"Tinta Epson Black 100ml=true" - "Toner Original Hp Laserjet 85A=true"	1.11111	80.000	88.889
14	"Refil Catridge Canon 745 Black=true" - "Toner Original Hp Laserjet 85A=true"	"Tinta Epson Black 100ml=true" - "Tinta Original Epson 664 Black=true"	1.11111	80.000	100.000
15	"Tinta Original Epson 664 Black=true"	"Tinta Canon Black 100ml=true" - "Refil Catridge Canon 745 Black=true" - "Toner Original Hp Laserjet 85A=true"	1.11111	80.000	88.889

Fig.4 Final Association Results

*Nur Azis



After testing, the results obtained meet min.support and min. Confidence, the result is 72 itemsets. Based on the 160 rules used in the study, the results of the final association show several items with the highest association results, namely the lift ratio value of 1.11111 with a support value of 80% and a confidence value of 100%.

DISCUSSIONS

In this research, the data that will be tested is past data. As for what was taken between May 2022 and April 2023. After collecting it, it turned out that we were looking for thousands so that if it was processed it would take a relatively long time. In order to simplify calculations, sample data will be taken. Namely, taking sales data for one of the months between a predetermined time period. The data taken is data for January 2023. Next, the data will be filtered by applying the Golden rule (threshold) or a minimum transaction point of 10% of the number of transactions. In this research, the number of transactions is the number of operations or openings, in other words the number of CV days. Delta Computer is open in January 2021, which is 27 days. So the Golden rule (threshold) value is 10% multiplied by 27 days. The minimum transaction amount is 2.7 so round it up to 3 items. In other words, only products with a minimum of 3 sales transactions will be followed up for the calculation process in determining the itemset or entered in tabular format.

Table 1 Recap of Sales Data for May 2022 to April 2023 After implementing the Golden Rule (threshold)

Number	Product name	Amount
1	Original Canon 810 Black cartridge	81
2	Refill cartridge Canon 745 Black	86
3	Canon 810 Black Cartridge Refill	84
4	HP 680 Black Cartridge Refill	91
5	HP Laserjet 85A toner refill	212
6	Canon Black 100ml ink	179
7	Epson Black 100ml ink	108
8	Original Epson 664 Black ink	79
9	Original HP Laserjet 107A toner	108
10	Original HP Laserjet 85A toner	115
Amount		1623

(Source: Author 2023)

Then, calculations are carried out to form itemsets. The formula for calculating 1 itemset is as follows:

$$\text{Support}(A) = \frac{\sum \text{transactions containing } A}{\text{transaction}} * 100\%$$

*Nur Azis



The following is a calculation of one itemset formation from the product above.

$$\begin{aligned} &\text{Original Canon 810 Black Support Cartridge Value} \\ &S (\text{Original Canon 810 Black Cartridge}) \\ &= *100\% \frac{\sum \text{transaction contains Canon Original Cartridge 810 Black}}{\text{transaction}} \\ &= *100\% = 21.60\% \frac{81}{375} \end{aligned}$$

After searching for the support value for 1 itemset, we find the support value for 2 itemsets using the following formula:

$$\text{Support}(A,B) = *100\% \frac{\sum \text{transaction contains A and B}}{\sum \text{Transaction}}$$

The following is a calculation of one itemset formation from the product above.

$$\begin{aligned} &\text{Support Value for the combination of Original Canon 810 Black Cartridge with Canon 745} \\ &\text{Black (AB) Refill Cartridge} \\ &\text{Support}(AB) = *100\% \frac{\sum \text{transaction contains AB}}{\sum \text{Transaction}} \\ &\text{Support}(AB) = *100\% \frac{167}{375} \\ &\text{Support}(AB) = 44.53\% \end{aligned}$$

After searching for the support values for 2 itemsets, it is found, and judging from the percentage results that there are already more than 100%, the frequency pattern can be determined, then look for association rules that meet the minimum requirements for confidence by calculating the confidence of the associative rule $A \rightarrow B$ or the combination of 2 itemsets is minimum. of $\pm 75\%$. The Confidence value of rule $A \rightarrow B$ can be seen from one of the following calculations:

$$\begin{aligned} &\text{Original Canon 810 Black cartridge with HP Laserjet 85A (AE) toner refill,} \\ &\text{Support}(AE) = *100\% \frac{\sum \text{transaction contains AE}}{\sum \text{Transaction}} \\ &\text{Support}(AE) = \frac{293}{375} *100\% \\ &\text{Support}(AE) = 78.14\% \end{aligned}$$

CONCLUSION

Based on the results of the research that has been done, it can be concluded that the most sold products in one year are more than 50 combination itemsets. The implications of the research show that the Apriori Algorithm can help find the types of sales transaction patterns that often occur at CV. Delta Komputer Cijantung, From the patterns presented there are several combinations of items that reach a confident value of 100% so that it can be an alternative decision in analyzing stock items. Future research can analyze by adding itemsets so that the addition of itemsets will affect the minimum support and maximum support values to produce the best confident value.

REFERENCES

- Akbar, R., Silvana, M., Hersyah, M. H., & Jannah, M. (2020). Implementation of business intelligence for sales data management using interactive dashboard visualization in XYZ stores. *2020 International Conference on Information Technology Systems and Innovation (ICITSI)*, 242–249.
- Akmaludin, A., Rinawati, R., Arisawati, E., & Dewi, L. S. (2023). Decision Support for Selection of The Best Teachers Recommendations MCDM-AHP and ARAS Collaborative Methods. *Sinkron: Jurnal Dan Penelitian Teknik Informatika*, 8(4), 2036–2048.
- Angela, R., Pranta, A., & Ginting, R. I. (2022). Sistem Pendukung Keputusan dalam Menentukan

*Nur Azis



- Kelayakan Penerima Kredit Barang Elektronik Dengan Metode Oreste. *Jurnal Cyber Tech*, 4(2).
- Asana, I. M. D. P., Sudipa, I. G. I., Mayun, A. A. T. W., Meinarni, N. P. S., & Waas, D. V. (2022). Aplikasi Data Mining Asosiasi Barang Menggunakan Algoritma Apriori-TID. *INFORMAL: Informatics Journal*, 7(1), 38–45.
- Asward, B., & Brotosaputro, G. (2021). Analisis Stok dan Penjualan Kaftan Menggunakan Metode Backpropagation (Studi Kasus Lala New Fashion). *Respati*, 16(2), 71–82.
- Azis, N. (2023). Analisa dan Pancangan Aplikasi Jasa Transaksi Keuangan Pada PT. Tri Perkasa Garmino. *IKRA-ITH Informatika: Jurnal Komputer Dan Informatika*, 7(2), 66–71.
- Azis, N., & Handoko, B. A. (2021). Analisa dan Perancangan Aplikasi Pengadaan Barang di PT. Sintra Sinarindo Elektrik Berbasis Web. *Jurnal Information System*, 1(2), 38–42.
- Damanik, N. S., Erwansyah, K., & Halim, J. (2022). Penerapan Data Mining Untuk Menentukan Pola Kombinasi Produk yang Dapat Diberikan Diskon pada PT. Home Center Indonesia (INFORMA) Menggunakan Algoritma Apriori. *Jurnal Cyber Tech*, 2(12).
- Dewantara, R., & Giovanni, J. (2023). Analisis Peramalan Item Penjualan dalam Optimalisasi Stok Menggunakan Metode Least Square. *Jurnal Krisnadana*, 3(1), 59–66.
- Fujioka, K., & Shirahama, K. (2022). Generic Itemset Mining Based on Reinforcement Learning. *IEEE Access*, 10, 5824–5841.
- Halim, C. (2021). *Dinamika Etos Kerja Masyarakat Tionghoa Yogyakarta*. Sanata Dharma University Press.
- Herlawati, H., Handayanto, R. T., Rejeki, S., Priatna, W., Atika, P. D., Rofiah, S., Retnoningsih, E., Saputra, F. A., & Pradana, G. A. (2022). Penentuan Pola Frekuensi Jenis Perawatan Kecantikan Berbasis Web Menggunakan Algoritma Apriori (Studi Kasus: Peterson Salon Bekasi). *Journal of Students' Research in Computer Science*, 3(2), 241–254.
- Hernando, L. (2019). Algoritma Apriori Dan Fp-Growth Untuk Analisa Perbandingan Data Penjualan Laptop Berdasarkan Merk Yang Diminati Konsumen (Studi Kasus: Indocomputer Payakumbuh). *J-Clik: Jurnal Sistem Informasi Dan Manajemen Informatika*, 6(1), 1–17.
- Kaunan, F., Kelen, Y. P. K., & Nababan, D. (2023). Sistem Pendukung Keputusan Pemilihan Calon Kepala Desa Menggunakan Metode Analitic Hierarchy Process (AHP) Berbasis Web (Studi Kasus: Desa Oesena). *Jurnal Krisnadana*, 2(3), 375–387.
- Kurniana, T., Lestari, A., & Oktaviyani, E. D. (2023). Penerapan Algoritma Apriori untuk Mencari Pola Transaksi Penjualan Berbasis Web pada Cafe Sakuyan Side. *KONSTELASI: Konvergensi Teknologi Dan Sistem Informasi*, 3(1), 13–23.
- Kwintiana, B., Nengsih, T. A., Baradja, A., Harto, B., Sudipa, I. G. I., Handika, I. P. S., Adhicandra, I., & Gugat, R. M. D. (2023). *DATA SCIENCE FOR BUSINESS: Pengantar & Penerapan Berbagai Sektor*. PT. Sonpedia Publishing Indonesia.
- Panjaitan, S., Sulindawaty, Amin, M., Lindawati, S., Watrianthos, R., Sihotang, H. T., & Sinaga, B. (2019). Implementation of apriori algorithm for analysis of consumer purchase patterns. *Journal of Physics: Conference Series*, 1255(1), 12057.
- Prasetya, T., Yanti, J. E., Purnamasari, A. I., Dikananda, A. R., & Nurdiawan, O. (2022). Analisis Data Transaksi Terhadap Pola Pembelian Konsumen Menggunakan Metode Algoritma Apriori. *INFORMATICS FOR EDUCATORS AND PROFESSIONAL: Journal of Informatics*, 6(1), 43–52.
- Ridzuan, F., & Zainon, W. M. N. W. (2019). A review on data cleansing methods for big data. *Procedia Computer Science*, 161, 731–738.
- Rosmayati, I., Wahyuningsih, W., Harahap, E. F., & Hanifah, H. S. (2023). Implementasi Data Mining pada Penjualan Kopi Menggunakan Algoritma Apriori. *Jurnal Algoritma*, 20(1), 99–107.
- Saptadi, N. T. S., Chyan, P., & Leda, J. M. (2023). Analysis of Supermarket Product Purchase Transactions With the Association Data Mining Method. *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)*, 7(3), 618–627.
- Saputra, A. D., & Qoiriah, A. (2022). Penerapan Algoritma C4. 5 Untuk Mengatur Persediaan Stok Barang Berbasis Website. *Journal of Informatics and Computer Science (JINACS)*, 3(04), 481–493.
- Sika, S. N. R. (2021). Sistem Informasi Persediaan Stok Barang Berbasis Web Pada Toko Putra Gresik. *Jurnal Fasilkom*, 11(3), 157–164.

*Nur Azis



- Sileyew, K. J. (2019). *Research design and methodology*. IntechOpen Rijeka.
- Sriyanto, S., Fauzi, A., & Zonyfar, C. (2022). Data Mining Penerapan Algoritma Apriori untuk Analisis Data Transaksi Sistem Inventory (Studi Kasus PT ABC President Indonesia). *Scientific Student Journal for Information, Technology and Science*, 3(2), 143–150.
- Sudipa, I. G. I., Udayana, I. P. A. E. D., Rizal, A. A., Kharisma, P. I., Indriyani, T., Asana, I. M. D. P., Ariana, A. A. G. B., & Rachman, A. (2023). *METODE PENELITIAN BIDANG ILMU INFORMATIKA (Teori & Referensi Berbasis Studi Kasus)*. PT. Sonpedia Publishing Indonesia.
- Sudipa, I. G. I., Wardoyo, R., Hatta, H. R., Sagena, U., Gunawan, I. M. A. O., Zahro, H. Z., & Adhicandra, I. (2023). *MULTI CRITERIA DECISION MAKING: Teori & Penerapan Metode Pengambilan Keputusan dengan MCDM*. PT. Sonpedia Publishing Indonesia.
- Sudjiman, P. E., & Sudjiman, L. S. (2018). Analisis sistem informasi manajemen berbasis komputer dalam proses pengambilan keputusan. *TeIKa*, 8(2), 55–66.
- Thariq, A. (2023). Implementasi Market Basket Analysis Menggunakan Algoritma Apriori pada Data Penjualan Buku. *Jurnal Kolaboratif Sains*, 6(3), 154–163.
- Valeska, D. A., Umbara, F. R., & Sabrina, P. N. (2022). Diagnosa Gejala yang Muncul Bersamaan pada Penderita Tuberculosis Menggunakan Algoritma Apriori dengan Substitusi Metode Bayesian pada Nilai Confidence. *Jurnal Teknologi Informatika Dan Komputer*, 8(1), 318–332.