International Journal of Artificial Intelegence Research
Vol. 6, No. 2, December 2022
ISSN 2579-7298

# Association Rule Mining System in Analyzing The Use Pattern of Drugs by Using Apriori

Herlina Latipa Sari [a,1], Ila Yati Beti [b,2]

[a,b] Faculty of Computer Science, Informatics, Dehasen University Bengkulu, Bengkulu, Indonesia.
[1] herlinalatipasari@unived.ac.id; [2] ilayb@unived.ac.id;
* corresponding author

ARTICLE INFO

**Article history**
Received 12 June 2022
Revised 30 Sep 2022
Accepted 01 December 2022

**Keywords**
Association Rule
Mining
Apriori
Medicine

ABSTRACT

Data mining is a process to support decision making in finding information patterns in the data. In this study, Association Rule Mining will be implemented as one of the data mining techniques to analyze data and assist data of scientists in compiling raw data, formulating it and recognizing various patterns through a priori algorithms. The method used in this study is the Cross Industry Standard Process for Data Mining (CRISP-DM) Method by analyzing drug use patterns in health centers. The results of the study shows that by using the apriori algorithm, it found patterns and rules of widely used drugs that will provide recommendations in supporting decision making by health centers to submit drug procurement so that they can improve the quality of health services and minimize the risk of shortages or excess drug supplies and help health centers in optimizing drug inventory management. The results of the analysis using the apriori algorithm on the combination pattern of 2 itemsets produced 2 association rules for drug use, they are "If using Amoxicillin caplets 500 mg, then you will use paracetamol" with a confidence value of 80% and "If using Dexamethasone tablets 0.5 mg, then you will use Ascorbic Acid (Vit C) tablets 50 mg" with a confidence value of 100%.



## 1. Introduction

The implementation of data mining has become one of the strategies in processing and analyzing big data to reveal hidden patterns, trends, and relationships [1]. With data mining, it can produce accurate and relevant information to support effective and efficient decision making in various fields such as banking, health, e-commerce and others [2]. In the health sector, health centers act as First Level Health Facilities (FKTP) in health programs, which provide health services for patients to consult with doctors and get medicine. Health centers are responsible for preparing the medicines needed by patients and minimizing the risk of shortages and excess supplies of drugs. Based on this and interviews that have been conducted, it was found that drug procurement at health centers is based on the average end-of-month usage [3].

The purpose of this study is to determine and analyze drug use from patient prescription services by using Association Rule Mining to find associative rules between a combination of items [4]. The Apriori algorithm is a type of association rule used in daily drug use data. The use of the Apriori algorithm in this study will determine the confidence and support values that affect the amount of drug use shown in the combination pattern for drug itemset, the smaller the confidence

and support values set, the more drug use is shown and the greater the confidence and support values set, the less use is shown.

It is expected that the Association Rule Mining system in analyzing drug usage patterns using apriori can provide a combination pattern of drug items that are frequently used in daily patient services, in addition to providing recommendations in drug procurement from the results of drug items that are widely used so as to optimize drug inventory management.

## 2. Method

### 2.1 Research Methodology

The research method used in this research is the Cross Industry Standard Process for Data Mining (CRISP-DM) method, a data mining methodology developed by a consortium of companies founded by the European Commission in 1996 and has been established as a standard process in data mining [5]. CRISP-DM has six phases as seen in Figure 1.

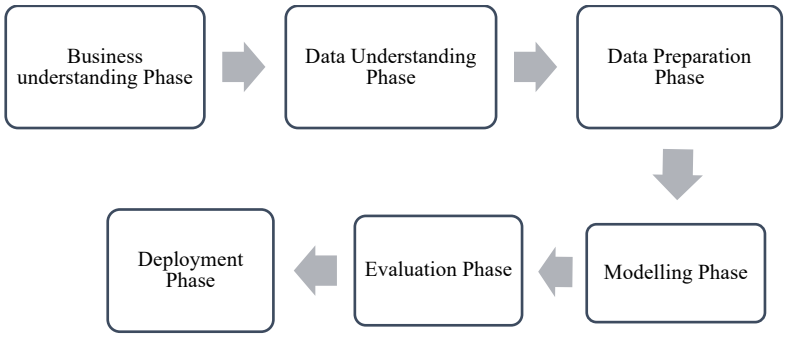


**Fig 1.** Research Phases

a. *Business Understanding Phase*: In this phase, the researcher understands the research objectives and the needs of the research objectives, the research objective is to obtain patterns or rules in the use of drugs in health centers from patient drug prescriptions so that the combination of drugs that are often dispensed from health center pharmacies is known.

b. *Data Understanding Phase*: in this phase the researcher creates data targets and focuses on variables or data samples to be taken then the data is cleaned target data with the aim of obtaining consistent data. The data in the database is not all used, therefore only data that is suitable for analysis will be taken.

c. *Data Preparation Phase*: After the data understanding phase is carried out, then in this data processing phase, the final dataset will be created which will be applied to the modeling tool. From the initial raw data and then the data mining process will be carried out. Data is changed or combined into a format suitable for processing in data mining. Data is changed or combined into a format suitable for processing in data mining. The data is by using the name of the drug used in the prescription of the Sukamerindu Health Center patient, Bengkulu City.

d. *Modelling Phase*: In this modeling phase, researchers select associative rule techniques using the apriori algorithm. This phase is the main process because when the method is applied, hidden knowledge is obtained from the data.

e. *Evaluation Phase*: This evaluation phase will evaluate and examine to ensure that the modeling stage used meets the objectives of the study, namely the phase to draw patterns into the knowledge found. This phase is an evaluation to assess whether the model that has been created is in accordance with the objectives of the study. This phase is to determine the accuracy of the model using *a confusion matrix*.

f. *Deployment Phase*: This phase is the final stage in creating a data mining activity report, a final report containing the knowledge gained or pattern recognition in data in the data mining process.

### 2.2 *Knowledge Discovery in Database* (KDD)

Data mining is also known as *Knowledge Discovery in Database* (KDD) with the process depicted in Picture 2.

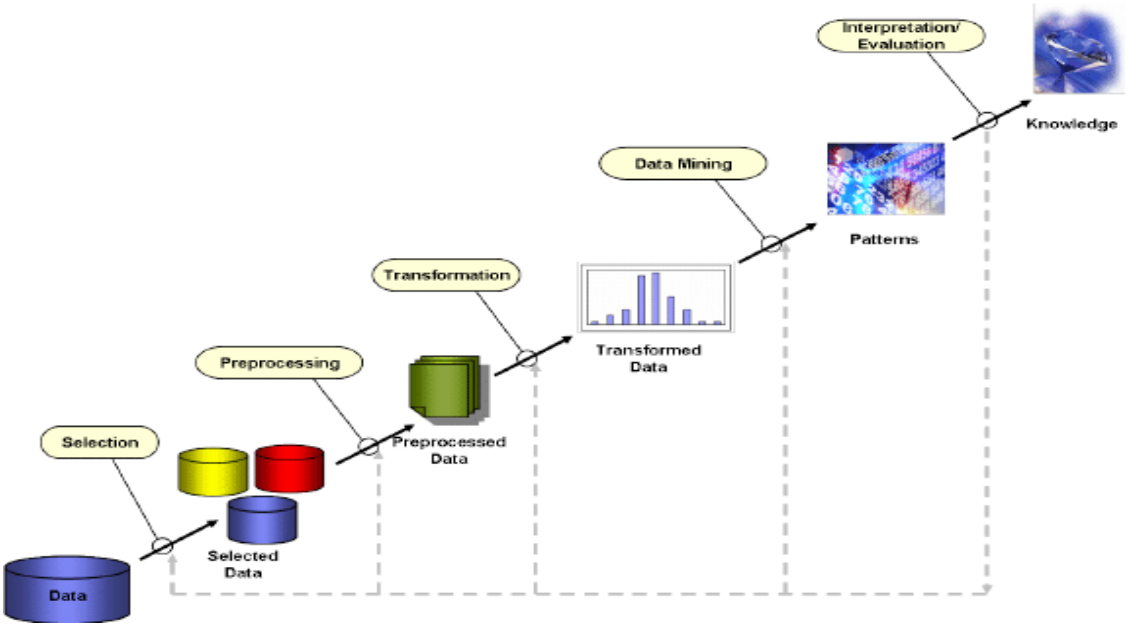


**Fig 2.** Process of *Knowledge Discovery in Database* (KDD)

From Picture 2 The iterative process chain of data mining in KDD is as follows [6,7]:

a. Data cleaning, eliminate noise and inconsistent data.
b. Data integration, eliminate noise and inconsistent data.
c. Data selection, retrieve data relevant to the analysis task from the database.
d. Data transformation, transforming or combining data into a form suitable for mining through summary or aggregation operations.
e. Data mining, an essential process to extract patterns from data with intelligent methods.
f. Pattern evaluation, identify interesting patterns and represent knowledge based on interestingness measures.
g. Knowledge presentation, presentation of excavated knowledge to users using visualization and knowledge representation techniques.

### 2.3 *Association Rule Mining of Algoritma Apriori*

*Association Rule Mining* is a data mining technique to find associative rules between a combination of items [8]. Association analysis is also known as one of the data mining techniques that is the basis of various other data mining techniques, especially one stage of association analysis called *high frequency pattern mining* attracts the attention of many researchers to produce efficient algorithms in finding rules for a combination of items [9,10]. The Apriori algorithm is an algorithm for running frequent itemset searches with association rules. The apriori algorithm is a type of association rule in data mining, rules that state associations between several attributes are often called affinity analysis or market basket analysis [11].

The apriori algorithm uses a level-wise search approach, where k-itemset is used to obtain (k+1)-itemset [12]. This process is carried out until no more combinations can be formed [13]. The importance of an association can be known by two benchmarks, namely support and confidence. Support (supporting value) is the percentage of the combination of items in the database while

confidence (certainty value) is the strength of a relationship between items in the association rules [14]. Based on the equations regarding support and confidence to obtain a value on an itemset, it is obtained using the formula [15]:

Persamaan (1) digunakan untuk menghitung nilai *support* untuk sebuah item.

$$Support\ (A) = \frac{number\ of\ transaction\ with\ A}{total\ of\ transaction} x100\% \quad \text{.....(1)}$$

Meanwhile, to calculate the support value for 2 (two) items, the formula is used: $support\ (A,B) = P(A \cap B)$

$$Support = \frac{\sum transactin\ contains\ A\&B}{total\ of\ transaction} x100\% \quad \text{.....(2)}$$

Equation (3) is used to calculate the confidence value which can be obtained using the formula:

$$Confidence = \frac{\sum transaction\ contains\ A\ and\ B}{\sum transaction\ contains\ A} x100\% \quad \text{.....(3)}$$

To determine the association rules to be selected, they must be sorted based on support x confidence, where the rules are taken as many as n rules that have the largest results. The stages of how the apriori algorithm works are as follows [16,17]:

a. Determining minimum support using the formula equation (1), the process of generating candidate *item sets* (*candidate generation*)
b. Data scanning process to calculate *support (support counting*)
c. Determine *candidate rules* that contain rule possibilities that have a support value > minimum support.
d. Determine the value of the combined antecedent support by comparing the rule support value with the value generated for the antecedent support.

## 3. Results and Discussion

Data analysis was carried out after the data was collected and in accordance with the association rule technique using the apriori algorithm with several iterations or stages, the data used were patient drug prescription data in 2023 at the Sukamerindu Health Center Pharmacy in Bengkulu City. The data sample used was 20 patient drug prescriptions which can be seen in Table 1.

**Table 1.** Patient Drug Prescription Data Sample

| Recipe | Drugs used in prescriptions |
|---|---|
| 1 | Amoxicillin Kaplet 500 mg, Parasetamol tablet 100 mg, Dexamethasone tablet 0,5 mg, Asam askorbat (Vit C) tablet 50 mg |
| 2 | Kotrimoksazol suspense, Paracetamol syrup 120 mg, Obat Batuk Hitam (OBH), Vitamin B Complex tab |
| 3 | Tetrasiklin HCL kapsul 500 mg, Antalgin tablet 500 mg, Vitamin B Complex tab |
| 4 | Amoxicillin kaplet 500 mg, Vitamin B Complex tab, Antalgin tablet 500 mg |
| 5 | Antalgin tablet 500 mg, Vitamin B Complex tab, Tetrasiklin HCL kapsul 500 mg |
| 6 | Paracetamol syrup 120 mg, Cetrizin tablet 10 mg, Gliseril guaikolat tablet 100 mg, Vitamin B Compleks tab |
| 7 | Paracetamol syrup 120 mg, Obat Batuk Hitam (OBH), Cetirizine tablet 10 mg, Amoxicillin kaplet 500 mg |
| 8 | Parasetamol syrup 120 mg, Benedict, Tetrasiklin HCL kapsul 500 mg |
| 9 | Cetirizine tablet 10 mg, Dexamethasone tablet 0,5 mg, Asam askorbat (Vit C) tablet 50 mg |

| | |
|---|---|
| 10 | Aminofilin tablet 200 mg, antalgin tablet 500 mg, Thiamin HCL (Vit B1) tablet 50 mg, Metformin tablet 500 mg |
| 11 | Obat Batuk Hitam (OBH), Cetirizine tablet 10 mg, Asam askorbat (Vit C) tablet 50 mg |
| 12 | Aminofilin tablet 200 mg, Metformin tablet 500 mg, Cetrizin tablet 10 mg, Vitamin B Complex tab |
| 13 | Amoxicillin kaplet 500 mg, Asam askorbat (Vit C) tablet 50 mg, Parasetamol syrup 120 mg |
| 14 | Thiamin HCL (Vit B) tablet 50 mg, Antalgin tablet 500 mg, Cetirizine tablet 10 mg |
| 15 | Paracetamol syrup 120 mg, Gliseril guaikolat tablet 100 mg, Vitamin B Complex tab |
| 16 | Antalgin tablet 500 mg, Vitamin B Complex tab |
| 17 | Aminofilin tablet 200 mg, Antalgin tablet 500 mg, Cetrizin tablet 10 mg, Vitamin B Complex tab |
| 18 | Amoxicillin kaplet 500 mg, Parasetamol syrup 120 mg |
| 19 | Paracetamol syrup 120 mg, Domperidon sirup, Thiamin HCL (Vit B) tablet 50 mg, Cetirizine tablet 10 mg |
| 20 | Ambroxol tablet 30 mg, Cetrizin tablet 10 mg, Asam askorbat (Vit C) tablet 50 mg, Vitamin B Complex tab |

From Table 1, there are 20 patient drug prescriptions, each of which contains several drug items used or issued by the Community Health Center Pharmacy. For the drug items issued, the support value will be calculated for 1 item set using equation (1), obtaining the quantity (amount) of 17 types of drugs in Table 1.

**Table 2.** Drug Data for itemset 1

| No | Drug Name | Total | Support |
|---|---|---|---|
| 1 | Amoxicillin caplet 500 mg | 5 | 0,25 |
| 2 | Paracetamol syrup 120 mg | 9 | 0,45 |
| 3 | Dexamethasone tablet 0,5 mg | 2 | 0,1 |
| 4 | Asam askorbat (Vit C) tablet 50 mg | 5 | 0,25 |
| 5 | Kotrimoksazol suspense | 1 | 0,05 |
| 6 | Obat Batuk Hitam (OBH) | 3 | 0,15 |
| 7 | Vitamin B Complex tab | 10 | 0,5 |
| 8 | Antalgin tablet 500 mg | 7 | 0,35 |
| 9 | Tetrasiklin HCL kapsul 500 mg | 3 | 0,15 |
| 10 | Gliseril guaikolat tablet 100 mg | 2 | 0,1 |
| 11 | Benedict | 1 | 0,05 |
| 12 | Cetrizin tablet 10 mg | 9 | 0,45 |
| 13 | Aminofilin tablet 200 mg | 3 | 0,15 |
| 14 | Thiamin HCL (Vit B1) tablet 50 mg, | 3 | 0,15 |
| 15 | Metformin tablet 500 mg | 2 | 0,1 |
| 16 | Domperidon sirup | 1 | 0,05 |
| 17 | Ambroxol tablet 30 mg | 1 | 0,05 |

The next stage is to calculate the support value on 2 itemsets obtained by using the equation formula (2) to find support (A,B) = P(A∩B), from the data the combination of 2 itemsets, they are 133 pairs as total, then select items that meet the value Φ >= 2 with minimum support ≥ 10% with the results in Table 3.

**Table 3.** Support Value on Combination Pattern of 2 Drug Items

| No | 2 Item Combination Pattern | Support | |
|---|---|---|---|
| 1 | Amoxicillin kaplet 500 mg, Parasetamol | 4/20 | 20% |
| 2 | Amoxicillin kaplet 500 mg, Asam askorbat (Vit C) tablet 50 mg | 2/20 | 10% |
| 3 | Parasetamol, Asam askorbat (Vit C) tablet 50 mg | 2/20 | 10% |
| 4 | Parasetamol, Obat Batuk Hitam (OBH) | 2/20 | 10% |
| 5 | Parasetamol, Vitamin B Complex tab | 3/20 | 15% |
| 6 | Parasetamol, Antalgin tablet 500 mg | 5/20 | 25% |
| 7 | Parasetamol, Cetrizin tablet 10 mg | 3/20 | 15% |
| 8 | Dexamethasone tablet 0,5 mg, Asam askorbat (Vit C) tablet 50 mg | 2/20 | 10% |
| 9 | Asam askorbat (Vit C) tablet 50 mg, Cetirizine tablet 10 mg | 3/20 | 15% |
| 10 | Obat Batuk Hitam (OBH), Cetirizine tablet 10 mg | 2/20 | 10% |
| 11 | Vitamin B Complex tab, Tetrasiklin HCL kapsul 500 mg | 2/20 | 10% |
| 12 | Vitamin B Complex tab, Antalgin tablet 500 mg | 5/20 | 25% |
| 13 | Vitamin B Complex tab, Cetrizin tablet 10 mg | 4/20 | 20% |
| 14 | Vitamin B Complex tab, Aminophilin tablet 200 mg | 2/20 | 10% |
| 15 | Tetrasiklin HCL kapsul 500 mg, Antalgin tablet 500 mg | 2/20 | 10% |
| 16 | Antalgin tablet 500 mg, Cetirizine tablet 10 mg | 2/20 | 10% |
| 17 | Antalgin tablet 500 mg, Aminofilin tablet 200 mg | 2/20 | 10% |
| 18 | Antalgin tablet 500 mg, Thiamin HCL (Vit B1) tablet 50 mg | 2/20 | 10% |
| 19 | Cetrizin tablet 10 mg, Aminophilin tablet 200 | 2/20 | 10% |
| 20 | Cetrizin tablet 10 mg, Thiamin HCL (Vit B1) tablet 50 mg | 2/20 | 10% |
| 21 | Aminofilin tablet 200, Metformin tablet 500 mg | 2/20 | 10% |

To find the Association Rule Mining rules by calculating the confidence value, it can be obtained by using the formula in equation (3), for all ∑ drugs used, there are 2 combination patterns of 2 items after setting the minimum combination value = 75% in table 4.

**Table 4.** Confidence value in the combination pattern of 2 itemsets

| No | 2 Items Combination Pattern | Confidence | |
|---|---|---|---|
| 1 | Amoxicillin kaplet 500 mg →Parasetamol | 4/5 | 80% |
| 2 | Amoxicillin kaplet 500 mg →Asam askorbat (Vit C) tablet 50 mg | 2/5 | 40% |
| 3 | Parasetamol →Asam askorbat (Vit C) tablet 50 mg | 2/9 | 22,2% |
| 4 | Parasetamol → Obat Batuk Hitam (OBH) | 2/9 | 22,2% |
| 5 | Paracetamol → Vitamin B Compleks tab | 3/9 | 33,3% |
| 6 | Paracetamol → Antalgin tablet 500 mg | 5/9 | 55,6% |
| 7 | Parasetamol → Cetrizin tablet 10 mg | 3/9 | 33,3% |
| 8 | Dexamethasone tablet 0,5 mg → Asam askorbat (Vit C) tablet 50 mg | 2/2 | 100% |
| 9 | Asam askorbat (Vit C) tablet 50 mg → Cetirizine tablet 10 mg | 3/5 | 60% |
| 10 | Obat Batuk Hitam (OBH) → Cetirizine tablet 10 mg | 2/3 | 66,7% |
| 11 | Vitamin B Complex tab → Tetrasiklin HCL kapsul 500 mg | 2/10 | 20% |
| 12 | Vitamin B Complex tab→ Antalgin tablet 500 mg | 5/10 | 50% |
| 13 | Vitamin B Complex tab → Cetrizine tablet 10 mg | 4/10 | 40% |
| 14 | Vitamin B Complex tab → Aminofilin tablet 200 mg | 2/10 | 20% |
| 15 | Tetrasiklin HCL kapsul 500 mg → Antalgin tablet 500 mg | 2/3 | 66,7% |
| 16 | Antalgin tablet 500 mg → Cetirizine tablet 10 mg | 2/7 | 28,6% |

| | | | |
|---|---|---|---|
| 17 | Antalgin tablet 500 mg → Aminophylline tablet 200 mg | 2/7 | 28,6% |
| 18 | Antalgin tablet 500 mg → Thiamin HCL (Vit B1) tablet 50 mg | 2/7 | 28,6% |
| 19 | Cetirizine tablet 10 mg → Aminophylline tablet 200 | 2/9 | 22,2% |
| 20 | Cetirizine tablet 10 mg → Thiamin HCL (Vit B1) tablet 50 mg | 2/9 | 22,2% |
| 21 | Aminofilin tablet 200 → Metformin tablet 500 mg | 2/3 | 66,7% |

From the stages that have been carried out, the items that meet the support x confidence with a minimum value = 75% in Table 5 contain 2 Association Rules that are formed, rhey are the first rule "If using Amoxicillin caplets 500 mg then will use paracetamol" with a confidence value of 80%, the second rule "If using Dexamethasone tablets 0.5 mg then will use Ascorbic Acid (Vitamin C) tablets 50 mg" with a confidence value of 100%.

**Tabel 5.** Association Rules that are formed

| No | 2 Item Combination Pattern | Support | Confidence |
|---|---|---|---|
| 1 | Amoxicillin kaplet 500 mg →Parasetamol | 20% | 80% |
| 2 | Dexamethasone tablet 0,5 mg → Asam askorbat (Vit C) tablet 50 mg | 10% | 100% |

## 4. Conclusion

According to the implementation of Association Rule Mining apriori algorithm to the problem of analyzing drug use patterns at Sukamerindu Health Center, Bengkulu City, with consideration of the increasing number or size of drug use data contained in patient prescriptions with the increasing number of patient services that occur. From the drug use data, rules will be produced that become useful information in recommending drugs that will be added to the monthly stock. The results of the analysis using the apriori algorithm on the combination pattern of 2 itemsets produce 2 drug use association rules, these rules are useful for making drug supply procurement effective and optimal, namely "If using Amoxicillin caplets 500 mg, then paracetamol will be used" with a confidence value of 80% and "If using Dexamethasone tablets 0.5 mg, then Ascorbic Acid (Vit C) tablets 50 mg" with a confidence value of 100%.

## References

[1] I. Priyadi, J. Santoni, and J. Naam, "Data Mining Predictive Modeling for Prediction of Gold Prices Based on Dollar Exchange Rates, Bi Rates and World Crude Oil Prices," *Indones. Artif. Intell. Data Min,* vol.2, no.2, pp. 93-100, 2019.

[2] R. Agrawal, et.al, "Mining Association Rules between Sets of Item in Large Database," *Proceeding of the 1993 ACM SIGMOD Conference Washington DC, USA,* 1993.

[3] D.P.A.Andira, and A. Ikhwan, "Implementasi Data Mining pada Sistem Persediaan Obat di Puskesmas Sei Berombang dengan Metode Algoritma Apriori," *Indonesian Journal of Computer Science (IJCS),* vol 11, no.3, pp.958-978, 2022.

[4] F. Jiang, and Y. Lu, " Software Testing Model Selection Research Based on Yin-Yang Testing Teory'" *In 2012 International Conference on Computer Science and Information Proceeding (CISP),* pp.590-594. 2012.

[5] A. Febriani, T. Rahmawati, Trimadya, and E. Sabna, "Implementation of Data Mining to Predict the Feasibility of Blood Donors Using C4.5 Algorithm," *Indones. Artif. Intell. Data Min,* vol.1, no.1, pp. 41-46, 2018.

[6] F. Handayanna and E. Irfiani, "Analisa Pola Penjualan Makanan dengan Penerapan Algoritma Apriori," *Techno. Com,* Vol.20, no.4, pp.478-488, 2021.

[7] R. Takdirillah, "Penerapan Data Mining Menggunakan Algoritma Apriori Terhadap Data Transaksi Sebagai Pendukung Informasi Strategi Penjualan," *Edumatic J. Pendidik. Informatika,* vol.4, no.1, pp.37-46, 2020, doi: 10.29408/edumatic.v41.2081.
[8] Z. Munawar, Y. Herdiana, and N.I. Putri, "Sistem Rekomendasi Hibrid Menggunakan Algoritma Apriori Mining Asosiasi," *TEMATIK,* vol.8, no.1, pp. 84-95, 2021.
[9] K. Erwansyah, B. Andika, and R. Gunawan, "Implementasi Data Mining Menggunakan Asosiasi Dengan Algortma Apriori Untuk Mendapatkan Pola Rekomendasi Belanja Produk Pada Toko Avis Mobile," *J. Teknol. Sist. Infdan Sist. Komput. TGD,* vol.4, no.1, pp. 148-161, 2021.
[10] H. Maulidiya and A. Jananto, "Asosiasi Data Mining Menggunakan Algoritma Apriori Dan Fpgrowth Sebagai Dasar Pertimbangan Penentuan Paker Sembako, 2020.
[11] F. Ardhy, O. J. F. Wassalam, T. H. Andika, and S. A. Salimu, "Penerapan Data Mining Menggunakan Algoritma Apriori Dalam Memprediksi Penjualan Produk," *Jurnal Informasi dan Komputer,* vol.10, no.2, 2022.
[12] M. S. Rindawati and H. Andriani, "Analisis Pengendalian Persediaan Obat Menggunakan Metode ABC, Safety Stock, EOQ, dan ROP di Instalasi Farmasi Rumah Sakit Pemerintah di Jakarta," *Syntax Literate: Jurnal Ilmiah Indonesia*, vol.7, no.10, 2022.
[13] U. Ependi and A. Putra, "Solusi Predikasi Persediaan Barang dengan Menggunakan Algoritma Apriori (studi kasus : regional part depo auto 2000 palembang)," *JEPIN (jurnal Edukasi dan Penelitian Informatika,* vol.5, no.2, 2019.
[14] L. Tommy, C. Kirana, and V. Lindawati, "Recommender System Dengan Kombinasi Apriori dan Content-Based Filtering pada Aplikasi Pemesanan Product," *Jurnal TEKNOINFO*, vol.13, no.2, 2019.
[15] K. S. Talwar, A. Oraganti, N. Mahajan, and P. Narsale, "Recommendation System Apriori Algorithm," *IJSRD- International Journal for Scientific Research & Development,* vol.3, issue.01, 2015.
[16] F.A.Saputra and A. Iskandar, "Data Mining Penerapan Asosiasi Apriori dalam Penentuan Pola Penjualan," *Journal of Computer System and Informatics (JoSYC),* vol.4 no.4, pp.778-788, 2023.
[17] A. Novianti and E. Elisa, "Penentuan Aturan Asosiasi Pola Pembelian pada Minimarket dengan Algoritma Apriori", *Build.informatic, Technol, Sci,* vol 2, no.1, pp 64-70, 2020.