Journal of Artificial Intelligence and Engineering Applications

Website: https://ioinformatic.org/

15th October 2025. Vol. 5. No. 1; e-ISSN: 2808-4519

Application of Data Mining Using Apriori to Find Patterns of Asthma in Medical Record Data at the Health Center (Case Study: Datar City Health Center)

Halimatussadiah[1]*, Relita Buaton[2], Siswan Syahputra[3]

[1,2,3]STMIK Kaputama

htussadiah338@gmail.com[1*], bbcbuaton@gmail.com[2], siswansyahputra90@gmail.com[3]

Abstract

Medical records are a very important source of information in the world of health. Medical records document a patient's medical history, diagnosis, treatment, and care patterns at health facilities. However, with the large amount of data that continues to grow every day, it is often difficult for medical personnel and health facility managers to manually analyze and find useful patterns. Community health centers, as primary healthcare facilities, play an important role in addressing public health issues. Community health centers often face limitations in effectively processing available data. Therefore, methods are needed to help uncover hidden information from medical record data. One approach that can be used to analyze big data is data mining. Data mining allows users to find patterns, trends, or certain relationships that were previously unseen. In medical records, the application of data mining techniques can help identify disease patterns, relationships between diseases, and risk factors that contribute to certain diseases by using the apriori method to obtain better health service planning. From testing using the RapidMiner application, this study identified complaints, medical history, and causal factors. The results showed that there were 5 association rules formed with the highest Best rule value of 14% support and 62% confidence. The rule was "If the causal factor is genetic, the complaint is dizziness, then the medical history includes a history of asthma since childhood."

Keywords: *Apriori Algorithm, RapidMiner, Medical Records*

1. Introduction

Medical record data is one of the most important sources of information in the world of health. Medical records record the history of the disease, diagnosis, treatment, and patient care patterns in health facilities. However, in the large and ever-growing amount of data, it is often difficult for medical personnel and healthcare facility managers to analyze and find useful patterns manually.

Puskesmas as a first-level health facility has an important role in dealing with public health problems. The problem that often occurs in health centers is the large amount of patient data from time to time and the more data that is collected causes high complexity because of the many variables involved. This makes it difficult to identify important patterns from medical record data and is often a problem in health centers. [1]

However, often Puskesmas face limitations in processing the available data effectively. Therefore, a method is needed that can help unearth hidden information from the medical record data. One approach that can be used to analyze big data is data mining. Data mining allows users to discover specific patterns, trends, or relationships that were previously unseen. In medical record data, the application of data mining techniques can help identify disease patterns, relationships between diseases, and risk factors that contribute to certain diseases by using a priori methods to obtain better health service planning.

2. Method

From the research conducted by [2] it can be concluded that If T1 and P3 then BT1 with a support value = 20% and Confidence 100% and a value of S*C = 20% if the soil type T1 is (Alluvial Soil) then the customer will buy P3 fertilizer (Organic Fertilizer Specific Formula Palm Oil) with the fertilizer brand that many customers buy is BT1 (BT_Kelapa Palm Oil). with a supporting value of 20%, a certainty value of 100%.
From the research conducted by [3], 8 products were produced that were intertwined with a *support* result of 0.333 and a *confidence* value of 71.4%. The results of this research can be used as a sales strategy to increase daily sales.

2.1. Data Mining

Data Mining is the process of extracting useful information and patterns from a very large amount of data. The data mining process consists of data collection, data extraction, data analysis, and data statistics. It is also commonly known as knowledge discovery, knowledge extraction, data/pattern analysis, information harvesting, and others. [4]

2.2. Algoritma apriori

A priori algorithm is a basic algorithm proposed by Agrawal & Skrikant in 1994 to determine Frequent itemsets for Boolean association rules. A priori algorithms belong to the type of association rules on data mining. Rules that state associations between several attributes are often called affinity analysis or market basket analysis. Association rule data mining is a data mining technique to find the rules of a combination of items. One of the stages of association analysis that attracts the attention of many researchers to produce efficient algorithms is frequent pattern mining analysis. The importance of an association can be determined by two benchmarks, namely: support and confidence, support is the percentage of the combination of these items in the database, while confidence is the strength of the relationship between items in the association rules. [5]

The formation of association rules that meet the minimum requirements for confidance by calculating the confidance of the A→B associative rule, where support is supporting data and confidance is confidence.
The confidance value of rule A→B is obtained from the following formula: [5]

$$Support\ (A) = \frac{\sum Transaksimengandung A dan B}{\sum Jumlah seluruh transaksi} \times 100\% \quad (1)$$

$$Confidance\ (A) = \frac{\sum Transaksimengandung A dan B}{\sum Transaksimengandung A} \times 100\% \quad (2)$$

Searches for the minimum qualifying combination of items from the *support* value in the database. The *support* value of an item is obtained using the following formula: [5]

$$Support\ (A) - \frac{Jumlah\ transaksi\ mengandung\ A}{Total\ Transaksi} \quad (3)$$

The support value of the 2 items is obtained using the formula:

$$Support\ (A,B) = p\ (A \cap B)$$

$$Support\ (A,B) - \frac{\sum Jumlah\ transaksi\ mengandung\ A\ dan\ B}{\sum transaksi} \quad (4)$$

After all high-frequency patterns are found, then an association rule that meets the minimum requirements for confidence is searched by calculating the confidence of the associative rule A B. The confidence value of rule A B is obtained by the following formula: ∪∪[5]

$$Confidence - P\ (B|A) - \frac{\sum J\ transaksi\ mengandung\ A\ dan\ B}{\sum transaksi} \quad (5)$$

2.3. Disease Definition

Illness is an abnormal condition in which a person's body or mind experiences discomfort or disorders in a person. Disease is the failure of an organism's adaptation mechanism to react appropriately to stimuli or pressures so that disturbances arise in the function or structure of the organization of body systems.[6]

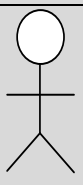
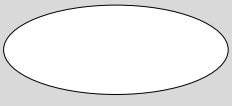
2.4. Medical Records

Medical records are facts related to the patient's condition, past history of illness and treatment and are currently written by the health profession that provides services to the patient. Medical record documentation is a file containing identity, anamnesis, physical determination, laboratory, diagnosis and medical actions against a patient that are recorded either in writing or electronically. Medical records have a very important role, because they store information about the patient's condition, so confidentiality must be guaranteed. With the implementation of good medical records, it will certainly support the implementation of efforts to improve services to the community.[7]

2.5. Use Case Diagram

Use Case Diargam is a modeling to carry out (behavior) information systems to be created. Use cases are used to find out what functions exist in an information system and who has the right to use those functions. Here are the symbols on the use case diagram. [8]

Table 1: Use Case Diagram

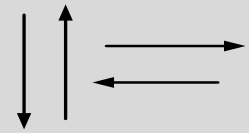
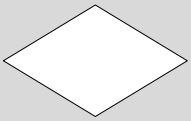
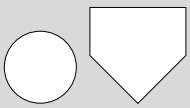
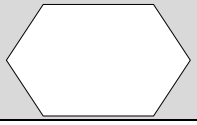
No.	Notasi	Information	Simbol
1	Actor	System users or those who interact directly with the system	
2	Use Case	An elliptical circle with the name of the use case written in the center of the circle	
3	Association	The line that functions connects the actor with the use case.	

No.	Notasi	Information	Simbol
4	Relationship	As a link between actors and use cases, as well as other use cases	
5	Include Relationship	Allows a use case to use functionality provided by another use case	<<include>>
6	Extend Relationship	Allows use cases to have the possibility to extend the functionality provided by other use cases.	<<extend>>

2.6. Flowchart

A flowchart is a diagram of the algorithms in a program that states the direction of the program's flow. It is used as a communication tool, process analysis and for documentation. Flowchart is a visual form of pseudocode that is easier and simpler. A program flowchart is a chart that describes the logical sequence of the program's processes in problem solving. The flowchart symbols can be seen in table below; [9]

Table 2: Flowchart

Information	Information
	Input/output; used to represent data I/O
	Flow line; Indicates the current of the process
	Process; used to represent a process
	Decision; used for a selection of conditions in the program
	Liaison; Shows links to the same page or another page
	Defined processes; indicates an operation whose details are shown elsewhere
	Terminal; indicates the beginning and end of a process
	Preparation; is used to give the initial value of a quantity

2.7. Rapidminer

Rapid Miner is a data mining tool developed using Java as a programming language. Java is considered one of the languages that has various advantages, such as simplicity, security, strength, impact, high-level object-oriented capabilities and many other advantages. Rapid miner was developed for general data mining tasks. Most of the previous versions are open source (lower than 5), but the sixth version is adopted by several licensing options (starter, personal, professional, enterprise). The tool has many properties and advantages such as a user-friendly interface and good community support. [10]

Fig. 1: Rapidminer

3. Results and Discussion

The research method carried out is for something systematically using scientific methods and applicable sources. In the process of this research, it is shown to provide more meaningful results for the agency in finding disease patterns in the medical record data at the health center. The results of the conceptualization will be made into a research method using the literature study pattern as shown in the image below.

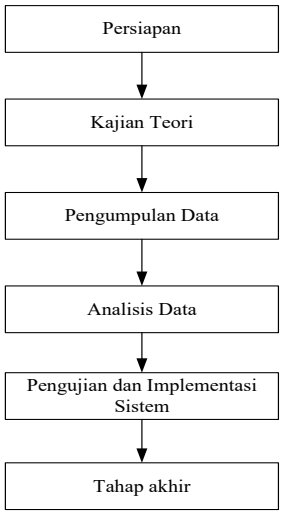


Fig. 2: Research Workflow

3.1. Research Supporting Data

The following is the data obtained from personnel transaction data as shown in the table below.

Table 3: Research Supporting Data

No	Patient Name	Complaints	Disease History	Causal Factors
1	Syabira Syahgara	Batuk	Tidak ada riwayat penyakit serius	Cuaca Dingin, Debu
2	Junaidi	Demam	Riwayat Asma sejak kecil	Polusi udara, infeksi virus
3	Fikri Pratama	Sesak napas	Tidak ada riwayat penyakit serius	Cuaca Dingin, Debu
4	Putri Arina	Batuk	Riwayat Asma sejak kecil	Genetik
5	Dias Irwansya	Mengi	Tidak ada riwayat penyakit serius	Cuaca Dingin, Debu
6	Nindi Adelia	Sesak napas	Riwayat Asma sejak kecil	Genetik
7	Erikha Novita	Mengi	Riwayat Asma sejak kecil	Genetik
8	Karen Mitchell	Nyeri di dada	Riwayat Asma sejak kecil	Genetik
9	Allison Brown	Mual	Riwayat Asma sejak kecil	Genetik
10	Sucita Manda Sari	Demam	Riwayat Asma sejak kecil	Genetik
11	Zuan Kelim Hutapea	Nyeri di dada	Tidak ada riwayat penyakit serius	Cuaca Dingin, Debu
12	Afatih	Mual	Tidak ada riwayat penyakit serius	Cuaca Dingin, Debu
13	Muhammad Gibran Alfatih	Demam	Tidak ada riwayat penyakit serius	Cuaca Dingin, Debu
14	Kanaya Ramadani	Sesak napas	Riwayat ISK sebelumnya	Alergen, cuaca dingin

No	Patient Name	Complaints	Disease History	Causal Factors
15	Veronica Dodson	Mengi	Riwayat ISK sebelumnya	Alergen, cuaca dingin
16	Vincent Vazquez	Nyeri di dada	Riwayat ISK sebelumnya	Alergen, cuaca dingin
17	Ghafar Atha Al Rizki	Mual	Riwayat ISK sebelumnya	Alergen, cuaca dingin
18	Ferdi Alamsyah	Batuk	Riwayat Asma sejak kecil	Polusi udara, infeksi virus
19	Pindira Al Zahra	Demam	Riwayat ISK sebelumnya	Alergen, cuaca dingin
20	Aprik	Sesak napas	Riwayat ISK sebelumnya	Genetik

3.2. Implementation Methode

From the research conducted to apply the association rule method to analyze disease patterns at the Datar City Health Center using a sample of patient medical record data that will be used as a research support, it can be seen in table 3. The data is a transactional database that will be represented as shown in the table below.

Table 4: Complaint Data

Code	Complaints
K1	Sesak Nafas
K2	Batuk
K3	Mengi (Suara seperti siulan saat bernapas)
K4	Nyeri di dada
K5	Demam
K6	Mual
K7	Nyeri perut
K8	Pusing
K9	Luka ringan
K10	Diare

Table 5: Disease History

Code	Disease History
R1	Tidak ada riwayat penyakit serius
R2	Riwayat Asma sejak kecil
R3	Riwayat ISK sebelumnya

Table 6: Causal Factors

Code	Causal Factors
F1	Tidak diketahui
F2	Kebersihan kurang, kurang minum
F3	Polusi udara, infeksi virus
F4	Genetik
F5	Cuaca Dingin, Debu
F6	Alergi, Cuaca Dingin

Table 7: Representasi Data

No	Complaints										History			Causal Factors					
	K1	K2	K3	K4	K5	K6	K7	K8	K9	K10	R1	R2	R3	F1	F2	F3	F4	F5	F6
1	0	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1	0
2	0	0	0	0	1	0	0	0	0	0	0	1	0	0	0	1	0	0	0
3	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1	0
4	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0
5	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1	0
6	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0
7	0	0	1	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0
8	0	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0
9	0	0	0	0	0	1	0	0	0	0	0	1	0	0	0	0	1	0	0
10	0	0	0	0	1	0	0	0	0	0	0	1	0	0	0	0	1	0	0
11	0	0	0	1	0	0	0	0	0	0	1	0	0	0	0	0	0	1	0

No	Complaints										History			Causal Factors					
	K1	K2	K3	K4	K5	K6	K7	K8	K9	K10	R1	R2	R3	F1	F2	F3	F4	F5	F6
12	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0	1	0
13	0	0	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	1	0
14	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1
15	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1
16	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1
17	0	0	0	0	0	1	0	0	0	0	0	0	1	0	0	0	0	0	1
18	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0
19	0	0	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	1
20	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0
∑	4	3	3	3	4	3	0	0	0	0	6	8	6	0	0	2	7	6	5

Then the process of forming *support* 1 *itemset will be carried* out with a minimum amount of *support* = 15%. With the following formula:

$$\textbf{\textit{Support}}\ (\mathbf{A}) = \frac{\Sigma \text{ transaksi mengandung A}}{\Sigma \text{ transaksi}} * 100\%$$

Table 8: Support 1 Itemset

ID	*Count*	*Support*
K1	4/20	20%
K2	3/20	15%
K3	3/20	15%
K4	3/20	15%
K5	4/20	20%
K6	3/20	15%
R1	6/20	30%
R2	8/20	40%
R3	6/20	30%
F4	7/20	35%
F5	6/20	30%
F6	5/20	25%

After obtaining 1 itemset selected some data that meets the predetermined value, the value itself is the limit of the number used to obtain the selected number, the *support value* is 15%. Furthermore, the process of forming C2 or called 2 itemset with a minimum amount of support = 15%.

$$\textbf{Support}\ (\mathbf{A}, \mathbf{B}) = \frac{\Sigma \text{ transaksi mengandung A dan B}}{\Sigma \text{ transaksi}} * 100\%$$

Combinations of 2 itemsets that do not meet the minimum support requirements will be eliminated.

Table 9: Support 2 Itemset

ID	*Count*	*Support*
K1, R1	1/20	5%
K1, R2	1/20	5%
K1, R3	2/50	10%
K1, F4	2/50	10%
K1, F5	1/20	5%
K1, F6	1/20	5%
K2, R1	1/20	5%
K2, R2	2/50	10%
K2, R3	0/20	0

ID	*Count*	*Support*
K2, F4	1/20	5%
K2, F5	1/20	5%
K2, F6	0/20	0
K3, R1	1/20	5%
K3, R2	1/20	5%
K3, R3	1/20	5%
K3, F4	1/20	5%
K3, F5	0/20	0
K3, F6	1/20	5%

ID	*Count*	*Support*
K4, R1	1/20	5%
K4, R2	1/20	5%
K4, R3	1/20	5%
K4, F4	1/20	5%
K4, F5	1/20	5%
K4, F6	1/20	5%
K5, R1	1/20	5%
K5, R2	2/50	10%
K5, R3	1/20	5%
K5, F4	1/20	5%
K5, F5	1/20	5%
K5, F6	1/20	5%
K6, R1	1/20	5%
K6, R2	1/20	5%

ID	*Count*	*Support*
K6, R3	1/20	5%
K6, F4	1/20	5%
K6, F5	1/20	5%
K6, F6	1/20	5%
R1, F4	0/20	0
R1, F5	6/20	30%
R1, F6	0/20	0
R2, F4	6/20	30%
R2, F5	0/20	0
R2, F6	0/20	0
R3, F4	1/20	5%
R3, F5	0/20	0
R3, F6	5/20	25%

After obtaining 2 itemsets selected some data that meet the predetermined value, the value itself is the limit of the number used to obtain the selected number, the support value is 10%. From the Table above, the element above T means interrelated items, while F means no related items. The number of frequencies of the set item must be greater than the number of frequencies of the Itemsset θ. From the table above, f3 is obtained as follows:

Table 10: Support 3 Itemset

No	K1	R3	F4	*f*
1	0	0	0	*F*
2	0	0	0	*F*
3	0	0	0	*F*
4	0	0	1	*F*
5	1	0	0	*F*
6	0	0	1	*F*
7	1	0	1	*F*
8	0	0	1	*F*
9	0	0	1	*F*
10	0	0	1	*F*
11	0	0	0	*F*

No	K1	R3	F4	*f*
12	0	0	0	*F*
13	0	0	0	*F*
14	0	1	0	*F*
15	1	1	1	*T*
16	0	1	0	*F*
17	0	1	0	*F*
18	0	0	0	*F*
19	0	1	0	*F*
20	1	1	1	*T*
Jumlah				2

The process of forming C3 or called the minimum amount of *support* = 10%, the results of the calculation can be seen in the table below with the following formula:

$$\boldsymbol{Support}\ (\boldsymbol{A},\boldsymbol{B}) = \frac{\Sigma \text{ transaksi mengandung A, B \& C}}{\Sigma \text{ transaksi}} * 100\%$$

Table 11: Support 3 Itemset

ID	Count	Support
K1, R3, F4	2/20	10%

After all the high frequency patterns are found, then look for the association rules with the results of the frequency patterns shown in the table below as follows:

Table 12: Highest Frequency Pattern Results

ID	Count	Support
K1, R3, F4	2/20	10%

After all high-frequency patterns are found, then the association rules that meet the minimum requirements for confidence are searched by calculating confidence or association A→B, with a minimum confidence of 10%.

Table 13: Association Final Results

Rule	Confidence	
Jika keluhan yang dialami oleh pasien yaitu (K1) Sesak Nafas, dengan Riwayat penyakit yang dimiliki riwayat ISK sebelumnya maka faktor penyebabnya adalah Genetik	2/2	100%

And after multiplying between Support and Confidence, the result of the multiplication is 10% for the rule which will be the Best Rule.

Table 14: Best Rule

If antecendent then consequent	Support	Confidence	S*C
Jika keluhan yang dialami oleh pasien yaitu (K1) Sesak Nafas, dengan Riwayat penyakit yang dimiliki riwayat ISK sebelumnya maka faktor penyebabnya adalah Genetik	10	100%	10%

3.3. Flowchart Planning

The flowchart design can be described as follows:

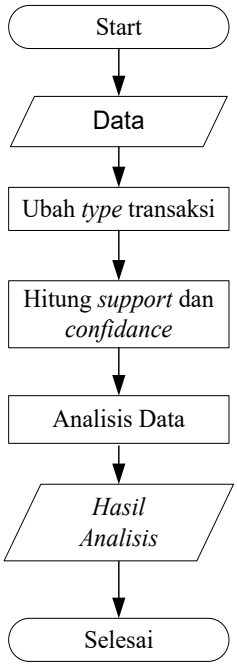


Fig. 3: Flowchart Rapidminer

Information:
In the image above, it is explained that it starts with start, Input data according to existing variables, Mans data transformation, Calculate support and confidance, Analyze data, Get analysis results, Then end with end.

3.4. Use Case Diagram

The following is a form of Diagram Use Case data in the form of a diagram that can explain a system process flow that will be built as shown in the image below.

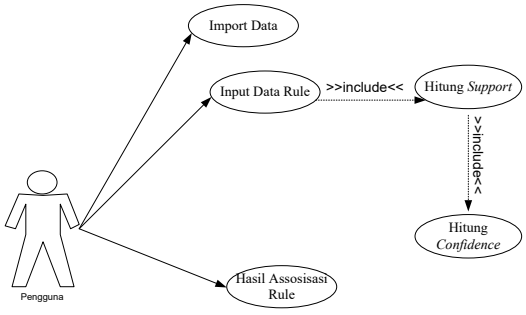


Fig. 4: Use Case Diagram

3.5. Results Overview

A description of the results that can be seen by the user and the commands or mechanisms used to control the operation and enter a data.

a. Import Data

The data processing process to find disease patterns in medical record data using a priori method is carried out using the rappid miner application with the following view:

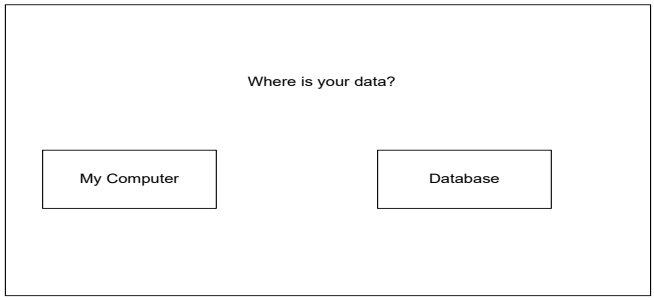


Fig. 5: Import Data

The process of retrieving data that has been stored in data format uses ms.excel. Add data is done by clicking the add data menu, then a data location will appear, then select the file to be added, which can be seen in the image above import data excel.

b. Apriori Process

In this view, a priori algorithm process will be carried out by dragging and dropping on the repository and operators tabs. The process of combining tabular data with operators related to a priori algorithms, can be seen in the view below.

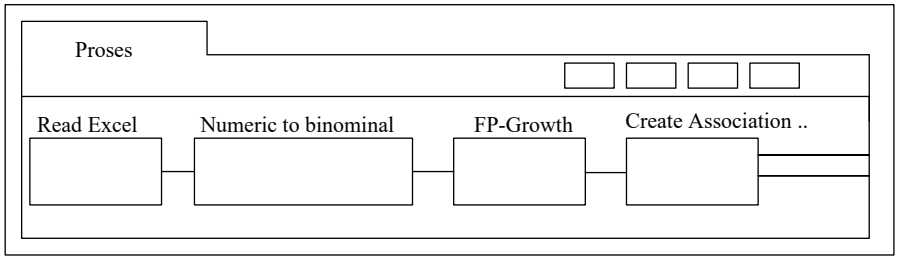


Fig. 6: Apriori process

3.6. Rapidminer Studio Display

The image below shows the results of the analysis of association rules in the form of text (text view) in the RapidMiner 5.3 application. Each row in this view shows an association rule formed from medical record data, with a "premise → conclusion" format and a confidence value that indicates the level of trust in the rule. For example, the rule "[Allergy, Cold weather] → [No history of serious illness]" with a confidence of 0.379 means that 37.9% of patients with the causative factors of Allergy, Cold weather have no history of serious illness. In addition, there are also rules such as "[diarrhea] → [No history of serious illness]" that show a pattern of association between several characteristics of the patient. The confidence value listed describes how strong the relationship between the items is in the data. This view helps users to understand asthma disease patterns, making them very useful in decision-making. The results of the association rule formation from the tests on the RapidMiner Studio application can be seen in the following image:

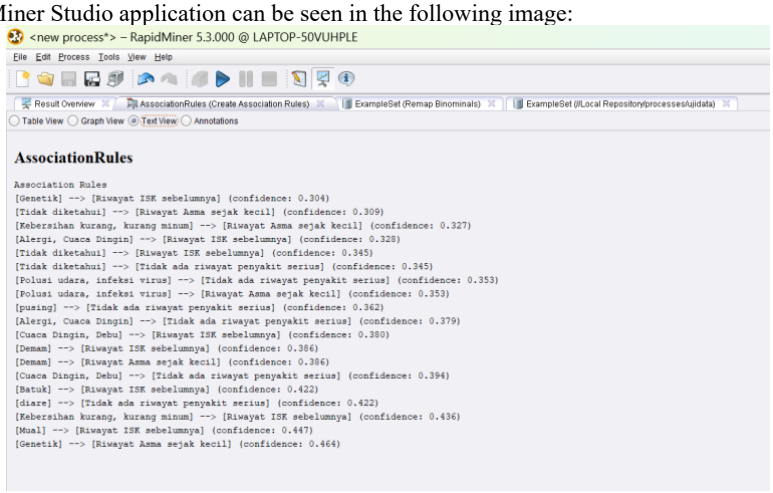


Fig. 7: Associate Rules view

4. Conclusions and Suggestions

4.1 Conclusion

Based on the results of the research and discussions that have been carried out, conclusions can be drawn, namely as follows:

1. From testing using the RapidMiner application , this study identified complaints, disease history, and causative factors. The results showed that there were 19 association rules formed with the Best rule value in the 2 item sets, the highest support value of 7%, and the confidance value of 46% in the 2 item sets. with the rule "If the Causative Factor is Genetic, Then the History of Asthma Disease History Since Childhood".
2. The 3 set items show that there are 5 association rules formed with the highest Best rule value, 14% support value, and 62% confidance value. with the rule "If the causative factor is Genetic, the complaint is dizzy, then the disease history has a history of asthma since childhood".
3. Using the A priori algorithm and the RapidMiner application, this study was able to provide information quickly about the patterns of asthma in the Datar City Health Center. As well as being able to find the association rule from the correlation of medical record data at the Datar City Health Center using the A priori method.

4.1 Suggestions

Given the limitations possessed by the author, both knowledge and thought, the author can provide suggestions that can be used as a reference for the results of this research in the future, which are as follows:

1. For further research, it is recommended to use various other methods in the application of data mining to medical record data at the Datar City Health Center in order to obtain a more diverse pattern.
2. The Datar City Health Center is advised to develop an automatic recommendation system based on algorithms or other data mining techniques.
3. For further research, it can be carried out to dig deeper or expand the findings of this study.
4. For Readers, it is recommended to understand the meaning of each of these metrics (Support, Confidence, Lift) in order to be able to interpret the strength and relevance of each rule found.

References

[1] Z. A. Sofyan, R. Kurniawan, and Y. A. Wijaya, "OPTIMASI PARAMETER ALGORITMA K-MEANS PADA DATA REKAM MEDIS KLINIK ANANDA MEDIKA RANCAH," *Jurnal Mahasiswa Teknik Informatika*, vol. 7, no. 6, pp. 3667–3672, Dec. 2023.

[2] Nola Ritha, E. Suswaini, and W. Pebriadi, "Penerapan Association Rule Menggunakan Algoritma Apriori Pada Poliklinik Penyakit Dalam (Studi Kasus: Rumah Sakit Umum Daerah Bintan)," *Jurnal Sains dan Informatika*, vol. 7, no. 2, pp. 222–230, Dec. 2021, doi: 10.34128/jsi.v7i2.329.

[3] A. Pirman, A. Hanifa, and G. Triyono, "Implementasi Algoritma Apriori Pada Penjualan Makanan Ringan dan Minuman Kesehatan," *Jurnal Pendidikan Teknologi Informasi*, vol. 4, no. 1, pp. 204–215, Mar. 2024.

[4] Amna *et al.*, *Data Mining Data mining*, 1st ed., vol. 1, no. 1. Padang: PT GLOBAL EKSEKUTIF TEKNOLOGI, 2023.

[5] Amna *et al.*, *Data Mining Data mining*, 1st ed., vol. 1, no. 1. Padang: PT GLOBAL EKSEKUTIF TEKNOLOGI, 2023.

[6] M. Hafiz and S. Anggraini, "Klasterisasi Penyakit Menular Di Indonesia Menggunakan Metode K-Means Clustering," *Journal of Computer)online)*, vol. 4, no. 1, pp. 50–57, Mar. 2024, Accessed: Mar. 11, 2025. [Online]. Available: https://jurnal.stmikroyal.ac.id/index.php/j-com/article/view/3033/1374

[7] F. Hidayat, *Konsep Dasar Sistem Informasi Kesehatan*, 1st ed., vol. 1. Yogyakarta: CV BUDI UTAMA), 2020. Accessed: May 24, 2025. [Online]. Available: https://www.google.co.id/books/edition/Konsep_Dasar_Sistem_Informasi_Kesehatan/nvhZEQAAQBAJ?hl=en&gbpv=0

[8] Pitrawati and A. Sanjaya, "REKAYASA PERANGKAT LUNAK PERHITUNGAN HARGA POKOK PRODUKSI METODE FULL COSTING PADA UMKM MITRA CAKE DI BANDAR LAMPUNG," *Jurnal informasi dan Komputer*, vol. 11, no. 2, pp. 154–162, Nov. 2021.

[9] A. Munazilin and F. Santoso, *logika dan algoritma pemrograman*, 1st ed. Situ Bondo: CV. AA. Rizky, 2021.

[10] R. Swastika, S. Mukodimah, F. Susanto, M. Muslihudin, and S. Ipnuwati, *IMPLEMENTASI DATA MINING (Clastering, Association, Prediction, Estimation, Classification)*, 1st ed., vol. 1. Indramayu Jawa Barat: CV. Adanu Abimata, 2023.