Minimum Threshold Determination Method based on Dataset Characteristics in Association Rule Mining

Erna Hikmawati (✉ ema@pasim.ac.id)

Bandung Institute of Technology: Institut Teknologi Bandung <https://orcid.org/0000-0002-6162-5341>

Nur Ulfa Maulidevi

Bandung Institute of Technology: Institut Teknologi Bandung

Kridanto Surendro

Bandung Institute of Technology: Institut Teknologi Bandung

Research

Keywords: Minimum threshold, adaptive rule, association rule

Posted Date: July 27th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-728509/v1>

License: This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Version of Record: A version of this preprint was published at Journal of Big Data on November 25th, 2021. See the published version at <https://doi.org/10.1186/s40537-021-00538-3>.

Minimum Threshold Determination Method based on Dataset Characteristics in Association Rule Mining

Erna Hikmawati^{1*}, Nur Ulfa Maulidevi², Kridanto Surendro³

School of Electrical Engineering and Informatics, Institut Teknologi Bandung, Indonesia

**Corresponding author: ernahikma21@students.itb.ac.id*

Abstract

The process of extracting data to obtain useful information is known as data mining. Furthermore, one of the promising and widely used techniques for this extraction process is association rule mining. This technique is used to identify interesting relationships between sets of items in a dataset and predict associative behavior for new data. The first step in association rule mining is the determination of the frequent item set that will be involved in the rule formation process. In this step, a threshold is used to eliminate items excluded in the frequent itemset which is also known as the minimum support. Furthermore, the threshold provides an important role in determining the number of rules generated. However, setting the wrong threshold leads to the failure of the association rule mining to obtain rules. Currently, the minimum support value is determined by the user. This leads to a challenge that becomes worse for a user that is ignorant of the dataset characteristics. In this study, a method was proposed to determine the minimum support value based on the characteristics of the dataset. Furthermore, this required certain criteria to be used as thresholds which led to more adaptive rules according to the needs of the user. The results of this study showed that 6 from 8 datasets, obtained a rule with lift ratio >1 using the minimum threshold value that was determined through this method.

Keywords: Minimum threshold, adaptive rule, association rule

1. Introduction

The current developments led to large quantities of data growth which were accessible from various locations everywhere at all times [1]. Data availability is a huge asset for an organization because it contains a lot of useful information. Therefore, Data Mining is a way to extract this useful information and patterns [1–6]. Furthermore, the Association rule is one of the most basic study topics widely used in Data Mining [1–5,7–11]. It is an important data analysis method that discovers and identifies interesting relationships between sets of items in a dataset and predicts association relationships for new data [7,8,12].

The basic concept of the association rule is to generate rules based on items with frequent occurrence in a transaction. This includes two main processes, namely the determination of frequent itemset and the process of forming rules. A frequent itemset is a collection of items that have a higher frequency of occurrence compared to the threshold value specified in the transaction. This value is also known as the minimum support. The real-time applications of the association rule have several challenges ranging from the presence of very large, multiple, and heterogeneous data sources to difficulty in the determination of the minimum support [7]. Furthermore, the first step in the association rule is to determine the minimum support value [13,14]. Currently, this value is determined by the user despite the difficulty associated with this process. Another difficulty is that the user is ignorant of the dataset characteristics to be searched during this process [3,6,11,15–18].

The minimum support value has an important role in determining the value of several rules generated [3]. However, determining a wrong value leads to the failure of the association rule to obtain the required rule [19]. Furthermore, determining a value that is too low leads to the involvement of too many items in the rule formation process. Conversely, when a value that is too high is obtained, fewer items are involved which leads to the loss of a lot of information. The minimum support values when compared in terms of two factors namely, time and memory showed that a value that is too low will require more of these two factors when compared to a higher value [6,11,20]. Currently, the method used to determine the minimum support is based on the intuitiveness of the user. The association rule process is repeated through the change of the minimum support value when the rule obtained does not match. However, this does not guarantee that a rule will be generated with the input minimum support value.

Decision-makers are required to make the right strategic choices which lead to the adjustment of current conditions which are full of volatility, uncertainty, complexity, and ambiguity. This is known as the VUCA (Volatility, Uncertainty, Complexity, Ambiguity) world [21]. Furthermore, the current conditions also demand the adaptive transformation of the association rules to meet the needs of the users. Decision-makers have different criteria which are used to obtain the information they need. Therefore the rule formation process should pay attention to the criteria desired by the users to obtain adaptive Rules [22].

Furthermore, to meet these needs, the determination of frequent itemset should not only be based on the frequency of occurrence of an item but also involve another criterion called item utility [10,23–25]. Therefore, the resulting rule will be adaptive according to the criteria desired by the user.

This study which proposes a method to determine the minimum support value based on the characteristics of the

dataset and other criteria can also be used as assessments in the process of forming the rule. The main contributions of this study include:

1. The proposition of a method to determine the minimum support value based on the characteristics of the dataset.
2. The generation of a minimum support value-based not only on the frequency of occurrence of items but also based on certain criteria affecting the rule formation process such as price, profit, or item usability value.
3. The generation of a minimum support value certainly forms a rule that does not require the user to determine this value at the beginning.

This study is divided into several parts. Parts II, III, IV and V describe and explains related studies, proposed methods, experimental results and discussion and conclusions respectively.

2. Related Work

2.1. Association Rule

The association rule mining was first proposed in a study by Agrawal [26,27]. This study developed a focus on different topics ranging from high utility itemset [23,25,28–31], Top-K [5,11,32], Skyline [6,10,33], Multicriteria [34–36] and Meta association rules [37–39]. Furthermore, the main concept of association rule mining can be formally defined as follows: $I = \{i_1, i_2, \dots, i_m\}$ is a set of items and D is a set of transactions T , where a set of transactions T is also a set of items, therefore, $T \subseteq I$. Furthermore, provided that A is a set of items, Transaction T is said to contain A if and only if $A \subseteq T$. Association rules are from $A \rightarrow B$, where $A \subseteq I$, $B \subseteq I$, and $A \cap B = \emptyset$. In addition, rule $A \rightarrow B$ has support in transaction set D provided that $s\%$ of transactions in D contain $A \cup B$ [4,22,40]. This support is shown by the occurrence frequency of items which is calculated by observing the ratio between the frequency of transactions containing itemset A divided by the total transactions. In general, it can be seen in the formula below [37]:

$$Supp(A \rightarrow B) = \frac{|\{t \in D \mid A \cup B \subseteq t\}|}{|D|} \quad [1]$$

Where:

- Supp is the support value
- A is the antecedent of the rule in the form of itemset
- B is consequent in the form of itemset
- t is a transaction containing A and B
- D is the total transaction

Confidence is another threshold used in determining the rule apart from the support value. It is the ratio between the number of transactions containing items A and B divided by transactions containing item A for rule $A \rightarrow B$.

The confidence value of rule $A \rightarrow B$ is obtained by the formula [37]:

$$Conf(A \rightarrow B) = \frac{|\{t \in D \mid A \cup B \subseteq t\}|}{|\{t \in D \mid A \subseteq t\}|} \quad [2]$$

Where:

- Conf is Confidence
- A is the antecedent of the rule in the form of itemset
- B is consequent in the form of itemset
- t is a transaction containing A and B
- D is the total transaction

The general framework of the association rule is to extract a rule with a support value for an item that exceeds the minimum support and confidence value. Therefore, this

rule exceeds the minimum confidence value specified by the user. In this case, it can be stated that $A \rightarrow B$ is included in the frequent and confidence category (strong rule) provided that $Supp(A \rightarrow B) \geq minsupp$ and $Conf(A \rightarrow B) \geq minconf$ respectively.

To evaluate rule, lift ratio can be used [35,36,41–48]. Lift ratio is the ratio between the support value of the rule with the antecedent and consequent support value. The higher of lift ratio value, the more interest rule or called the strong rule. A rule is called interest and strong rule if the lift ratio value > 1 because it shows that is a positive correlation between the premise and conclusion of this association rule [36,48]. The lift ratio value can be calculated by the formula [36,41–43,47,48]:

$$Lift(A \rightarrow B) = \frac{sup(A \cup B)}{sup(A) \times sup(B)} = \frac{conf(A \rightarrow B)}{sup(B)} \quad [3]$$

Where:

- Lift is Lift Ratio value
- A is the antecedent of the rule in the form of itemset
- B is consequent in the form of itemset
- Supp is the support value
- Conf is the confidence value

2.2. Minimum Support

The basic concept of the association rule requires the user to specify a minimum support value at the beginning. This value usually applies uniformly to all items, although in reality, different items may have different criteria for assessing them. Therefore, studies about multiple minimum support which state that the support value should vary for different items have emerged [9,23,49,50]. However, the implementation of this system adds a task for the user, which is to determine the minimum support for each item.

Furthermore, the difficulty of determining the minimum support value by the user has led to a new field of study, namely association rules without the use of minimum support such as Top-K [5,11,19,32,51] and Skyline [6,10,33,52]. The Top-k association rule does not require the minimum support value because in this method the user is only asked to determine the k value, which is the number of rules that will be generated in the rule formation process. Therefore, it is easier for users to determine the k value because they explicitly know the number of rule results they want to obtain.

The Skyline algorithm was first proposed by Borzsony [52] and later developed by Goyal [33]. It is a point that is not dominated by other points [52]. This algorithm was combined by Jerry Chun-Wei Lin [10] and Jeng-Shyang Pan [6] to produce association rules. Furthermore, this study does not use minimum support but instead makes use of the maximum utility (utilmax) obtained from each iteration performed by the utility list structure.

Based on the results of the previous review [20], the association rules require a minimum support value to reduce the items involved in the rule formation process. Furthermore, the existence of this threshold can reduce the time and memory required in the rule formation process. Therefore, regardless of the term, starting from Minimum Utility, Maximum Utility to Minimum Support, this threshold is required for the association rule process. Meanwhile, processes without the threshold will involve all the items present and require more time and memory, which will lead to rules that may not be as desired.

3. The Proposed Model

To determine the minimum support value in the association rule, a special method is needed to determine the value according to the characteristics of the dataset which involves certain criteria according to the desires of the user. This study proposed an adaptive support method to answer these needs, as shown in Figure 1.

Fig.1 Adaptive support method flow

Furthermore, the algorithm used to determine the minimum support value based on the characteristics of the database and certain items criteria (utility) is shown in algorithm 1.

```

Pseudocode adaptive_support
Declare
     $D = \text{Dataset}$ 
     $S = \text{support value for each item}$ 
     $U = \text{utility value for each item}$ 
     $n = \text{frequency of item}$ 
Begin
    Foreach  $D$  as  $d$ :
         $S(d) = n(d) / |D|$ 
         $U(t) = S(d) * U(d)$ 
         $\text{Sum} = \text{Sum} + U(t)$ 
     $\text{Avesup} = \text{Sum} / |D|$ 
     $\text{Min\_threshold} = \text{Avesup} / |D|$ 
End

```

Algorithm.1 Adaptive Support Algorithm

In the adaptive support model, the determination of the minimum support value is based on several factors which include:

1. Characteristics of the dataset

In determining the minimum support value, the user should know in advance, the characteristics of the dataset that will be processed for rule formation. This is because several characteristic factors of the dataset will affect the suitable minimum support value. Furthermore, for an adaptive support mode, the factor used to determine the minimum support value is shown from the number of items contained in the transaction, the number of transactions, the average number of items in each transaction, and the support value for each item.

2. Specific criteria (utility for each item)

To be able to produce adaptive rules, a criterion is required as an assessment for an item known as a frequent itemset with the right to be involved in the rule formation process. Furthermore, these criteria can be determined by the user and each item has its own criteria values. For example, a user who wants to get a rule for the most expensive item has an involved criterion such as, the item price. In addition, another example is when a user wants a rule with the biggest profit then, the criterion for the item involved is the profit from each item.

Based on both types of inputs, the calculation process is carried out according to algorithm 1 to obtain adaptive support. The calculation stages are as follows:

1. Calculation of the support value for each item in the dataset with the following formula:

$$Sup(d) = \frac{n(d)}{|D|} \quad [4]$$

2. Calculation of the utility for each item in the dataset with the following formula:

$$Util(d) = Sup(d) \times U(d) \quad [5]$$

3. Calculation of the average utility for the entire transaction with the following formula:

$$ave\ sup = \frac{\sum Util(d)}{|D|} \quad [6]$$

4. Calculation of the minimum threshold value used for the rule formation process which is the average utility value divided by the total existing transactions through the following formula:

$$\min\ sup = \frac{ave\ sup}{|D|} \quad [7]$$

5. Performance of the rule formation process with existing methods through the apriori algorithm, fpgrowth, or other algorithms

Where:

- $Sup(d)$ = support value for an item
- $n(d)$ = number of occurrences for an item
- $|D|$ = total transaction
- $U(d)$ = utility value for an item
- $Util(d)$ = utility and support value for an item
- $Avesup$ = Average utility of the item
- $Minsup$ = minimum threshold value (item density level)

4. Experiment and Result

4.1. Data Experiment

The special dataset for the association rule case used in this study was obtained from SPMF [53]. The dataset criteria used can be seen in table 1.

Table 1 Characteristics of the Dataset

Dataset	Number of transactions	Number of Items	Average Number of Items on each Transaction
Mushrooms	8,416	119	23
Chess	3,196	75	37
Skin	245,057	11	4
Tes	3	5	3
Recordlink	574913	29	10
Retail	88,162	16,470	10.30
Accidents	340,183	468	33.8
Connect	67,557	129	43

The study environment used was a laptop with an Intel Core i7-8550U CPU @ 1.80 GHz 1.99GHz, 16 GB of Installed memory (RAM), and a 500GB SSD Hard drive.

4.2. Result

The proposed adaptive support method was carried out on 8 datasets according to table 1. In this study, other criteria used to determine the minimum threshold were the same for all items, namely valued 1. Furthermore, the results from all 6 of the 8 datasets obtained the appropriate minimum support value and rules with lift ratio>1. The results of this adaptive support calculation process are shown in table 2.

Table 2 Adaptive Support Calculation Results

Dataset	Density Rate (Average occurrence of each item in all transactions)	Minimum Support (Density / Number of Transactions)	Number of Rules with lift ratio >1
Mushrooms	1626.445	19.32%	19011494
Chess	1577.413	49.35%	320577653
Retail	74.61	0.08%	33170
Accidents	36992.1812	11.59%	-
Connect	22519	33.33%	-
Skin	8165.67	33.36%	2
Tes	1.8	60%	4
Recordlink	213691.1154	37.16%	9892

The trial of the proposed adaptive support method used two basic algorithms in the association rule, namely Apriori and Fpgrowth [10]. The Apriori algorithm used a level-wise approach and generated candidate items for each level [26]. Meanwhile, the fpgrowth algorithm used the pattern growth method but did not generate candidates for each level. After the elimination of itemset with minimum support, the next step was to obtain a Frequent Pattern-Tree [54]. The trial was carried out from the highest minimum support value at 100% and then repeated while a reduction of this value occurred. All minimum support values are not tested on all datasets because the number of rules is different. The more rules that are generated, the more difficult it will be in the

decision-making process. For example, in the connect dataset with a minimum support value of 70%, it produces a rule 392,469,141 using the fpgrowth algorithm and consumes a lot of resources (runtime and memory) so the test is stopped. Table 3-6 shows the test results of each dataset for each minimum support value using the a priori and fpgrowth algorithms. The blue color in the table is the minimum support value proposed by the adaptive support method.

Table 3 Number of Rules for each minimum support value with apriori algorithm

Minimum Support value	Number of Rules							
	Chess	Mushroom	Skin	Accidents	Connect	Retail	RecordLink	Test
0%						35166		
5%			256			32	231656	32
10%		393890460	166			8	208204	32
20%		19174370	84			2	86252	32
30%		78888	8			2	59754	32
40%		5904	2	2764708		0	14012	4
50%	880936478	1248	0	375774		0	5906	4
60%	83864464	266	0	54170		0	3268	4
70%	8111370	180	0	8226		0	1310	0
80%	552564	52	0	1432		0	180	0
90%	10742	22	0	180	3640704	0	180	0
100%	0	0	0	0	0	0	0	0

Table 4 Number of Rules for each minimum support value with the fpgrowth algorithm with lift

ratio>1

Minimum Support value	Number of Rules							
	Chess	Mushroom	Skin	Accidents	Connect	Retail	RecordLink	Test
0%			474			33170	1,640,695	25
5%			209			30	130,524	25
10%		370,981,205	156			8	121,129	25
20%		18,982,160	76	209,815,664		2	50,750	25
30%		58,930	8	14,910,124		2	29,580	25
40%		3,046	2	1,574,428		0	5,611	4
50%	280,026,979	515	0	209,256		0	3,070	4
60%	32,975,301	129	0	30,945		0	2,134	4
70%	3,232,975	104	0	5,036	392,469,141	0	966	0
80%	189,334	45	0	841	24,086,200	0	110	0
90%	2,564	22	0	154	341,469	0	110	0
100%	0	0	0	0	0	0	0	0

Table 5 Runtime for each minimum support value with apriori algorithm

Minimum Support Value	Runtime (MS)							
	Chess	Mushroom	Skin	Accidents	Connect	Retail	RecordLink	Test
0%						1002		
5%			4			140	216	1
10%		800330	1			94	184	0
20%		10030	0			109	131	0
30%		531	0			117	85	0
40%		94	0	3341		93	12	0
50%	880936478	31	0	306		94	20	0
60%	109935	20	0	58		120	8	0
70%	8340	16	0	14		78	1	0
80%	469	16	0	2		78	0	0
90%	29	16	0	1	4820	78	1	0
100%	11	12	0	0	0	78	0	0

Table 6 Runtime for each minimum support value with fpgrowth algorithm with lift ratio>1

Minimum Support Value	Runtime (MS)							
	Chess	Mushroom	Skin	Accidents	Connect	Retail	RecordLink	Test
0%			4				2683	0
5%			0			0	175	0
10%		1213180	0			0	168	0
20%		37757	0	1058700		0	77	0
30%		73	0	64794		0	44	0
40%		4	0	6297		0	7	0
50%	1431849	3	0	816		0	7	0
60%	121658	0	0	159		0	7	0
70%	6125	0	0	35	996275	0	4	0
80%	367	0	0	22	45541	0	0	0
90%	4	0	0	9	546	0	0	0
100%	0	0	0	2	0	0	0	0

Based on tables 3-6, it was shown that for each different dataset, the same minimum support value led to a varied number of rules. This was because of the dataset characteristics consisting of the number of items, the average item in each transaction, and the different density values. Furthermore, the trials that have been carried out showed that the higher minimum support value led to a smaller number of rules and runtimes and vice versa. The minimum support value generated by the adaptive rule method when compared with the results of trials using the fpgrowth and apriori algorithms is shown in Figure 2.

Fig.2 Comparison of minimum support values using Apriori, Fpgrowth, and Adaptive Support

Figure 2 shows that 6 out of 8 datasets or about 75% of the total experimental data obtained a minimum support value between the minimum and maximum values for each algorithm. This means that the value of the proposed adaptive support can generate a rule and if viewed from the quality, adaptive support produces a rule that has a lift ratio value > 1 . In addition, with the proposed adaptive support value, users no longer need to guess the appropriate minimum support value that can generate rules. So that there is no longer a possibility that with the specified minimum support value, it does not get a rule, or the number of rules is 0 and must repeat the experiment with another minimum support value.

Figure 3 shows that the smaller the minimum support value, the more rules are generated. The red line in Figure 3 shows the proposed minimum support value based on the adaptive support method.

Fig.3 Test results of the number of rules for each minimum support value

5. Conclusion and Future Work

The minimum support value is important and has a big influence on the rule formation process in the association rule. The errors obtained when determining this resulting rule were not desired. Furthermore, the determination of this value was difficult especially if the user is ignorant of the dataset characteristics as a factor which was used as a reference in determining the minimum support value. This is because datasets with a high level of density required

a larger minimum support value when compared to datasets with low density. This density level is the average value of the occurrence frequency of each item in the dataset. Therefore, an adaptive support model which determined the minimum support value based on the characteristics of the dataset was proposed. In addition, certain criteria leading to adaptive rules following the desires of the user were included.

From the study results on 9 datasets, 8 datasets obtained a rule based on the minimum support value suggested by the adaptive support model. In the future, the author will integrate the multicriteria system model [55] and this adaptive support method to obtain a completely adaptive rule model [22].

DECLARATIONS

Acknowledgments

We would like to thank Institut Teknologi Bandung and Pasim National University for supporting this research.

Authors' contributions

The author confirms the sole responsibility for this manuscript fully as a sole author for the following: study conception and design, data collection, analysis and interpretation of results, and manuscript preparation. The author read and approved the final manuscript.

Funding

I sincerely thank LPDP (Indonesia Endowment Fund for Education), Ministry of Finance, Republic Indonesia for providing me with the financial support for this research. This research was funded by LPDP (Indonesia Endowment Fund for Education), Ministry of Finance, Republic Indonesia.

Availability of data and materials

The original data used for this study is available in: <http://www.philippe-fournier-viger.com/spmf/index.php?link=datasets.php>.

Competing interests

The author reports no potential conflict of interest.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

References

1. Luna JM, Fournier-Viger P, Ventura S. Frequent itemset mining: A 25 years review. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery* [Internet]. 2019 [cited 2019 Sep 25]; Available from: <https://onlinelibrary.wiley.com/doi/abs/10.1002/widm.1329>
2. Prajapati DJ, Garg S, Chauhan NC. Interesting association rule mining with consistent and inconsistent rule detection from big sales data in distributed environment. *Future Computing and Informatics Journal*. 2017;2:19–30.
3. Selvi CSK, Tamilarasi A. An Automated Association Rule Mining Technique With Cumulative Support Thresholds. :12.
4. Zhang C, Zhang S. *Association rule mining: models and algorithms*. Berlin ; New York: Springer; 2002.
5. Ryang H, Yun U. Top-k high utility pattern mining with effective threshold raising strategies. *Knowledge-Based Systems*. 2015;76:109–26.
6. Pan J-S, Lin JC-W, Yang L, Fournier-Viger P, Hong T-P. Efficiently mining of skyline frequent-utility patterns. *Intelligent Data Analysis*. 2017;21:1407–23.
7. Zhang S, Wu X. *Fundamentals of association rules in data mining and knowledge discovery: Fundamentals of association rules*. *WIREs Data Mining Knowl Discov*. 2011;1:97–116.
8. Weng C-H, Chen Y-L. Mining fuzzy association rules from uncertain data. *Knowl Inf Syst*. 2010;23:129–52.
9. Dahbi A, Balouki Y, Gadi T. Using Multiple Minimum Support to Auto-adjust the Threshold of Support in Apriori Algorithm. In: Abraham A, Haqiq A, Muda AK, Gandhi N, editors. *Proceedings of the Ninth International Conference on Soft Computing and Pattern Recognition (SoCPaR 2017)* [Internet]. Cham: Springer International Publishing; 2018 [cited 2021 Jan 27]. p. 111–9. Available from: http://link.springer.com/10.1007/978-3-319-76357-6_11
10. Lin JC-W, Yang L, Fournier-Viger P, Hong T-P. Mining of skyline patterns by considering both frequent and utility constraints. *Engineering Applications of Artificial Intelligence*. 2019;77:229–38.
11. Duong Q-H, Liao B, Fournier-Viger P, Dam T-L. An efficient algorithm for mining the top- k high utility itemsets, using novel threshold raising and pruning strategies. *Knowledge-Based Systems*. 2016;104:106–22.
12. Vu L, Alaghband G. An efficient approach for mining association rules from sparse and dense databases. *2014 World Congress on Computer Applications and Information Systems (WCCAIS)* [Internet]. Hammamet, Tunisia: IEEE; 2014 [cited 2021 Jun 20]. p. 1–8. Available from: <http://ieeexplore.ieee.org/document/6916550/>
13. Boley M, Grosskreutz H. Approximating the number of frequent sets in dense data. *Knowl Inf Syst*. 2009;21:65–89.
14. Wazir S, Beg MMS, Ahmad T. Comprehensive mining of frequent itemsets for a combination of certain and

uncertain databases. *Int j inf tecnol.* 2020;12:1205–16.

15. Zhang S, Wu X, Zhang C, Lu J. Computing the minimum-support for mining frequent patterns. *Knowl Inf Syst.* 2008;15:233–57.

16. Alias S, Razali MN, Tan Soo Fun, Sainin MS. Sequential pattern mining using personalized minimum support threshold with minimum items. 2011 International Conference on Research and Innovation in Information Systems [Internet]. Kuala Lumpur, Malaysia: IEEE; 2011 [cited 2021 Jun 20]. p. 1–6. Available from: <http://ieeexplore.ieee.org/document/6125688/>

17. Ghafari SM, Tjortjis C. Association Rules Mining by Improving the Imperialism Competitive Algorithm (ARMICA). In: Iliadis L, Maglogiannis I, editors. *Artificial Intelligence Applications and Innovations* [Internet]. Cham: Springer International Publishing; 2016 [cited 2020 Sep 10]. p. 242–54. Available from: http://link.springer.com/10.1007/978-3-319-44944-9_21

18. Lin W-Y, Tseng M-C. Automated support specification for efficient mining of interesting association rules. *Journal of Information Science.* 2006;32:238–50.

19. Salam A, Khayal MSH. Mining top-k frequent patterns without minimum support threshold. *Knowl Inf Syst.* 2012;30:57–86.

20. Hikmawati E, Surendro K. How to Determine Minimum Support in Association Rule. *Proceedings of the 2020 9th International Conference on Software and Computer Applications* [Internet]. Langkawi Malaysia: ACM; 2020 [cited 2020 Aug 15]. p. 6–10. Available from: <https://dl.acm.org/doi/10.1145/3384544.3384563>

21. Giones F, Brem A, Berger A. Strategic decisions in turbulent times: Lessons from the energy industry. *Business Horizons.* 2019;62:215–25.

22. Hikmawati E, Maulidevi NU, Surendro K. Adaptive rule: A novel framework for recommender system. *ICT Express.* 2020;S2405959520300916.

23. Krishnamoorthy S. Efficient mining of high utility itemsets with multiple minimum utility thresholds. *Engineering Applications of Artificial Intelligence.* 2018;69:112–26.

24. Liu M, Qu J. Mining high utility itemsets without candidate generation. *Proceedings of the 21st ACM international conference on Information and knowledge management - CIKM '12* [Internet]. Maui, Hawaii, USA: ACM Press; 2012 [cited 2019 Sep 26]. p. 55. Available from: <http://dl.acm.org/citation.cfm?doid=2396761.2396773>

25. Nguyen LTT, Nguyen P, Nguyen TDD, Vo B, Fournier-Viger P, Tseng VS. Mining high-utility itemsets in dynamic profit databases. *Knowledge-Based Systems.* 2019;175:130–44.

26. Agrawal R. Fast Algorithms for Mining Association Rules. *Proceedings of the 20th International Conference on Very Large Data Bases.* San Francisco, CA, USA; 1994. p. 487–99.

27. Agrawal R, Imielinski T, Swami A, Road H, Jose S. Mining Association Rules between Sets of Items in Large Databases. *Proceedings of the 1993 ACM SIGMOD International Conference on Management of Data.* Washington, DC, USA; 1993. p. 207–61.

28. Yao H, Hamilton HJ. Mining itemset utilities from transaction databases. *Data & Knowledge Engineering.* 2006;59:603–26.

29. Ahmed CF, Tanbeer SK, Byeong-Soo Jeong, Young-Koo Lee. Efficient Tree Structures for High Utility Pattern Mining in Incremental Databases. *IEEE Transactions on Knowledge and Data Engineering.* 2009;21:1708–21.

30. Lin C-W, Hong T-P, Lu W-H. An effective tree structure for mining high utility itemsets. *Expert Systems with Applications.* 2011;38:7419–24.

31. Tseng VS, Shie B-E, Wu C-W, Yu PS. Efficient Algorithms for Mining High Utility Itemsets from Transactional Databases. *IEEE Transactions on Knowledge and Data Engineering*. 2013;25:1772–86.
32. Fournier-Viger P, Wu C-W, Tseng VS. Mining Top-K Association Rules. In: Kosseim L, Inkpen D, editors. *Advances in Artificial Intelligence* [Internet]. Berlin, Heidelberg: Springer Berlin Heidelberg; 2012 [cited 2019 Sep 26]. p. 61–73. Available from: http://link.springer.com/10.1007/978-3-642-30353-1_6
33. Goyal V, Sureka A, Patel D. Efficient Skyline Itemsets Mining. *Proceedings of the Eighth International C* Conference on Computer Science & Software Engineering - C3S2E '15* [Internet]. Yokohama, Japan: ACM Press; 2008 [cited 2019 Sep 26]. p. 119–24. Available from: <http://dl.acm.org/citation.cfm?doid=2790798.2790816>
34. Choi DH, Ahn BS, Kim SH. Prioritization of association rules in data mining: Multiple criteria decision approach. *Expert Systems with Applications*. 2005;29:867–78.
35. Ait-Mlouk A, Gharnati F, Agouti T. An improved approach for association rule mining using a multi-criteria decision support system: a case study in road safety. *European Transport Research Review* [Internet]. 2017 [cited 2019 Sep 25];9. Available from: <http://link.springer.com/10.1007/s12544-017-0257-5>
36. El Mazouri FZ, Abounaima MC, Zenkour K. Data mining combined to the multicriteria decision analysis for the improvement of road safety: case of France. *Journal of Big Data* [Internet]. 2019 [cited 2019 Sep 25];6. Available from: <https://journalofbigdata.springeropen.com/articles/10.1186/s40537-018-0165-0>
37. Ruiz MD, Gómez-Romero J, Molina-Solana M, Campaña JR, Martín-Bautista MJ. Meta-association rules for mining interesting associations in multiple datasets. *Applied Soft Computing*. 2016;49:212–23.
38. Ruiz MD, Gómez-Romero J, Molina-Solana M, Ros M, Martín-Bautista MJ. Information fusion from multiple databases using meta-association rules. *International Journal of Approximate Reasoning*. 2017;80:185–98.
39. Xiong J, Liu Z. Fuzzy meta association rules based on hierarchy theory based analysis of epidemic incidence of hand, foot and mouth disease in children. *Future Generation Computer Systems*. 2019;91:574–8.
40. Kantardzic M. *DATA MINING Concepts, Models, Methods, and Algorithms*. Second Edition. Hoboken, New Jersey: John Wiley & Sons, Inc.; 2011.
41. Alam TM, Shaukat K, Hameed IA, Khan WA, Sarwar MU, Iqbal F, et al. A novel framework for prognostic factors identification of malignant mesothelioma through association rule mining. *Biomedical Signal Processing and Control*. 2021;68:102726.
42. Hong J, Tamakloe R, Park D. Application of association rules mining algorithm for hazardous materials transportation crashes on expressway. *Accident Analysis & Prevention*. 2020;142:105497.
43. Kim YS, Yum B-J. Recommender system based on click stream data using association rule mining. *Expert Systems with Applications*. 2011;38:13320–7.
44. Lakshmi KS, Vadivu G. Extracting Association Rules from Medical Health Records Using Multi-Criteria Decision Analysis. *Procedia Computer Science*. 2017;115:290–5.
45. Mirhashemi SH, Mirzaei F. Extracting association rules from changes in aquifer drawdown in irrigation areas of Qazvin plain, Iran. *Groundwater for Sustainable Development*. 2021;12:100495.
46. Mukhopadhyay A, Maulik U, Bandyopadhyay S, Coello CAC. A Survey of Multiobjective Evolutionary Algorithms for Data Mining: Part I. *IEEE Transactions on Evolutionary Computation*. 2014;18:4–19.
47. Telikani A, Gandomi AH, Shahbahrami A. A survey of evolutionary computation for association rule mining. *Information Sciences*. 2020;524:318–52.
48. Tseng M-C, Lin W-Y. Efficient mining of generalized association rules with non-uniform minimum support.

Data & Knowledge Engineering. 2007;62:41–64.

49. Gan W, Lin JC-W, Fournier-Viger P, Chao H-C, Zhan J. Mining of frequent patterns with multiple minimum supports. *Engineering Applications of Artificial Intelligence*. 2017;60:83–96.

50. Wu CW, Shie B-E, Tseng VS, Yu PS. Mining top-K high utility itemsets. *Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining - KDD '12* [Internet]. Beijing, China: ACM Press; 2012 [cited 2019 Sep 26]. p. 78. Available from: <http://dl.acm.org/citation.cfm?doid=2339530.2339546>

51. Lee Y-C, Hong T-P, Lin W-Y. Mining association rules with multiple minimum supports using maximum constraints. *International Journal of Approximate Reasoning*. 2005;40:44–54.

52. Borzsony S, Kossmann D, Stocker K. The Skyline operator. *Proceedings 17th International Conference on Data Engineering* [Internet]. Heidelberg, Germany: IEEE Comput. Soc; 2001 [cited 2019 Sep 26]. p. 421–30. Available from: <http://ieeexplore.ieee.org/document/914855/>

53. Viger PF. SPMF: An Open-Source Data Mining Library [Internet]. Available from: <http://www.philippe-fournier-viger.com/spmf/index.php?link=datasets.php>

54. Pei JHJ. Mining Frequent Patterns without Candidate Generation. *Proceedings of the 2000 ACM SIGMOD international conference on Management of data*. Dallas, Texas, USA; 2000. p. 1–12.

55. Hikmawati E, Maulidevi NU, Surendro K. A multi-criteria recommender system model for determining lockdown decision of Covid-19 cases in Indonesia. *Proceedings of the 2021 10th International Conference on Software and Computer Applications*. 2021. p. 9.