ATLANTIS PRESS

Advances in Social Science, Education and Humanities Research, volume 354

2nd International Conference on Applied Science and Technology 2019 - Social Sciences Track (iCASTSS 2019)

Analysis of SP Students using AHP-Apriori Combination

Yeni Rokhayati
Multimedia and Networking Engineering
Politeknik Negeri Batam
Batam, Indonesia
yeni@polibatam.ac.id

Uswatun Hasanah Binti Rusdi
Informatics Engineering
Politeknik Negeri Batam
Batam, Indonesia
uswatunhasanah1896@gmail.com

Dwi Ely Kurniawan
Informatics Engineering
Politeknik Negeri Batam
Batam, Indonesia
dwialikhs@polibatam.ac.id

Nur Zahrati Janah
Informatics Engineering
Politeknik Negeri Batam
Batam, Indonesia
nur.zahrati@polibatam.ac.id

Sudra Irawan
Geometics Engineering
Politeknik Negeri Batam
Batam, Indonesia
sudra@polibatam.ac.id

***Abstract*— Unauthorized absences in lectures causes students to get a warning letter (SP). The SP issued in each semester is numerous. It increases the work of related parties to deal with compensation issues. Thus, the factors that cause students to be absent without permission are interesting material to be explored more deeply, i.e. what factors have the most influence and whether there are relationships among these factors. The required data is collected through interviews and online questionnaires. After that, the factors that cause the absence are ranked to find out which factors are most influential using the AHP method by considering the level of SP (SP1, SP2, and SP3) as the criteria, and the factors as the alternatives. After getting the 5 highest factors, they were analyzed for their relationship using the Apriori method.**

Keywords— AHP, analysis, Apriori, factor, SP students

I. INTRODUCTION

Warning Letter (SP) is a form of warning given to students who have several absences without permission in lecture sessions that exceeds the specified limit. This SP has 3 levels, namely SP1, SP2, and SP3. According to Article 17 of the Politeknik Negeri Batam academic regulation, SP1, SP2, and SP3 will be issued to students who have a number of unauthorized absences exceeding 5%, 20%, and 40% of the total number of sessions in one semester that had planned. Every student who gets this SP is required to undergo compensation duties.

Even though the compensation task has been given according to the SP level, this still does not deter students from being absent without permission in lectures. This is proven by the steady number of SP students in each semester. Based on the Teaching and Learning Implementation (PBM) report of the Informatics Engineering Department from 2015 to 2017, the average student who gets SP1 is 77 students, SP2 is 22 students, and SP3 is 11 students per semester. Consequently, it increases the workload of the department, especially the administration, academic lecturers and other staffs involved in the compensation process. Furthermore, it also hinders the academic performance of the students. So, the factors that cause students to get SP become important and useful to be analyzed.

Many previous researches related to SP analysis have analyzed whether a given warning letter has a positive impact on students or employees, both in terms of discipline and performance. SP can be effectively used as an effort to habituate student discipline and performance [1], [2]. On the other hand, other researches state that it is not effective in improving students or employee's behavior [3], [4]. However, research that examines the causes of a person exposed to the SP still does not exist. Therefore, this research will analyze what factors have the most influence and whether there are relationships among these factors, especially in the Informatics Engineering Department, Politeknik Negeri Batam.

Reference [5] conducted a ranking of insurance companies using AHP and proceeded by analyzing the factors that influence the ranking. While reference [6] analyzed the factors that had a major contribution to the data of severe accidents in China using association rules. Analysis of Mathematics achievement using the two levels of the Bernoulli model, which results in defining factors that have a positive effect and factors that have a negative influence [7]. The other research analyzed the causes of traffic accidents in China, by ranking the factors using AHP, and then determining the level based on the factors using Apriori [8].

This study will rank the influence of the factors causing SP, and then analyze the relationship among the factors. Based on the reference mentioned above, we have chosen the AHP method to rank the factors, followed by the implementation of the Apriori method in analyzing the relationship of these factors.

II. RESEARCH METHODOLOGY

This research has been started with the study of problems, that is by conducting interviews to administration staff, to students, and to counselling sections of the Politeknik Negeri Batam. The interviews intended to find out the basic problems related to SP students, especially about the causal factors. The study continued with literature review to learn more about previous related research, AHP method, Apriori method, and other theories needed. After that, data was collected through online questionnaires to SP students. The result of this data is

processed by the AHP method to get the factor ranking and continued with the implementation of the Apriori method to find out the relationship of these factors. The research process flow is depicted in Fig. 1.

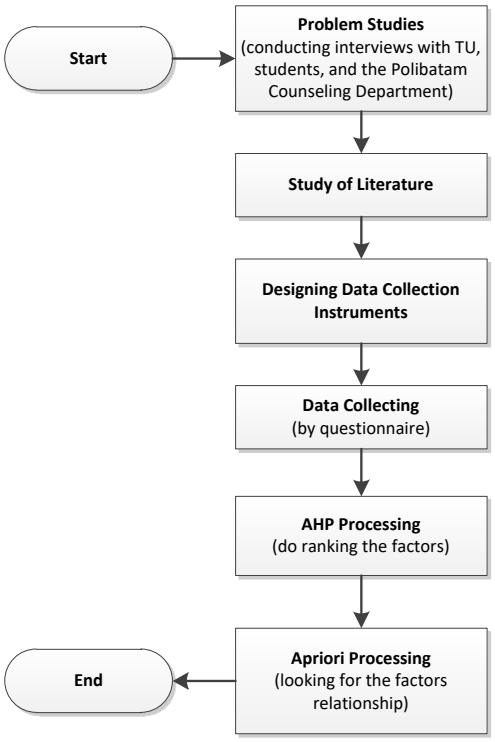

Fig. 1. The research process flow.

A. Design of Data Collection

There are two stages of data collection in this study. The first step was to conduct interviews with the administration staff, the counselling of Politeknik Negeri Batam, lecturers, and SP students. The expected outcome of this interview is to gather their opinions on what factors cause students getting SP. In addition, from the administration staff, a list of SP students will also be obtained to be used as respondents in collecting data using an online questionnaire.

After the factors that caused SP had been collected from the results of the interview, the second step is collecting data with an online questionnaire. The core question asked is about what factors cause them getting SP. The results of this data collection will be the initial data in AHP and Apriori processing.

B. Design of AHP Process

AHP is a ranking method that involves more than one criterion and the criteria are arranged hierarchically [9-12]. Ranking with AHP in this study uses criteria SP1, SP2, and SP3, where SP3 is considered more important than SP2, and SP2 is considered more important than SP1. For example, if matrix A is a comparison matrix of criteria SP1, SP2, and SP3, then matrix A can be written in equation (1).

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \quad (1)$$

Where element a_{ij} is the comparison of SP_i to SP_j, with i, j = 1, 2, and 3.

This comparison value uses numbers 1 to 9 or vice versa; and these comparison figures means:

1: equally important
3: slightly more important
5: more important
7: much more important
9: absolutely more important
2, 4, 6, 8: value in between

After the comparison, the Eigen vector of each criterion is then calculate, which is obtained from the line average of the normalization of matrix A, or symbolized as A_n. The Eigen vector, which is the average row of the A_n matrix, is symbolized by E.

After getting the Eigen vector, the next step is to look at the comparison consistency. Consistency is measured by a consistency ratio (CR), which is a comparison of the consistency index (CI) with a random index of consistency (RI). If CR less then 0.1, then the comparison is said to be consistent [13],[14],[15].

Based on the AHP process shown in Fig. 2, after a consistent comparison of criteria is obtained, it is followed by the alternative comparison process per criteria SP1, SP2, and SP3. The process is like comparing the criteria, that is calculating the Eigen vector, and then proceeding with the calculation of the comparison consistency.

If the criteria and alternatives are compared and known consistent, an H matrix will be obtained as in equation 2, which is an assembly matrix of Eigen vectors.

$$H = \begin{bmatrix} e_{11} & e_{12} & e_{13} \\ e_{21} & e_{22} & e_{23} \\ e_{31} & e_{32} & e_{33} \end{bmatrix} \quad (2)$$

The matrix element e_{ij} states the Eigen vector of alternative i in the criteria j.

This H matrix is then multiplied by the Eigen vector from the comparison of criteria, so that the final score vector is obtained which indicates which alternative has a higher or lower score than the others, and this also tells about the ranking of alternatives.

If consistent, then it will be followed by a comparison of the causes of SP that are used as alternatives. After Eigen vectors are calculated and their consistency is known, then these factors are ranked. The 5 factors with the highest ranking will be used as analysis data for the Apriori process. Completely, the flow of the AHP process can be seen in Fig. 2.

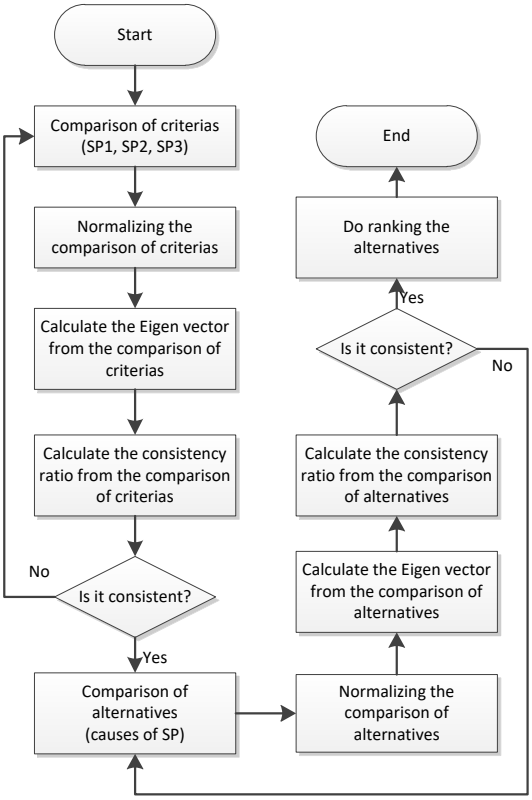

Fig. 2. The flow of AHP process.

C. Design of Apriori Process

The Apriori process begins with recapitulation of the data with the 5 highest factors, those have been obtained previously through the AHP process. Next, determine the k-itemset candidate and calculate the support. Support calculations use the formula in equation 3.

$$\sup(A) = \frac{n(A)}{N} \times 100\% \qquad (3)$$

where:
sup(A) = support of A
$n(A)$ = the number of transactions containing A
N = the number of all transactions

k-itemset whose support value is ≥ minimum support, referred to as frequent k-itemset , and symbolized by Fq_k [16],[17],[18].

After all Fq_k is found, it is combined into the association rules candidates, which then calculated the value of the confidence. Calculation of this confidence value uses the formula in equation 4.

$$Confidence\ A \rightarrow B = \frac{n(A \text{ and } B)}{n(A)} \times 100\% \qquad (4)$$

where:
$n(A \text{ and } B)$= the number of transactions containing A and B
$n(A)$ = the number of transactions containing A

If the confidence value is ≥ confidence minimum, then the candidate will be the association rules, which between the items are interconnected [19],[20]. The flow of this Apriori process can be seen in Fig. 3.

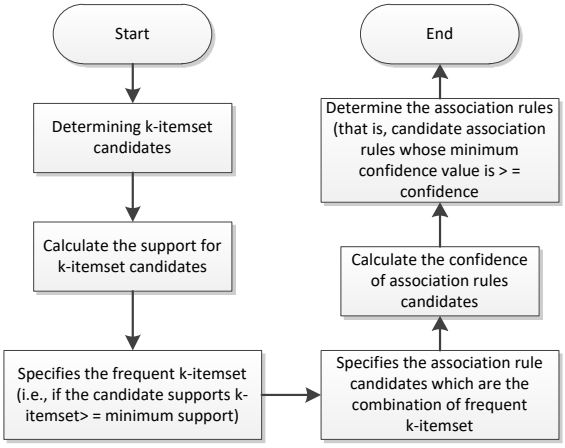

Fig. 3. The flow of Apriori process.

III. Result and Discussion

The things discussed in the results section and this discussion are the result of data collecting, both through interviews and questionnaires, then implementing the ranking process with AHP and implementing analysis with Apriori.

A. Result of Data Collecting

The results of data collecting in this section are divided into 2, that are, the results of data collecting from interviews, and the results of data collecting from online questionnaire.

Interviews that have been conducted to administration staff, counselling department, several lecturers, and several students, it was found that there are 9 factors that often cause students getting SP.

F1. It's hard to get up early
F2. Traffic jam on the way
F3. Lack of interest in some subjects
F4. Dislike some lecturers
F5. There is a family problem
F6. There are some friendship issues
F7. The distance from home to campus is too far
F8. Difficulty finding transportation to campus
F9. Others

In addition, the results of the interview with administration staff also obtained data on the list of Informatics Engineering students who get SP from year 2015 to 2017. There were 70 students. After getting the data, then an online questionnaire was made which was aimed at SP students, to find out the causes of their SP, personally. There were 35 SP students who filled out the questionnaire.

B. Implementation of AHP Process

The data recapitulation in Table I that will be used in the AHP process is a recap of questionnaire, which is more intended for the AHP process.

TABLE I. ALTERNATIVE DATA BASED ON CRITERIA

		Criteria		
		SP1	**SP2**	**SP3**
Alternative	**F1**	13	3	0
	F2	12	2	0
	F3	7	1	0
	F4	7	1	0
	F5	3	2	0
	F6	4	1	0
	F7	11	2	0
	F8	3	1	0
	F9	11	2	0

The implementation of the AHP process in this sub-section writes a complete comparison of criteria including the comparison matrix, normalization, calculation of the Eigen vector and its consistency ratio. Whereas alternative comparisons are the same as the comparison of criteria, so that only the matrix H is written which is a collection of Eigen vectors.

The comparison criteria matrix SP1, SP2, and SP3 are written in equation 5, normalized in equation 6, and the Eigen vectors in equation 7.

$$A = \begin{bmatrix} 1 & 0.33 & 0.2 \\ 3 & 1 & 0.33 \\ 5 & 3 & 1 \end{bmatrix} \quad (5)$$

$$A_N = \begin{bmatrix} 0.111 & 0.077 & 0.130 \\ 0.333 & 0.231 & 0.217 \\ 0.556 & 0.692 & 0.652 \end{bmatrix} \quad (6)$$

$$E = \begin{bmatrix} 0.106 \\ 0.261 \\ 0.633 \end{bmatrix} \quad (7)$$

The result of the calculation $\lambda_{max} = 3.055$, CI=0.028, so the obtained ratio is obtained, that is CR = 0.048. The value CR≤0.1, therefore the comparison of the criteria is consistent.

The three matrix of alternative comparison for criteria SP1, SP2, and SP3 are normalized, then their consistency is searched, and finally the Eigen vectors are collected in the H matrix as written in equation 8.

$$H = \begin{bmatrix} 0.18 & 0.20 & 0.11 \\ 0.17 & 0.13 & 0.11 \\ 0.10 & 0.07 & 0.11 \\ 0.10 & 0.07 & 0.11 \\ 0.04 & 0.13 & 0.11 \\ 0.06 & 0.07 & 0.11 \\ 0.15 & 0.13 & 0.11 \\ 0.04 & 0.07 & 0.11 \\ 0.15 & 0.13 & 0.11 \end{bmatrix} \quad (8)$$

After the matrix H is multiplied by the Eigen vector in equation 8, the final score is obtained, then it can be ranked as shown in Table II.

TABLE II. THE FINAL SCORE AND RANKING

	Final Score	**Ranking**
F1	0.14191	1
F2	0.12305	2
F3	0.09820	6
F4	0.09820	7
F5	0.10959	5
F6	0.09372	8
F7	0.12155	3
F8	0.09222	9
F9	0.12155	4

Looking at the results in Table II, the 5 factors with the highest rank in a row are F1, F2, F7, F9, and F5. After this is known, then the next step is analyzing the relationship between these factors using the Apriori method.

C. Implementation of Apriori Process

Through the results of factor ranking with AHP, it was found that the 5 highest factors causing SP were F1, F2, F5, F7, and F9. As preliminary data for Apriori process, from 35 data of respondents, only those with the highest factors were analyzed. There are 5 candidates 1-itemsets, namely F1, F2, F5, F7, and F9. This support is calculated first, and the results are in Table III.

TABLE III. 1-ITEMSET CANDIDATE AND THE SUPPORT

1-Itemset Candidate	**Support**
F1	46.00%
F2	43.00%
F5	11.00%
F7	39.00%
F9	39.00%

In calculating the causes of SP1 using a minimum support of 15%, which means an item is said to often appear if more than 4 times of the total transaction 28, because 15% of 28 is 4.2.

So, the frequent 1-itemset (Fq_1) are {F1}, {F2}, {F7}, dan {F9}; where the support value is more than the minimum support. The 2-itemset candidate is a combination of (Fq_1), where the support calculation presented in Table IV.

TABLE IV. 2-ITEMSET CANDIDATE AND THE SUPPORT

2-Itemset Candidate	**Support**
F1, F2	18.00%
F1, F7	18.00%
F1, F9	10.71%
F2, F7	28.57%
F2, F9	10.71%
F7, F9	7.14%

Looking at the support values in table 4, it can be determined that what becomes Fq_2, that is {F1, F2}, {F1, F7}, and {F2, F7}. Furthermore, the 3-itemset candidate which combination of Fq_2 is only 1, namely {F1, F2, F7}, where the value of the support is 14.28%. This does not exceed the minimum support value. So, nothing become Fq_3. The same process is applied to the causes of SP2 in Table V.

TABLE V. CAUSES OF SP2

ID	Causes SP
006	F1, F9
024	F9
027	F1, F2, F5
031	F7, F9
033	F1, F2, F5, F7

Using a minimum support of 50%, obtained that Fq_1 is {F1} and {F9}; where there is no Fq_2 dan Fq_3.

The process is continued with the calculation of confidence from the candidate association rules formed from a combination of Fq. Table 6 is the result of confidence calculation for candidate association rules of Fq_1 and Fq_2, both from the causes of SP1 and SP2.

TABLE VI. CANDIDATE OF RULE ASSOCIATION AND THE CONFIDENCE

No	Candidate of Rule Association	*Confidence*
1	F1 → SP1	100%
2	F2 → SP1	100%
3	F7 → SP1	100%
4	F9 → SP1	100%
5	F1 → F2	38.5%
6	F2 → F1	41.7%
7	F1 → F7	38.5%
8	F7 → F1	45.5%
9	F2 → F7	66.7%
10	F7 → F2	72.7%
11	F1, F2 → SP1	100%
11	F1, F7 → SP1	100%
12	F2, F7 → SP1	100%
13	F1 → SP2	100%
14	F9 → SP2	100%

Using a minimum confidence of 50%, the association rules are 9. They are F1 → SP1, F2 → SP1, F7 → SP1, F9 → SP1, F2 → F7, F7 → F2, F1, F2 → SP1, F1, F7 → SP1, F2, F7 → SP1, F1 → SP2, dan F9 → SP2. The meaning of these rules will be written in conclusion, in the second paragraph.

IV. CONCLUSION

Data of factors that cause students getting SP have been found through administration interviews, lecturers, students, and counselling sections. There are 9 factors. They are: difficulty getting up early (F1), traffic jams (F2), lack of interest in some subjects (F3), disliking some lecturers (F4), family problems (F5), friendship issues (F6), the distance from home to campus is too far (F7), difficulty finding transportation to campus (F8), and others (F9).

This factor was analyzed in two stages. The first is to rank these factors in influencing students to get SP using the AHP method, which results in sequential order starting from the highest ranks are F1, F2, F7, F9, F5, F3, F4, F6, and F8. The 5 factors with the highest ranking were further analyzed in the second stage, that is to find out the relationship between influencing students to get SP. The result can be stated that:

- If students have difficulty getting up early (F1), or don't like some subjects (F2), or have difficulty finding transportation to campus (F8), or others (F9), students tend to get SP1.
- If students have difficulty getting up early (F1), or others (F9), students will tend to get SP2.
- Apparently, if the distance of the house to campus is too far (F7), then students tend to experience traffic jams on the trip (F2), and vice versa.

This will help the academic lecturer in anticipating or giving advice to students so as not to be exposed to SP. For example, if there are students who are indicated to have these 5 highest factors, then immediately call their students or even parents to change their behavior that will lead to getting SP.

ACKNOWLEDGMENT

Our main gratitude goes to our institution, Batam State Polytechnic, for all the support and facilities that have been given to complete this research.

REFERENCES

[1] B. G. Moss and W. H. Yeaton, "Failed warnings: Evaluating the impact of academic probation warning letters on student achievement," Evaluation Review, vol. 39, no. 5, pp. 501-524, 2015.

[2] J. A. Taylor and J. Lawrence, "Making students aware: An online strategy for students given academic warning," Studies in Learning, Evaluation, Innovation and Development, vol. 4, no. 2, pp. 39-52, 2007.

[3] W. V. Epperson and R. M. Harano, "An evaluation of some additional factors influencing the effectiveness of warning letters," Accident Analysis & Prevention, vol. 7, no. 4, pp. 239-247, 1975.

[4] D. Shatin, J. S. Gardner, A. Stergachis, D. Blough, and D. Graham, "Impact of mailed warning to prescribers on the co-prescription of tramadol and antidepressants," Pharmacoepidemiology and Drug Safety, vol. 14, no. 3, pp. 149-154, 2005.

[5] M. K. Valahzaghard and M. Ferdousnejhad, "Ranking insurance firms using AHP and factor analysis," Management Science Letters, vol. 3, no. 3, pp. 937-942, 2013.

[6] C. Xu, J. Bao, C. Wang, and P. Liu, "Association rule analysis of factors contributing to extraordinarily severe traffic crashes in China," Journal of Safety Research, vol. 67, pp. 65-75, 2018.

[7] I. Demir and S. Kilic, "Using PISA 2003, Examining the Factors Affecting Student's Mathematics Achievement," Hacettepe Üniversitesi Eğitim Fakültesi Dergisi, vol. 38, no. 38, pp. 44-45, 2010.

[8] J. Xi, Z. Zhao, W. Li and Q. Wang, "A Traffic Accident Causation Analysis Method Based on AHP-Apriori," Procedia Engineering, vol. 137, pp. 680-687, 2016.

[9] T. L. Saaty, "How to make a decision: The Analytic Hierarchy Process," European Journal of Operational Research, vol. 48, no. 1, pp. 9-26, 1990.

[10] T. L. Saaty, "Decision making—the analytic hierarchy and network processes (AHP/ANP)," Journal of systems science and systems engineering, vol. 13, no. 1, pp. 1-35, 2004.

[11] C. Qiansheng, "Analytic Hierarchy Process (AHP) and Attribute Hierarchical Model (AHM)," Systems Engineering---Theory & Practice, vol. 11, pp. 56-59, 1997.

[12] M. I. A. Khalil, "Selecting the appropriate project delivery method using AHP," International journal of project management, vol. 20, no. 6, pp. 469-474, 2002.

[13] T. L. Saaty, "Decision-making with the AHP: Why is the principal eigenvector is necessary," European Journal of Operational Research, vol. 145, no. 1, pp. 85-91, 2003.

[14] Z. Xu, "On consistency of the weighted geometric mean complex judgement matrix in AHP," European Journal of Operational Research, vol. 126, no. 3, pp. 683-687, 2000.

[15] E. F. Lane and W. A. Ferdini, "A consistency test for AHP decision makers," Decision Sciences, vol. 20, no. 3, pp. 575-590, 1989.

[16] C. Borgelt and R. Kruse, "Induction of association rules: Apriori implementation," Compstat, pp. 395-400, 2002.

[17] S. Kotsiantis and D. Kanellopoulos, "Association rules mining: A recent overview," GESTS International Transactions on Computer Science and Engineering, vol. 32, no. 1, pp. 71-82, 2006.

[18] N. Pasquier, Y. Bastide, R. Taouil and L. Lakhal, "Discovering frequent closed itemsets for association rules," in International Conference on Database Theory, 1999.

[19] M. Y. Lin, P. Y. Li and S. C. Hsueh, "Apriori-based frequent itemset mining algorithms on MapReduce," in Proceedings of the 6th international conference on ubiquitous information management and communication, 2012.

[20] M. H. Margahny and A. Shakour, "Fast algorithms for mining association rules," in The conference proceedings of AIML, 2005.