The association rules search of Indonesian university graduate’s data using FP-growth algorithm

S Faza, R F Rahmat*, E B Nababan, D Arisandi, S Effendi
Faculty of Computer Science and Information Technology, Universitas Sumatera Utara, Medan, Indonesia

*E-mail: romi.fadillah@usu.ac.id

Abstract. The attribute varieties in university graduates data have caused frustrations to the institution in finding the combinations of attributes that often emerge and have high integration between attributes. Association rules mining is a data mining technique to determine the integration of the data or the way of a data set affects another set of data. By way of explanation, there are possibilities in finding the integration of data on a large scale. Frequent Pattern-Growth (FP-Growth) algorithm is one of the association rules mining technique to determine a frequent itemset in an FP-Tree data set. From the research on the search of university graduate’s association rules, it can be concluded that the most common attributes that have high integration between them are in the combination of State-owned High School outside Medan, regular university entrance exam, GPA of 3.00 to 3.49 and over 4-year-long study duration.

1. Introduction
Association rules mining is a data mining technique to determine the integration of the data or the way of a data set affects another set of data [1]. By way of explanation, there are possibilities in finding the integration of data on a large scale.

University graduates data is one of the large-scale data of an educational institution consisting of identity data of the students that were collected during the freshman year and academic data that were collected throughout the years in the university. From the existing graduates’ data, some attributes can deduct student’s length of studies such as previous school, entrance exam types, faculties, and GPA.

The attribute varieties in university graduate’s data, e.g., previous school, entrance exam types, faculties, and GPA, have caused frustrations to the institution in finding the combinations of attributes that often emerge and have high integration between attributes of university graduate’s data. Therefore, it is necessary to have an alternative to obtain the combination of common attributes that have a high integration between attributes using association rules mining. Association rules mining can generate high-integrated attributes combination by searching for attribute combinations that often appear first.

Frequent Pattern-Growth (FP-Growth) Algorithm is an algorithm that can be used in association rules mining technique to determine the frequent itemset in a set of data [2]. The algorithm only performs the database scanning process twice to determine the frequent itemset [3] and represents the transaction using FP-Tree data structure [4]. The result is expected to generate numerous rules that have high integration on the attributes in university graduates’ data.
Research on university graduates has been performed using various methods. In 2011, Siregar analysed the integration determinant between student and length of study data using multiple linear regression algorithms [5]. This research generates a result at 61% of the integration of student’s data toward study period data. Parameters set as the integration determinants are GPA, average national exams score, amount of credit hours, and parents background education. The research produces a prediction without yet knowing which combination of parameters which frequently emerge.

Huda, in 2010, used Apriori algorithm to show the information of student graduation rates [6]. The result of data mining process in the research can be applied as consideration in taking further decisions on the factors that affect the graduation rates, especially the one in student’s primary data of a faculty. The student’s primary data in the process of mining include data admission process, data from school, student city data, and study program data. The research advocates future researcher to use FP-Growth algorithm since the proposed method required numerous iteration of searches to the database.

In 2013, Ma’ruf also used Apriori algorithm to determine the relationship of admission process with student graduation rates [7]. The research generates information on the origin of student admission achievement which set as a reference in maximizing the advertisement in particular area. Apriori algorithm performed candidate generate of each item, so it required many iterations of searches to the database.

The attribute varieties in university graduate’s data, e.g., previous school, entrance exam types, faculties, and GPA, have caused frustrations to the institution in finding rules that consist of the combinations of attributes that often emerge and have high integration between attributes. Therefore, an alternative is required to generate association rules for university graduate’s data.

2. Methodology
The proposed method to search the association rules on university graduate’s data consists of several phases, i.e., data pre-processing by determining the attributes, namely school types, university entrance types, GPA, and length of study, then classifying the attributes into several groups. The next phase is data mining process to obtain association rules. Association rules can be obtained by two steps [3], which are the search of the frequent itemset using FP-Growth algorithm then followed by the creation of association rules to search the integration of each item in an item set.

In determining association rules, there is an interestingness measure obtained from the data processing with specific calculations [3], namely support as the supporting value and confidence as an absolute value. Both will be applied to determine the interesting association rules, by comparing the threshold specified by the user. The limit usually contains minimum support and confidence. According to [8], association rules mining aims to determine all rules that have the value of support >= min_support and confidence >= min_confidence.

Prediction matching will be performed based on the obtained association rules to generate suggestion as a consideration for the faculties and universities to take further decisions about factors affecting the passing rate. The general architecture of the proposed method is shown in figure 1.

![Figure 1. General architecture of the proposed method](image)
2.1. Data pre-processing

Data pre-processing is a data selection process that aims to obtain the data being used in this research by eliminating some attributes that will not be analyzed since it is not suitable to be the factor of student length of study.

2.1.1. Attributes determination

In university graduates data, there are several attributes such as student ID, alumnus ID, name, gender, place, and date of birth, religion, city, zip code, previous school, faculty, major, university entrance types, GPA, length of study, and graduation period. The attributes that will be analyzed are school types, faculties, university entrance types, GPA and length of study, while the rest of the attributes will not be analyzed.

There are various types of formal education institutions taken before entering universities such as state or privately owned high schools which located in the city or outside the city. University entrance types are varied to enter the university, as for the first type is only entitled to those who have excellent grades throughout high school years. GPA (Grade Point Average) is the accumulated score obtained by students from each semester. Length of study is the time spent to finish the bachelor degree. The standard time for a student to graduate in four years. However, numerous students spent over 4 years to achieve the bachelor degree, yet the number of students who spend less than 4 years to obtain the degree is quite numerous.

2.1.2. Attributes grouping

There are several item groups in each attribute, which are:

1. School types
   1.1. State-owned high school (HS)/Islamic high school (IHS) in Medan
   1.2. State-owned HS/ IHS outside Medan
   1.3. Private-owned HS/ IHS in Medan
   1.4. Private-owned HS/ IHS outside Medan
   1.5. Other HS/ IHS

2. University entrance types
   2.1. High school scores based entry (HSSBE)
   2.2. Regular country-subsidized entry (RCSE)
   2.3. Independent entry (IE)
   2.4. Independent country-subsidized entry (ICSE)

3. GPA
   3.1. GPA < 3.00
   3.2. GPA = 3.00-3.49
   3.3. GPA > 3.49

4. Length of study
   4.1. < 4 years
   4.2. 4 years
   4.3. > 4 years

2.2. Data processing

There are several phases in processing the data to generate the desired data, which is a combination of rules or association rules of the university graduates.

2.2.1. Data mining

Data mining is a process of collecting and using the historical data to find the regularity of patterns and relationships in a large dataset [9]. This phase is performed to obtain association rules. Below in Table 1 is the list of 30 university graduates data that will be used to search the association rules.
| Student ID | School type                        | Entrance type | GPA     | Length of study |
|-----------|------------------------------------|---------------|---------|-----------------|
| 1         | State-owned HS/ IHS outside Medan  | RCSE          | 2.97    | > 4 Years       |
| 2         | Private-owned HS/ IHS in Medan     | RCSE          | 2.92    | > 4 Years       |
| 3         | State-owned HS/ IHS outside Medan  | RCSE          | 3.29    | 4 Years         |
| 4         | State-owned HS/ IHS in Medan       | RCSE          | 3.41    | < 4 Years       |
| 5         | State-owned HS/ IHS in Medan       | IE            | 2.92    | < 4 Years       |
| 6         | State-owned HS/ IHS in Medan       | RCSE          | 2.78    | > 4 Years       |
| 7         | State-owned HS/ IHS in Medan       | IE            | 3.05    | < 4 Years       |
| 8         | State-owned HS/ IHS in Medan       | RCSE          | 3.22    | < 4 Years       |
| 9         | State-owned HS/ IHS outside Medan  | RCSE          | 3.30    | < 4 Years       |
| 10        | State-owned HS/ IHS outside Medan  | RCSE          | 3.10    | < 4 Years       |
| 11        | State-owned HS/ IHS outside Medan  | RCSE          | 3.36    | < 4 Years       |
| 12        | Private-owned HS/ IHS in Medan     | RCSE          | 3.07    | > 4 Years       |
| 13        | State-owned HS/ IHS outside Medan  | RCSE          | 3.02    | > 4 Years       |
| 14        | State-owned HS/ IHS in Medan       | IE            | 3.06    | < 4 Years       |
| 15        | State-owned HS/ IHS in Medan       | RCSE          | 3.29    | < 4 Years       |
| 16        | State-owned HS/ IHS outside Medan  | RCSE          | 3.44    | < 4 Years       |
| 17        | State-owned HS/ IHS outside Medan  | RCSE          | 3.24    | < 4 Years       |
| 18        | State-owned HS/ IHS in Medan       | RCSE          | 3.08    | < 4 Years       |
| 19        | Private-owned HS/ IHS in Medan     | RCSE          | 3.03    | > 4 Years       |
| 20        | State-owned HS/ IHS in Medan       | RCSE          | 3.46    | 4 Years         |
| 21        | State-owned HS/ IHS in Medan       | IE            | 2.97    | > 4 Years       |
| 22        | Other HS/ IHS                     | RCSE          | 3.59    | 4 Years         |
| 23        | State-owned HS/ IHS in Medan       | IE            | 2.79    | > 4 Years       |
| 24        | State-owned HS/ IHS outside Medan  | RCSE          | 2.95    | > 4 Years       |
| 25        | Private-owned HS/ IHS in Medan     | RCSE          | 3.02    | 4 Years         |
| 26        | State-owned HS/ IHS outside Medan  | RCSE          | 2.99    | 4 Years         |
| 27        | Private-owned HS/ IHS outside Medan| RCSE          | 3.04    | 4 Years         |
| 28        | State-owned HS/ IHS outside Medan  | RCSE          | 3.54    | 4 Years         |
| 29        | Private-owned HS/ IHS in Medan     | RCSE          | 2.90    | > 4 Years       |
| 30        | State-owned HS/ IHS outside Medan  | RCSE          | 3.06    | 4 Years         |

According to Table 1, 30 university graduates have unique attribute groups individually. Therefore, each attribute item is given a code to simplify the search process rules as in table 2.

| Attribute         | Item                              | Code |
|-------------------|-----------------------------------|------|
| School types      | State-owned HS/ IHS in Medan      | A    |
|                   | State-owned HS/ IHS outside Medan | B    |
|                   | Private-owned HS/ IHS in Medan    | C    |
|                   | Private-owned HS/ IHS outside Medan| D   |
|                   | Other HS/ IHS                     | E    |
| Entrance types    | HSSBE                              | F    |
|                   | RCSE                              | G    |
|                   | IE                                | H    |
|                   | ICSE                              | I    |
| GPA               | GPA < 3.00                        | J    |
|                   | GPA = 3.00-3.49                   | K    |
|                   | GPA > 3.49                        | L    |
Length of study

| < 4 Years | M |
|-----------|---|
| 4 Years   | N |
| > 4 Years | O |

Based on table 2, the item set of university graduates data or code based item set can be seen in table 3.

| Student ID | Item set | Student ID | Item set | Student ID | Item set |
|------------|----------|------------|----------|------------|----------|
| 1          | BGJO     | 11         | BGKM     | 21         | AHJO     |
| 2          | CGJO     | 12         | EGKO     | 22         | EGLN     |
| 3          | BGKN     | 13         | BGKO     | 23         | AHJO     |
| 4          | AGKM     | 14         | AHKM     | 24         | BGJO     |
| 5          | AHJM     | 15         | AGKM     | 25         | CGKN     |
| 6          | AGJO     | 16         | BGKM     | 26         | BGJN     |
| 7          | AHKO     | 17         | BGKM     | 27         | DGKN     |
| 8          | AGKM     | 18         | AGKM     | 28         | BGLN     |
| 9          | BGKM     | 19         | CGKO     | 29         | CGJO     |
| 10         | BGKM     | 20         | AGKN     | 30         | BGKN     |

This stage has the purpose of searching the item combinations that meet the minimum requirements of the support value in the database. Support value of each item can be obtained using equation 1 below:

\[
Support(A) = \frac{\text{Total Transaction } A}{\text{Total Transaction}} \times 100\% \quad (1)
\]

Meanwhile, the support value of 2 items can be retrieved using the equation below:

\[
Support(A, B) = \frac{\text{Total Transaction } A \text{ and } B}{\text{Total Transaction}} \times 100\% \quad (2)
\]

Based on Table 3, the frequent itemset can be determined, suppose given a minimum value of support = 10% and then search the support value of each item with the equation 1 as in table 4.

Step 1: Finding frequent item

| No | Code | Total | Support     |
|----|------|-------|-------------|
| 1  | A    | 11    | 11/30 * 100% = 36% |
| 2  | B    | 12    | 12/30 * 100% = 40% |
| 3  | C    | 4     | 4/30 * 100% = 13% |
| 4  | D    | 1     | 1/30 * 100% = 3% |
| 5  | E    | 2     | 2/30 * 100% = 6% |
| 6  | F    | 0     | 0/30 * 100% = 0% |
| 7  | G    | 25    | 25/30 * 100% = 83% |
| 8  | H    | 5     | 5/30 * 100% = 16% |
| 9  | I    | 0     | 0/30 * 100% = 0% |
Based on the data in table 4, the common items that meet the minimum support value of 10% are school types A, B, and C, entrance types G and H, GPA J and K, and length of study M, N, and O.

Step 2: Building FP-Tree

FP-Tree is a data storage structure formed by a null-labeled root [3]. FP-Tree constructed consists of frequent items on every student, so the rare items are omitted. University graduates data with dataset can be seen in table 5.

**Table 5. University graduates item set data**

| Students ID | Item set | Students ID | Item set | Students ID | Item set |
|-------------|----------|-------------|----------|-------------|----------|
| 1           | BGJO     | 11          | BGKM     | 21          | AHJO     |
| 2           | CGJO     | 12          | GKO      | 22          | GN       |
| 3           | BGKN     | 13          | BGKO     | 23          | AHJO     |
| 4           | AGKM     | 14          | AHKM     | 24          | BGJO     |
| 5           | AHJM     | 15          | AGKM     | 25          | CGKN     |
| 6           | AGJO     | 16          | BGKM     | 26          | BGJN     |
| 7           | AHKO     | 17          | BGKM     | 27          | GKN      |
| 8           | AGKM     | 18          | AGKM     | 28          | BGN      |
| 9           | BGKM     | 19          | CGKO     | 29          | CGJO     |
| 10          | BGKM     | 20          | AGKN     | 30          | BGJO     |

FP-Tree that can be constructed from the existing item set can be seen in figure 2.

![FP-Tree Diagram](image-url)
Figure 2. (a) Tree after reading itemset-1 (BGJO), (b) Tree after reading itemset-2 (CGJO), (c) Tree after reading itemset-3 (BGKN), and (d) Tree after reading itemset-30 (BGKN).

Step 3: Searching frequent item set

Frequent itemset search is conducted through several stages [10].

1. Conditional FP-Tree evocation, which only contains the same item suffix on each item set as in figure 3.

   (a)                                (b)

   (c)

Figure 3. Conditional FP-Tree (a) suffix M, (b) suffix N, (C) suffix (O).

2. Checking whether the support value of the item set suffix is higher or less than the minimum support of 10% by the equation (1), if it meets the requirements, the second check will be performed on the second suffix continuously until the prefix using equation (2).

From the calculation of frequent itemset search obtained AGKM and BGKM item set that meets the minimum support as in table 6.

| No | Suffix | Total | Support     |
|----|--------|-------|-------------|
| 1  | AGKM   | 4     | 4/30 * 100% = 13% |
| 2  | BGKM   | 5     | 5/30 * 100% = 16% |
2.3. Creating association rules

Creating association rules aims to find the associative rule \( A \rightarrow B \) that meets the minimum requirement of confidence value, i.e., to look for a connection between items in an item set. The confidence value is obtained using equation 3 below:

\[
\text{Confidence}_{A \rightarrow B} = \frac{\text{Total Transaction containing } A \& B}{\text{Total Transaction containing } A} \times 100\%
\] (3)

This research looks at the relations between a student length of study to school types, entrance types, and GPA. Suppose that given a 60% confidence value, the search for confidence can be calculated using the equation (3) and can be seen in table 7.

| Table 7. Confidence value |
|---------------------------|
| No | Item set | \( A \rightarrow B \) | Confidence |
|----|---------|-----------------|------------|
| 1  | AGKM   | \( AGK \rightarrow M \) | 5/8 * 100% = 62% |
| 2  | BGKM   | \( BGK \rightarrow M \) | 4/5 * 100% = 80% |

Based on Table 7, AGKM and BGKM item sets have confidence values that meet the minimum confidence value. Therefore, AGKM and BGKM can be used as the rules of a university graduates data, which have a high degree of emergence and connectivity in a university graduates data.

Association rules obtained are:
1. Students originated from State-owned HS/IHS in Medan with entrance type RCSE and GPA range of 3.00 s/d 3.50, will spend <4 years to obtain the degree.
2. Students originated from State-owned HS/IHS outside Medan with entrance type RCSE and GPA range of 3.00 s/d 3.50, will spend <4 years to obtain the degree.

2.4. Suggestion Matching

Suggestion matching can for university or faculty be conducted based on the obtained association rules to see the suffix on the rules that are by the length of study. The advice given can be useful for universities to accelerate student’s length of study, or to reduce the student's rate with over 4 years-long studies.

Suggestion matching can be identified from the obtained rules based on the length of study, if the length of study is < 4 years or = 4 years, then the initial two combinations of the obtained rules, namely school, and entrance types. It can be selected as the suggestion for the faculty. The faculty can provide information to prospective students from particular schools using obtained rules for school types. Faculty can also offer numerous opportunities to enter the faculty to prospective students with particular entrance types that based on the obtained rules.

If the length of study is > 4 years, the rules suggested the faculty to avert or limit the promotion to students with school types specified by the rules, as for the entrance type rules, it is wise for the faculty to limit the opportunities for the prospective students with specified entrance types to enter the faculty.

Suggestion matching also considers the GPA from the obtained rules. However, it is not in accordance with the length of study attribute. If GPA is lower than 3.00, then it is implied that faculty needs to give extra attention and guidance to the students about the courses to be taken for the thesis. If the GPA range between 3.00 to 3.49 that means faculty should guide the students to improve their grades while those who have GPA above 3.50 should be advised to maintain their excellent work by the faculty.

3. Experimental Results

The test of the association rules search was conducted in every faculty. It generated rules with support value range of 1 to 23% depends on the faculty. The smaller the support value due to numerous
attribute combinations, the less the combination appearance rate of a combination. Meanwhile, the confidence value with the highest support value in all faculties, has an average value at over 70%. It is implied that the obtained association rules have high integration between each attribute. The test result of the highest support value can be seen in Table 8.

Table 8. Highest support value

| Faculty          | Item set | Highest support value | Confidence value |
|------------------|----------|-----------------------|------------------|
| Medical          | BGKM     | 6,12                  | 98,59            |
| Dentistry        | CGJO     | 5,9                   | 57,81            |
| Pharmacy         | CGJO     | 11,83                 | 96,88            |
| Public health    | BGKO     | 8,98                  | 64,13            |
| Nursery          | BGKM     | 11,42                 | 85,29            |
| Mathematics & science | BGKO | 10,76                 | 80,73            |
| Computer science | AGKO     | 12,01                 | 92,42            |
| Engineering      | BGKO     | 14,37                 | 91,8             |
| Psychology       | AGJO     | 8,21                  | 97,06            |
| Agriculture      | BGJO     | 12,13                 | 99,29            |
| Law              | BGKO     | 5,62                  | 69,9             |
| Cultural literature | BGKO | 9,7                   | 77,1             |
| Economy          | BGKM     | 7,17                  | 58,04            |
| Social and politics | BGKO | 8,08                  | 57,14            |

According to Table 8, the faculty of Medical, Nursery and Economy attained BGKM item set (State-owned HS/ IHS outside Medan, RCSE, GPA range of 3.00 to 3.49, length of study < 4 years), while Dentistry and Pharmacy faculties obtained CGJO item set (Private-owned HS/ IHS in Medan, RCSE, GPA < 3.00, length of study > 4 years). The faculty of Public Health, Mathematics and Science, Engineering, Law, Cultural Literature and Social Politics obtained BGKO item set (State-owned HS/ IHS outside Medan, RCSE, GPA range of 3.00 to 3.49, length of study > 4 years). AGKO item set (State-owned HS/ IHS in Medan, RCSE, GPA range of 3.00 to 3.49, length of study > 4 years) were acquired for Computer science faculty and Psychology obtained AGJO item set (State-owned HS/ IHS in Medan, RCSE, GPA < 3.00, length of study < 4 years). Lastly is faculty of Agriculture which attained BGJO item set (State-owned HS/ IHS outside Medan, RCSE, GPA < 3.00, length of study < 4 Years).

Based on the item set, the system can deliver recommendations to faculty or university. Below is the suggestion for the faculty of Medical, Nursery, and Economy with BGKM item set:

a. The three faculties should provide information or persuade the prospective students originated from State-owned HS/ IHS outside Medan to choose them.

b. Universities should offer numerous opportunities to the prospective students with RCSE entrance type to enter the three faculties.

c. These faculties should guide the students who have GPA range between 3.00 to 3.49 to improve their grades and to start thinking and deepen the courses to be taken for the thesis.

Suggestions for Dentistry and Pharmacy faculties with CGJO item set are:

a. These faculties should provide information or persuade the prospective students originated from Private-owned HS/ IHS in Medan to choose them.

b. Universities should offer numerous opportunities to the prospective students with RCSE entrance type to enter these faculties.
c. These faculties should give extra attention to the students who have GPA lower than 3.00 improve their grades and to start thinking and deepen the courses to be taken for the thesis.

These are the suggestion obtained for Public Health, Mathematics and Science, Engineering, Law, Cultural Literature and Social Politics faculties with BGKO item set:

a. These faculties can provide information or persuade the prospective students originated from all high school types excluding the State-owned HS/ IHS outside Medan to choose them.

b. Universities should limit the available opportunities to the students with RCSE entrance type to enter these faculties.

c. These faculties should guide the students who have GPA range between 3.00 to 3.49 to improve their grades and to start thinking and deepen the courses to be taken for the thesis.

Computer Sciences faculty with AGKO item set is presented with suggestions, which are:

a. Computer Sciences faculty should provide information or persuade the prospective students originated from State-owned HS/ IHS in Medan to choose the faculty.

b. Universities should offer numerous opportunities to the prospective students with RCSE entrance type to enter the faculty.

c. This faculty should guide the students who have GPA range between 3.00 to 3.49 to improve their grades and to start thinking and deepen the courses to be taken for the thesis.

Suggestions for Psychology faculty with AGJO items set are:

a. Psychology faculty should provide information or persuade the prospective students originated from all high school types excluding the State-owned HS/ IHS in Medan to choose the faculty.

b. Universities should offer numerous opportunities to the prospective students with RCSE entrance type to enter the faculty.

c. This faculty should give extra attention to the students who have GPA lower than 3.00 improve their grades and to start thinking and deepen the courses to be taken for the thesis.

At last is the suggestions for Agriculture faculty with BGJO item set which are:

a. Agriculture faculty should provide information or persuade the prospective students originated from all high school types excluding the State-owned HS/ IHS outside Medan to choose the faculty.

b. Universities should offer numerous opportunities to the prospective students with RCSE entrance type to enter the faculty.

c. This faculty should give extra attention to the students who have GPA lower than 3.00 improve their grades and to start thinking and deepen the courses to be taken for the thesis.

4. Conclusion and Future Research

Based on the result of the association rules search process using the FP-Growth algorithm, it can be concluded that the use of the FP-Growth algorithm to search association rules can generate association rules using minimum support and confidence values as the reference values. The most frequent attribute combination to appear which has high inter-attribute connectivity in all faculties is school type of State-owned HS/ IHS outside Medan, RCSE entrance types, GPA range of 3.00 to 3.49, and length of study > 4 Years.

From the result, the faculties can determine which school types that are suitable to promote the faculties to the prospective student. Moreover, the university can see which entrance type that has a significant impact on passing rate to determine the number of accepted students proposed by the faculties.

For further research, authors are expected to predict the student’s length of study using prediction method to support the result of this research. It can be done using Evolving Connectionist Method that has been implemented in [11] with fast learning [12] or using Distributed Adaptive Engine [13-15]. Further research can also add more parameters, such as national high school exams score and high school major which can be used as the factor of student’s study period.
5. References

[1] Said A M, Dominic P D D and Abdullah A B 2009 A Comparative Study of FP-growth Variations. *IJCSNS International Journal of Computer Science and Network Security* 9

[2] Yongmei L and Yong Guan 2008 FP-Growth Algorithm for Application in Research of Market Basket Analysis. *IEEE 6th International Conference on Computational Cybernetics* Stara Lesna Slovakia

[3] Han J and Kamber M 2006 *Data Mining: Concepts and Techniques* Second Edition (Elsevier: USA)

[4] Han J, Pei J, and Yin Y 2000 Mining Frequent Patterns without Candidate Generation. *Proceedings of the ACM SIGMOD International Conference on Management of Data* pp 1-12

[5] Siregar A S 2011 Application Development Of Interrelationship Determinant Between Student Data And Length Of Study Using Multiple Linear Regression Algorithm (STT Harapan Medan Students Case Study) *Thesis* Universitas Sumatera Utara: Medan

[6] Huda N M 2010 Data Mining Application To Show Information of Students Graduation Rates. *Thesis* Universitas Diponegoro: Semarang

[7] Ma’ruf F A 2013 Data Mining Application To Identify The Relationship Of Students Graduation Rates and Admission Type (Case Study: STMIK AMIKOM Yogyakarta) *Journal STMIK AMIKOM*: Yogyakarta.

[8] Avadh K S, Ajeet K, Ashish K 2014 An Empirical Analysis and Comparison of Apriori and FP-Growth Algorithm for Frequent Pattern Mining. *IEEE International Conference on Advanced Communication Control and Computing Technologies (ICACCT)*

[9] Pel W A, Chunhong A and Lei W 2014 An Improved Algorithm For Mining Association Rule In Relational Database. *Proceedings of the 2014 International Conference on Machine Learning and Cybernetics*

[10] Tan P N, Steinbach M and Kumar V 2005 *Introduction to Data Mining* (Addison-Wesley: United States)

[11] Rahmat R F, Rizki A, Alharthi A F and Budiarto R 2016 Big Data Forecasting Using Evolving Multi-layer Perceptron *4th Saudi International Conference on Information Technology* (Big Data Analysis) (KACSTIT) pp 1–5

[12] Alwajeeh T, Alharthi A F, Rahmat R F and Budiarto R 2017 Fast Learning for Big Data Using Dynamic Function *IOP Conference Series: Materials Science and Engineering* 190 p 12015

[13] Rahmat R, Pasha M F, Syukur M, Budiarto R 2015 A gene-regulated nested neural network. *International Arab Journal of Information Technology* 12 pp 532–539

[14] Pasha M F, Rahmat R F, Budiarto R, and Syukur M 2010 A distributed autonomous neuro-gen learning engine and its application to the lattice analysis of cubic structure identification problem. *International Journal of Innovative Computing, Information and Control* 6 pp 1005–1022

[15] Aaron, Sitompul O S and Rahmat R F 2014 Distributed autonomous Neuro-Gen Learning Engine for content-based document file type identification. *International Conference on Cyber and IT Service Management* pp 63–68