# HASIL CEK_Zahrotun, Jones_Library, Association Rule, FP-Growth, study program.

*by* Lisna Zahrotun, Anna Hendri Soleliza Jones Fp-growth Algorithm For Searching

**Submission date:** 20-Sep-2022 01:01PM (UTC+0700)
**Submission ID:** 1904328551
**File name:** orrowing_Transactions_PAtterns_And_Study_Program_Suitability.pdf (661.81K)
**Word count:** 2546
**Character count:** 13351

International Journal of Information System & Technology
Akreditasi No. 158/E/KPT/2021 | Vol. 5, No. 5, (2022), pp. 564-569

# Fp-Growth Algorithm For Searching Book Borrowing Transaction Patterns And Study Program Suitability

*Lisna Zahrotun[1], Anna Hendri Soleliza Jones[2]*

*Department of Informatic, Faculty of Industrial Engineering, Universitas Ahmad Dahlan*
*Ringroad Selatan, Kampus 4 UAD, Yogyakarta 55191*

*e-mail: lisna.zahrotun@tif.uad.ac.id[1*], annahendri@tif.uad.ac.id[2]*

***Abstrak***

*The current development of data has reached a sizeable amount. This is due to the development of the world of information technology which consists of data in it. One technique that can handle abundant data is data mining. Data mining methods are widely used to perform large amounts of data analysis. In the academic field, analysis can be used to determine the patterns of students and lecturers. Whereas in library transactions, analysis can be carried out to determine the patterns of existing book borrowing. This is done to determine the tendency of students with certain study programs to borrow any uku transactions. In this study, the aim of this research is to analyze the patterns of borrowing books from the Ahmad Dahlan University library, which includes borrowing transaction data and the book owner's study program. In addition, in this study, a percentage analysis of the suitability of the book borrower study program and the book owner's study program was also carried out. The stages in this research include data collection, data cleaning, data selection, data transformation, searching for association patterns using the FP-Growth method and pattern evaluation. The test used in this research is the lift ratio. The results of this study are publications in international journals that are in the draft process. Apart from that, the results of this study provide information on the analysis of patterns of lending books in libraries using the FP-Growth method. The resulting pattern is 103 patterns with a support count value of 5 and a confident 10% with the 2 itemset rule, this means that the level of book borrowing is still low. While the results of the analysis of the suitability of books in the study program with the borrower were 31% in accordance with the study program, namely Pharmacy and Public Health Sciences, meaning that there were 69% of students who borrowed books from the library that were not in accordance with their study program.*

***Keywords***: *Library, Association Rule, FP-Growth, study program.*

## 1. Introduction

The library is a collection of information that can be accessed by the public to find references or borrow materials / information. The information provided by the library can be physical or digital. One of the functions of the library is a place for borrowing books which is a provider of information and knowledge. The quality of the library can be seen from how well the book lending and borrowing transaction function is carried out. Based on library interviews, a university in Yogyakarta also has more than 24,397 collections of book titles totaling 67,453 copies. Transactions that occur in the library every day reach 100 transactions, all of which have been recorded in the system. However, the large number of books means that there are also many books that are not actively borrowed by students. Therefore, it takes a step to find book borrowing patterns to determine the level of student habits in borrowing books. Apart from that every days reached 100 transactions which have all been recorded in the system. However, the large number of books means that there are also many books that are not actively borrowed by students. Therefore, it

ISSN : 2580-7250

takes a step to find book borrowing patterns to determine the level of student habits in borrowing books. In addition, each study program recommends procurement of books for each year. However, this available book has never been analyzed the absorption of its borrowing. Therefore it is important to know the absorption of borrowed books based on the borrower's study program and the book owner's study program. In previous research, we have conducted research related to data mining, namely searching for patterns of alumni data associations [1], drug data patterns [2], grouping customer data [3]. Apart from data mining, research on text mining is also carried out [4][5][6].

One way to solve the problems in the library is by implementing the Association rule. association rule is a data mining technique used to find associative rules between a combination of items. Frequent Pattern Growth (FP-Growth) is an a priori algorithm that can be used to determine the most frequent itemsets in a set of data. FP-Growth is faster than the Apriori algorithm [7]. By using the FP-Tree concept, the FP Growth algorithm also has a good ability to look for association patterns [8]. Several studies have been conducted to find out patterns of lending library books. As research conducted by [9] and on other studies conducted [10]. Both studies have been carried out and both are still using the a priori algorithm, which is basically the process of calculating the frequent itemset is still relatively slow, so this study has determined the eclat algorithm so that the frequent itemset calculation process is faster. In addition, research conducted at the library has not implemented the suitability of study program books with borrower study programs and still uses book IDs for pattern search calculations.

## 2. Research Methodology

### 2.1. Data Mining

Data mining is a process of finding important knowledge from large collections of data that have been stored in databases [11]. Such knowledge cannot be found in simple ways. Data mining uses certain techniques (methods) to find knowledge including classification, clustering and linkage analysis. According to [12] Data mining tasks can be grouped into two categories namely:

a) Predictive mining is concerned with classification methods, regression and deviation detection.
b) Descriptive mining will derive and investigate patterns (correlation, trend, cluster) of important relationships from data sets. The main method of descriptive mining is clustering, association rules and sequential mining.

### 2.2. FP-Growth

FP-Growth is an improvement on the a priori algorithm. This algorithm is used to define the most frequent itemset in a data set [13]. The FP-Growth algorithm uses the concept of a development tree in searching for frequent itemsets. The characteristic of the FP-Growth algorithm is that the data structure used is a tree called the FP-Tree. By using the FP-Tree, the FP-Growth algorithm can directly extract itemset from the FP-Tree. Excavation of frequent itemset using the FP-Growth algorithm will be carried out by generating a data structure tree or called the FP-Tree [14][15].

### 2.3. Factor of Interest

Support is an occurrence rule for all transactions in the dataset [11]. Minimum support Is a value that must be met by the rules:

$$Support\ (A \rightarrow B) = P(A \cup B) = \frac{\sum A \cup B}{\sum Transaction} \quad (1)$$

### 2.4. Confidence

Confidence is the opportunity for an item to appear together with other items to appear. If it is formed with notation in A → B, then trust is defined as how often B occurs

when A also occurs. Minimum Confidence value is defined as the minimum to be fulfilled by the rules.

$$Confidence\ (A \rightarrow B) = P(A|B) = \frac{P(A \cup B)}{P(A)} = \frac{support\ count\ (A \cup B)}{support\ count\ (A)} \quad (2)$$

### 2.5. Lift Rasio

Lift Ratio is used to find out the correlation between the items in the rule. If the value of the lift rule is> 1 then Positively correlated, if the value of the lift rule <1 then negatively correlated, if the value of the lift rule = 1 then it's independent (not combined).

$$Lift(|A \rightarrow B|) = P(A|B) = \frac{P(A \cup B)}{P(A).P(B)} = \frac{confidence\ (A \cup B)}{support\ (A)} \quad (3)$$

### 2.6. Frekuent Pattern Growth

FP-Growth is one of the algorithms used to solve the Association Rule case. This algorithm has two steps. The first two steps, compression is performed on the database based on frequently occurring items to create a Frequent Pattern Tree (FP-Tree). Second, separation is performed on the database results into a compressed database condition. FP-Tree is a special feature that distinguishes the FP-Growth algorithm from the a priori algorithm. FP-Tree has two features, first starting from the root which is named "null". Then from the root to form a sub-tree consisting of certain items. Second, each node contains three important information, namely the item label (indicates the type of item (item ID) it is represented by the node), support (the number shows the number of transaction paths through that node or also known as frequency), connecting pointers (link nodes) as a liaison. Between vertices and items in each path between cells, marked with a line connecting the pointer arrows.

## 3. Result and Discussion

The results of this study are the association pattern and the percentage of suitability of book borrowers with the book owner's study program at a university library. Before the pattern search is carried out, the first step taken is taking the data of the library book lending transaction that will be processed. The data that has been successfully displayed is then carried out by a process of selecting and transforming the data. The results of the selection and data transformation then carried out a pattern search using the FP Growth method. After finding several patterns, the next step the borrower's suitability pattern and graph with the book owner's program of study. An image of the research is to check the correlation of the patterns produced using the lift ratio. And the final step is to display the results of steps is shown in Figure 1.

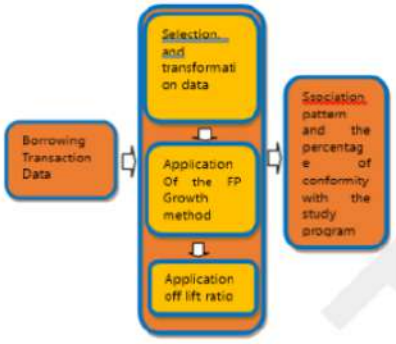


**Figure 1.** Research Diagram

### 3.1. Implementation

The first step is data retrieval which is the initial process by entering an excel file in the form of initial book loan transaction data. The transaction data table can be seen in Figure 2.

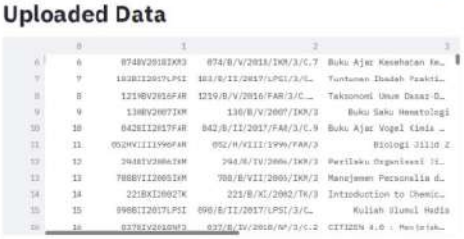


**Figure 2.** Load Data

The next stage, namely data selection is carried out before the information mining stage. The selected data is stored in the file. at this stage remove unused variables. The data that has been successfully selected can be seen in Figure 3.

**Figure 3.** Selected Data

The next stage, namely this transformation is carried out by changing the data variable into valid or ready for the data mining process. The results can be seen in Figure 4


**Figure 4.** Data Transformation

The final step is Result FP-growth to display the results of the association rules where there is an input form to input minimum confidence and minimum support with the minimum confidence 0.7 and minimum support conditions 8. The results of the FP-growth result can be seen in Figure 5.

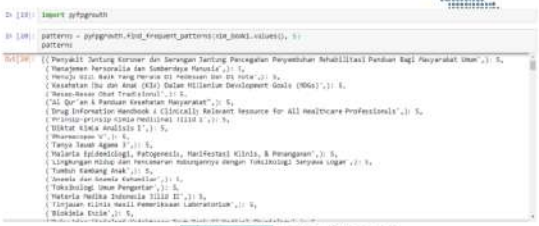

**Figure 5.** FP-Growth

### 3.2. Result

From the association patterns obtained, namely 103 patterns, the patterns formed from lending transactions are usually dominated by sustainable or similar books, for example in rule nos. 102 and 103, namely if you borrow morals lecture books, you also borrow morals lectures. Another example is rule no.95, which is if you borrow a book on organic biology 3 and biology, then you borrow an organic biology book 2. In addition to the search for association patterns, a suitability analysis is also carried out between the borrower's study program and the book owner's study program. The results of the analysis are shown in Figure 6. and Figure 7.

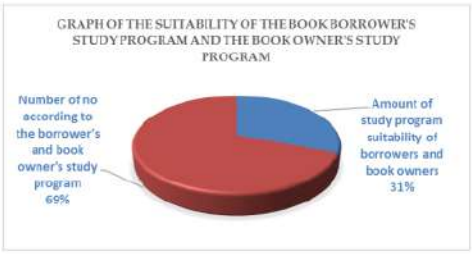


**Figure 6.** Suitability Results of the Study Program

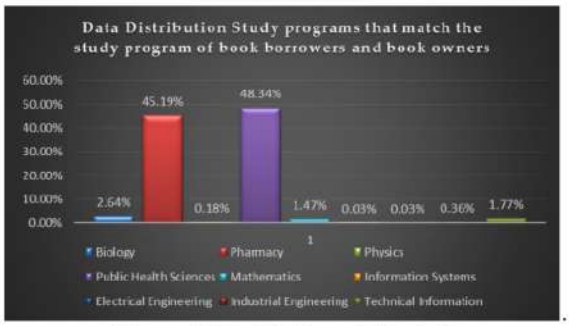


**Figure 7.** Distribution of Study Programs

## 4. Conclusion

A program has been made with the python programming language that is able to produce rules which can be used by librarians to determine the distribution of library book lending and the distribution of the suitability of the study program. The resulting pattern is 103 patterns with a support count value of 5 and a confident 10%, this means that the level of book borrowing is still low. From the association pattern obtained, the pattern formed from lending transactions is usually dominated by books that are sustainable or similar, for example in rules nos. 102 and 103, namely if you borrow morals lecture books, you also borrow lectures on morals. Another example is rule no.95, that is, if you

borrow a book on organic biology 3 and biology, then you borrow an organic biology book 2.The results of the analysis of the suitability of the book in the study program with the borrower are 31% that is appropriate, meaning that 69% of students borrow books from the library are not according to the study program. From the appropriate data, it is dominated by the Pharmacy and Public Health Sciences study programs.

## References

[1] H. Yuliansyah and L. Zahrotun, "Designing web-based data mining applications to analyze the association rules tracer study at university using a FOLD-growth method," *Int. J. Advanced Comput. Res.*, vol. 6, no. 27, pp. 215–221, 2016.

[2] L. Zahrotun, D. Soyusiawaty, R. S. Pattihua, and U. A. Dahlan, "The Implementation of Data Mining for Association Patterns Determination Using Temporal Association Methods in Medicine Data," pp. 6–11, 2016.

[3] L. Zahrotun, "Implementation of Data Mining Technique for Customer Relationship Management ( CRM ) on Online Shop Tokodiapers . com With Fuzzy C-Means Clustering," in *2017 2nd International Conferenceson Information TEchnology, Information System, and Electrical Engeenering (ICITISEE) yogyakarta, Indonesia*, 2017, pp. 298–302.

[4] L. Zahrotun, N. hutami Putri, and A. N. Khusna, "The Implementation of K-Means Clustering Method in Classifying Undergraduate Thesisi Titles," in *12th International Conference on Telecommunication Systems, Services, and Applications (TSSA)*, 2018.

[5] L. Zahrotun, "Text Mining for Internship Titles Clustering Using Shared Nearest-Neighbor Method," *Comput. Eng. Appl.*, vol. 6, no. 3, 2017.

[6] L. Zahrotun, "Comparison Jaccard similarity , Cosine Similarity and Combined Both of the Data Clustering With Shared Nearest Neighbor Method," *Comput. Eng. Appl.*, vol. 5, no. 1, pp. 11–18, 2016.

[7] M. Kavitha and M. S. T. T. Selvi, "Comparative Study on Apriori Algorithm and Fp Growth Algorithm with Pros and Cons," *Int. journa Comput. Sci. Trends Technol. (IJCST*, vol. 4, no. 4, pp. 161–164, 2016.

[8] C. M. Velu, M. Ramakrishnan, V. Somu, and P. Loganathan, P. Vivekanandan, "Efficiency Association Rule for Data Mining," *Int. J. Soft Comput.*, vol. 2, no. 1, pp. 21–36, 2007.

[9] naufal afif, "Association Rule Mining Peminjaman Buku Perpustakaan Universitas Ahmad Dahlan Menggunakan Algoritma Apriori," 2017.

[10] G. F. Mandias *et al.*, "Analisa Pola Peminjaman Buku di Pepustakaan Universitas Klabat Menggunakan Algoritma Apriori," *Konf. Nas. Sist. Inf.*, pp. 8–9, 2018.

[11] J.Han, Kamber.M, Jian Pie, *Data Mining: Concepts and Techniques*. 2011.

[12] T. Slimani, "Efficient Analysis of Pattern and Association Rule Mining Approaches," *Coll. Comput. Sci. Inf. Technol. , Taif Univ. , KSA*.

[13] T. Andi and E. Utami, "Association rule algorithm with FP growth for book search," in *3rd Annual Applied Science and Engineering Conference (AASEC 2018)*, 2018, pp. 1–6.

[14] F. Feng, J. Cho, W. Pedrycz, H. Fujita, and T. Herawan, "Soft Set Based Association Rule Mining," *Knowledge-Based Sist.*, vol. 111, pp. 268–282, 2016.

[15] C. Angeli, S. K. Howard, M. Jun, J. Yang, and P. A. Kirschner, "Data mining in educational technology classroom research : Can it make a contribution ?," *Comput. Educ.*, vol. 113, pp. 226–242, 2017.

# HASIL CEK_Zahrotun, Jones_Library, Association Rule, FP-Growth, study program.

ORIGINALITY REPORT

| 10% | 8% | 8% | 5% |
|---|---|---|---|
| SIMILARITY INDEX | INTERNET SOURCES | PUBLICATIONS | STUDENT PAPERS |

PRIMARY SOURCES

| | | |
|---|---|---|
| 1 | ojs.stmikpringsewu.ac.id<br>Internet Source | 4% |
| 2 | Frederick F. Patacsil. "Analysis of Cyberbullying Incidence among Filipina Victims: A Pattern Recognition using Association Rule Extraction", International Journal of Intelligent Systems and Applications, 2019<br>Publication | 1% |
| 3 | A A Hidayat, A Rahman, R M Wangi, R J Abidin, R S Fuadi, W Budiawan. "Implementation and comparison analysis of apriori and fp-growth algorithm performance to determine market basket analysis in Breiliant shop", Journal of Physics: Conference Series, 2019<br>Publication | 1% |
| 4 | scitepress.org<br>Internet Source | 1% |
| 5 | uniassignment.com<br>Internet Source | 1% |

| | | |
|---|---|---|
| 6 | www.iaeme.com<br>Internet Source | 1 % |
| 7 | jurnal-ppi.kominfo.go.id<br>Internet Source | 1 % |
| 8 | doaj.org<br>Internet Source | 1 % |

| | | | |
|---|---|---|---|
| Exclude quotes | On | Exclude matches | < 1% |
| Exclude bibliography | On | | |

International Journal of Information System & Technology
Akreditasi No. 158/E/KPT/2021 | Vol. 5, No. 5, (2022), pp. 564-569

# Fp-Growth Algorithm For Searching Book Borrowing Transaction Patterns And Study Program Suitability

*Lisna Zahrotun[1], Anna Hendri Soleliza Jones[2]*

*Department of Informatic, Faculty of Industrial Engineering, Universitas Ahmad Dahlan*
*Ringroad Selatan, Kampus 4 UAD, Yogyakarta 55191*

*e-mail: lisna.zahrotun@tif.uad.ac.id[1*], annahendri@tif.uad.ac.id[2]*

***Abstrak***

*The current development of data has reached a sizeable amount. This is due to the development of the world of information technology which consists of data in it. One technique that can handle abundant data is data mining. Data mining methods are widely used to perform large amounts of data analysis. In the academic field, analysis can be used to determine the patterns of students and lecturers. Whereas in library transactions, analysis can be carried out to determine the patterns of existing book borrowing. This is done to determine the tendency of students with certain study programs to borrow any uku transactions. In this study, the aim of this research is to analyze the patterns of borrowing books from the Ahmad Dahlan University library, which includes borrowing transaction data and the book owner's study program. In addition, in this study, a percentage analysis of the suitability of the book borrower study program and the book owner's study program was also carried out. The stages in this research include data collection, data cleaning, data selection, data transformation, searching for association patterns using the FP-Growth method and pattern evaluation. The test used in this research is the lift ratio. The results of this study are publications in international journals that are in the draft process. Apart from that, the results of this study provide information on the analysis of patterns of lending books in libraries using the FP-Growth method. The resulting pattern is 103 patterns with a support count value of 5 and a confident 10% with the 2 itemset rule, this means that the level of book borrowing is still low. While the results of the analysis of the suitability of books in the study program with the borrower were 31% in accordance with the study program, namely Pharmacy and Public Health Sciences, meaning that there were 69% of students who borrowed books from the library that were not in accordance with their study program.*

***Keywords***: *Library, Association Rule, FP-Growth, study program.*

## 1. Introduction

The library is a collection of information that can be accessed by the public to find references or borrow materials / information. The information provided by the library can be physical or digital. One of the functions of the library is a place for borrowing books which is a provider of information and knowledge. The quality of the library can be seen from how well the book lending and borrowing transaction function is carried out. Based on library interviews, a university in Yogyakarta also has more than 24,397 collections of book titles totaling 67,453 copies. Transactions that occur in the library every day reach 100 transactions, all of which have been recorded in the system. However, the large number of books means that there are also many books that are not actively borrowed by students. Therefore, it takes a step to find book borrowing patterns to determine the level of student habits in borrowing books. Apart from that every days reached 100 transactions which have all been recorded in the system. However, the large number of books means that there are also many books that are not actively borrowed by students. Therefore, it

ISSN : 2580-7250

takes a step to find book borrowing patterns to determine the level of student habits in borrowing books. In addition, each study program recommends procurement of books for each year. However, this available book has never been analyzed the absorption of its borrowing. Therefore it is important to know the absorption of borrowed books based on the borrower's study program and the book owner's study program. In previous research, we have conducted research related to data mining, namely searching for patterns of alumni data associations [1], drug data patterns [2], grouping customer data [3]. Apart from data mining, research on text mining is also carried out [4][5][6].

One way to solve the problems in the library is by implementing the Association rule. association rule is a data mining technique used to find associative rules between a combination of items. Frequent Pattern Growth (FP-Growth) is an a priori algorithm that can be used to determine the most frequent itemsets in a set of data. FP-Growth is faster than the Apriori algorithm [7]. By using the FP-Tree concept, the FP Growth algorithm also has a good ability to look for association patterns [8]. Several studies have been conducted to find out patterns of lending library books. As research conducted by [9] and on other studies conducted [10]. Both studies have been carried out and both are still using the a priori algorithm, which is basically the process of calculating the frequent itemset is still relatively slow, so this study has determined the eclat algorithm so that the frequent itemset calculation process is faster. In addition, research conducted at the library has not implemented the suitability of study program books with borrower study programs and still uses book IDs for pattern search calculations.

## 2. Research Methodology

### 2.1. Data Mining

Data mining is a process of finding important knowledge from large collections of data that have been stored in databases [11]. Such knowledge cannot be found in simple ways. Data mining uses certain techniques (methods) to find knowledge including classification, clustering and linkage analysis. According to [12] Data mining tasks can be grouped into two categories namely:

a) Predictive mining is concerned with classification methods, regression and deviation detection.
b) Descriptive mining will derive and investigate patterns (correlation, trend, cluster) of important relationships from data sets. The main method of descriptive mining is clustering, association rules and sequential mining.

### 2.2. FP-Growth

FP-Growth is an improvement on the a priori algorithm. This algorithm is used to define the most frequent itemset in a data set [13]. The FP-Growth algorithm uses the concept of a development tree in searching for frequent itemsets. The characteristic of the FP-Growth algorithm is that the data structure used is a tree called the FP-Tree. By using the FP-Tree, the FP-Growth algorithm can directly extract itemset from the FP-Tree. Excavation of frequent itemset using the FP-Growth algorithm will be carried out by generating a data structure tree or called the FP-Tree [14][15].

### 2.3. Factor of Interest

Support is an occurrence rule for all transactions in the dataset [11]. Minimum support Is a value that must be met by the rules:

$$Support\ (A \rightarrow B) = P(A \cup B) = \frac{\sum A \cup B}{\sum Transaction} \quad (1)$$

### 2.4. Confidence

Confidence is the opportunity for an item to appear together with other items to appear. If it is formed with notation in A → B, then trust is defined as how often B occurs

when A also occurs. Minimum Confidence value is defined as the minimum to be fulfilled by the rules.

$$Confidence\ (A \rightarrow B) = P(A|B) = \frac{P\ (A \cup B)}{P(A)} = \frac{support\ count\ (A \cup B)}{support\ count\ (A)} \quad (2)$$

### 2.5. Lift Rasio

Lift Ratio is used to find out the correlation between the items in the rule. If the value of the lift rule is> 1 then Positively correlated, if the value of the lift rule <1 then negatively correlated, if the value of the lift rule = 1 then it's independent (not combined).

$$Lift(|A \rightarrow B|) = P(A|B) = \frac{P\ (A \cup B)}{P(A).P(B)} = \frac{confidence\ (A \cup B)}{support\ (A)} \quad (3)$$

### 2.6. Frekuent Pattern Growth

**F**P-Growth is one of the algorithms used to solve the Association Rule case. This algorithm has two steps. The first two steps, compression is performed on the database based on frequently occurring items to create a Frequent Pattern Tree (FP-Tree). Second, separation is performed on the database results into a compressed database condition. FP-Tree is a special feature that distinguishes the FP-Growth algorithm from the a priori algorithm. FP-Tree has two features, first starting from the root which is named "null". Then from the root to form a sub-tree consisting of certain items. Second, each node contains three important information, namely the item label (indicates the type of item (item ID) it is represented by the node), support (the number shows the number of transaction paths through that node or also known as frequency), connecting pointers (link nodes) as a liaison. Between vertices and items in each path between cells, marked with a line connecting the pointer arrows.

## 3. Result and Discussion

The results of this study are the association pattern and the percentage of suitability of book borrowers with the book owner's study program at a university library. Before the pattern search is carried out, the first step taken is taking the data of the library book lending transaction that will be processed. The data that has been successfully displayed is then carried out by a process of selecting and transforming the data. The results of the selection and data transformation then carried out a pattern search using the FP Growth method. After finding several patterns, the next step the borrower's suitability pattern and graph with the book owner's program of study. An image of the research is to check the correlation of the patterns produced using the lift ratio. And the final step is to display the results of steps is shown in Figure 1.

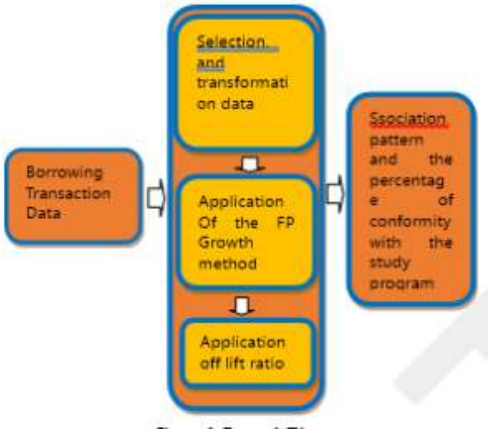


**Figure 1.** Research Diagram

### 3.1. Implementation

The first step is data retrieval which is the initial process by entering an excel file in the form of initial book loan transaction data. The transaction data table can be seen in Figure 2.

**Uploaded Data**

| | 0 | 1 | 2 | 3 |
|---|---|---|---|---|
| 6 | 6 | 0748V2018IKM3 | 074/B/V/2018/IKM/3/C.7 | Buku Ajar Kesehatan Ke… |
| 7 | 7 | 103BII2017LPSI | 103/B/II/2017/LPSI/3/C… | Tuntunan Ibadah Prakti… |
| 8 | 8 | 1219BV2016FAR | 1219/B/V/2016/FAR/3/C.… | Taksonomi Umum Dasar-D… |
| 9 | 9 | 138BV2007IKM | 138/B/V/2007/IKM/3 | Buku Saku Hematologi |
| 10 | 10 | 042BII2017FAR | 042/B/II/2017/FAR/3/C.9 | Buku Ajar Vogel Kimia … |
| 11 | 11 | 052HVIII1996FAR | 052/H/VIII/1996/FAR/3 | Biologi Jilid 2 |
| 12 | 12 | 294BIV2006IKM | 294/B/IV/2006/IKM/3 | Perilaku Organisasi Ji… |
| 13 | 13 | 708BVII2005IKM | 708/B/VII/2005/IKM/3 | Manajemen Personalia d… |
| 14 | 14 | 221BXI2002TK | 221/B/XI/2002/TK/3 | Introduction to Chemic… |
| 15 | 15 | 098BII2017LPSI | 098/B/II/2017/LPSI/3/C… | Kuliah Ulumul Hadis |
| 16 | 16 | 037BIV2018NP3 | 037/B/IV/2018/NP/3/C.2 | CITIZEN 4.0 : Menjejak… |

**Figure 2.** Load Data

The next stage, namely data selection is carried out before the information mining stage. The selected data is stored in the file. at this stage remove unused variables. The data that has been successfully selected can be seen in Figure 3.

**Figure 3.** Selected Data

The next stage, namely this transformation is carried out by changing the data variable into valid or ready for the data mining process. The results can be seen in Figure 4

**Data Transformation**

| | Inventaris | Prodi |
|---|---|---|
| 0 | Farmasi | Farmasi |
| 1 | Ilmu Kesehatan Masyarakat | Manajemen Pendidikan S2 |
| 2 | Ilmu Kesehatan Masyarakat | Ilmu Kesehatan Masyarakat |
| 3 | Ilmu Kesehatan Masyarakat | Ilmu Kesehatan Masyarakat |
| 4 | Ilmu Kesehatan Masyarakat | Pasca Sarjana Psikologi Sains |
| 5 | Ilmu Kesehatan Masyarakat | Pasca Sarjana Psikologi Sains |
| 6 | Lembaga Pengembangan Studi Islam | Biologi |
| 7 | Farmasi | Biologi |
| 8 | Ilmu Kesehatan Masyarakat | Biologi |
| 9 | Farmasi | Teknik Kimia - S1 |
| 10 | Farmasi | Pend. Biologi |

**Figure 4.** Data Transformation

The final step is Result FP-growth to display the results of the association rules where there is an input form to input minimum confidence and minimum support with the minimum confidence 0.7 and minimum support conditions 8. The results of the FP-growth result can be seen in Figure 5.

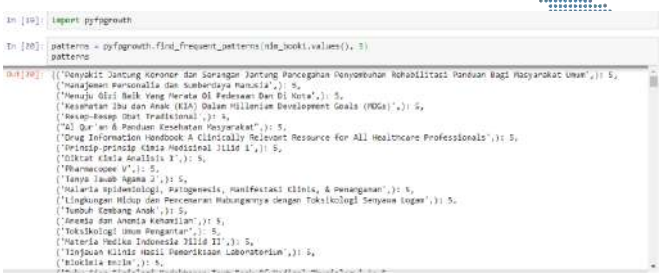

**Figure 5.** FP-Growth

### 3.2. Result

From the association patterns obtained, namely 103 patterns, the patterns formed from lending transactions are usually dominated by sustainable or similar books, for example in rule nos. 102 and 103, namely if you borrow morals lecture books, you also borrow morals lectures. Another example is rule no.95, which is if you borrow a book on organic biology 3 and biology, then you borrow an organic biology book 2. In addition to the search for association patterns, a suitability analysis is also carried out between the borrower's study program and the book owner's study program. The results of the analysis are shown in Figure 6. and Figure 7.

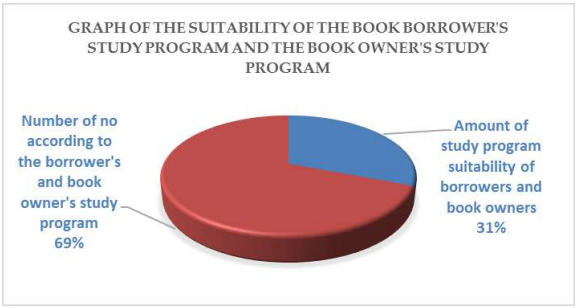


**Figure 6.** Suitability Results of the Study Program

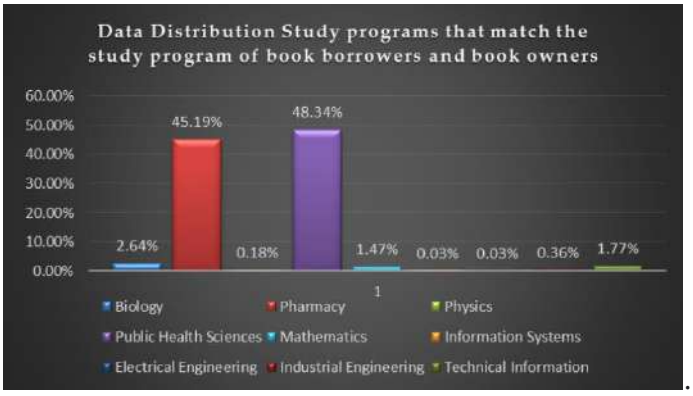


**Figure 7.** Distribution of Study Programs

## 4. Conclusion

A program has been made with the python programming language that is able to produce rules which can be used by librarians to determine the distribution of library book lending and the distribution of the suitability of the study program. The resulting pattern is 103 patterns with a support count value of 5 and a confident 10%, this means that the level of book borrowing is still low. From the association pattern obtained, the pattern formed from lending transactions is usually dominated by books that are sustainable or similar, for example in rules nos. 102 and 103, namely if you borrow morals lecture books, you also borrow lectures on morals. Another example is rule no.95, that is, if you

borrow a book on organic biology 3 and biology, then you borrow an organic biology book 2.The results of the analysis of the suitability of the book in the study program with the borrower are 31% that is appropriate, meaning that 69% of students borrow books from the library are not according to the study program. From the appropriate data, it is dominated by the Pharmacy and Public Health Sciences study programs.

## References

[1] H. Yuliansyah and L. Zahrotun, "Designing web-based data mining applications to analyze the association rules tracer study at university using a FOLD-growth method," *Int. J. Adnvanced Comput. Res.*, vol. 6, no. 27, pp. 215–221, 2016.

[2] L. Zahrotun, D. Soyusiawaty, R. S. Pattihua, and U. A. Dahlan, "The Implementation of Data Mining for Association Patterns Determination Using Temporal Association Methods in Medicine Data," pp. 6–11, 2016.

[3] L. Zahrotun, "Implementation of Data Mining Technique for Customer Relationship Management ( CRM ) on Online Shop Tokodiapers . com With Fuzzy C-Means Clustering," in *2017 2nd International Conferenceson Information TEchnology, Information System, and Electrical Engeenering (ICITISEE) yogyakarta, Indonesia*, 2017, pp. 298–302.

[4] L. Zahrotun, N. hutami Putri, and A. N. Khusna, "The Implementation of K-Means Clustering Method in Classifying Undergraduate Thesisi Titles," in *12th International Conference on Telecommunication Systems, Services, and Applications (TSSA)*, 2018.

[5] L. Zahrotun, "Text Mining for Internship Titles Clustering Using Shared Nearest-Neighbor Method," *Comput. Eng. Appl.*, vol. 6, no. 3, 2017.

[6] L. Zahrotun, "Comparison Jaccard similarity , Cosine Similarity and Combined Both of the Data Clustering With Shared Nearest Neighbor Method," *Comput. Eng. Appl.*, vol. 5, no. 1, pp. 11–18, 2016.

[7] M. Kavitha and M. S. T. T. Selvi, "Comparative Study on Apriori Algorithm and Fp Growth Algorithm with Pros and Cons," *Int. journa Comput. Sci. Trends Technol. (IJCST*, vol. 4, no. 4, pp. 161–164, 2016.

[8] C. M. Velu, M. Ramakrishnan, V. Somu, and P. Loganathan, P. Vivekanandan, "Efficiency Association Rule for Data Mining," *Int. J. Soft Comput.*, vol. 2, no. 1, pp. 21–36, 2007.

[9] naufal afif, "Association Rule Mining Peminjaman Buku Perpustakaan Universitas Ahmad Dahlan Menggunakan Algoritma Apriori," 2017.

[10] G. F. Mandias *et al.*, "Analisa Pola Peminjaman Buku di Pepustakaan Universitas Klabat Menggunakan Algoritma Apriori," *Konf. Nas. Sist. Inf.*, pp. 8–9, 2018.

[11] J.Han, Kamber.M,.Jian Pie, *Data Mining: Concepts and Techniques*. 2011.

[12] T. Slimani, "Efficient Analysis of Pattern and Association Rule Mining Approaches," *Coll. Comput. Sci. Inf. Technol. , Taif Univ. , KSA*.

[13] T. Andi and E. Utami, "Association rule algorithm with FP growth for book search," in *3rd Annual Applied Science and Engineering Conference (AASEC 2018)*, 2018, pp. 1–6.

[14] F. Feng, J. Cho, W. Pedrycz, H. Fujita, and T. Herawan, "Soft Set Based Association Rule Mining," *Knowledge-Based Sist.*, vol. 111, pp. 268–282, 2016.

[15] C. Angeli, S. K. Howard, M. Jun, J. Yang, and P. A. Kirschner, "Data mining in educational technology classroom research : Can it make a contribution ?," *Comput. Educ.*, vol. 113, pp. 226–242, 2017.