Implementation of ECLAT Algorithm Technology: Determining Books Borrowing Pattern in University Library

Lisnawita 1 & Mariza Devega 2

1Universitas Lancang Kuning, Pekanbaru, 28265, Indonesia 
E-mail: lisnawita@unilak.ac.id
2Universitas Lancang Kuning, Pekanbaru, 28265, Indonesia 
E-mail: marizadevega@unilak.ac.id

Abstract. Library is one place to get lots of information. Every transaction information recorded in a fairly large database, the large data if it can’t be used it will make the problem for the librarian. This research aims to determine the books that are often borrowed when borrowing, using the rules of the mining association using the ECLAT algorithm. Equivalence Class Transformation Algorithm (ECLAT) performs frequent itemset search from the bottom. This algorithm will only scan the data once, the scanning process will not be repeated to get frequent k-itemset. The database scan process in the ECLAT algorithm is not repeated because on the search itemset not pay attention the sequence from the item. The result of this study indicate the performance of the ECLAT algorithm is good by requiring 15 ms in the execution process. 102 transaction data the same minimum support is 1%, resulting in 21 frequent itemset.

Keywords: library, book borrowing, data mining, eclat

1. Introduction

The Library Unit (Unilak) has 11,274 overall book titles, UPT Unilak Library can serve 5 to 20 book lending transactions every day. The library UPT should be able to generate more transactions, even though in the book lending transaction Unilak UPT library can reach thousands of transactions each year, so that the library has limitations in providing information on book recommendations that should be collected in the library, because there are so many members who want to borrow books, but the available book collections are still not well targeted, so that the books collected can only fulfill certain fields. With the existence of these problems, the authors want to help the manager of the library UPT to always be able to provide the best service by analyzing the book lending pattern so that later it will provide useful information for managers, so the UPT Unilak Library can provide a collection of books according to the needs of the academic community, as well as parties the manager can establish the policies in the fdeveloping the UPT Unilak library.
2. **Theory: the review**

Joseph Eric Samodra, Budi Susanto, Willy Sudiarto Raharjo (2015) Data analysis process using relatively short algorithm (2304 transactions within 61 ms and 1984 transactions within 19ms at 10% minimum support). Garima Sinha, Dr. S. M. Ghosh (2014) This study compares three algorithms, FP Growth, Apriori and Eclat. Based on the data mining process, the fastest eclat based on execution time. And in association rules generated by these three algorithms can be combined to help an efficient algorithm. Miranda Nur Qolbi Aprilina, Wiranto, Widodo (2016) In the same year the pattern of book lending transactions generated every month is not always the same. The pattern of book lending transactions every month influences the formation of transaction patterns by semester and year. The pattern that often appears in the month of the same year then the pattern also appears in the semester and the year shows the pattern is consistent

3. **Research Methods**

The steps taken in this study can be seen in Figure 3.1 below

![Research Step Diagram](image)

After formulating the problem and collecting data, the Knowledge Discovery in Database stages are carried out as follows:

3.1. **Data Selection**

This step to analyze the data needed for the next process.

3.2. **Data Pre-processing**

This step to eliminating irrelevant attributes whose existences will reduce the accuracy of data mining

3.3. **Data Transformation**

The transformation stage is a stage by doing the attribute construction that is to divide the data into several ranges. In this study the variables that use the range are book title.
3.4. Data Mining

Equivalence Class Transformation (ECLAT) algorithm is a method that changes the horizontal TID itemset data format to a vertical TID itemset data format. The itemlist search process is carried out from the itemlist that most often appears until the most rarely appears without needing to pay attention to the order so that there is no need to repeat the search.

3.5. Implementation, Testing dan Interpretation

The stages of implementation, testing and interpretation of results are frequent itemsets search stages from research data and research sample data using SPMF tools. Then search frequent 4-itemsets at different minimum support. The results obtained from the previous process will be interpreted as information.

4. Result and discussion

Borrowing the data transaction used from 2015-2017 with many transactions 1502 transactions. So from that data, we can look for patterns of book borrowing that have occurred over the past 3 years. To find the pattern, the stages of knowledge data discovery are also called data mining. The data recorded so far has been kept by borrowing from Unilak library books, namely NIM, student names, Borrowed Book Titles.

4.1. Data Selection

In this study the variable used is the name of the book borrowed at the Unilak library. The book data taken is data that borrows more than one so that the association can be made.

| No | Book Lending Data                                   | Buku |
|----|----------------------------------------------------|------|
| 1  | Pengantar Ilmu Komunikasi                          | Buku |
| 2  | Pendidikan Ilmu Manajemen                         | Buku |
| 3  | Lingkungan Dan Kesehatan Environmet And Health     | Buku |
| 4  | Akuntansi Pengantar Satu                           | Buku |
| 5  | Manajemen Sumber Daya Manusi                       | Buku |
| 6  | Pokok- Pokok Materi Statistik 2 (Statistik Inferens) | Buku |
| 7  | Manajemen Keuangan Berbasis Balanced Scorecard     | Buku |
| 8  | Teori Teori Psikologi                              | Buku |

*Figure 1* Book Lending Data

4.2. Data Preprocessing

At this stage data cleaning do from noise or data duplication. The clean data obtained after cleaning data is 1005 datasets.
4.3. Data Transformation

At the transformation data stage, the title of the book is initialized so that there is no mistake in the title of the book that is too long. The transformation made is that the book title is made into all variable names, then initializes each book loan transaction data with Y.

4.4. Data Mining

The process of data mining using Weka software, by changing the CSV file to Arff format. Furthermore, by using the SPMF Tools data mining process is carried out with the following pattern search process.

| Support | 100% | 90% | 80% | 70% | 60% | 50% | 40% | 30% | 20% | 10% | .... | 3% | 2% | 1% |
|---------|------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|----|----|----|
| Frequent Itemset | -    | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   | 3  | 21 |    |

The following is the datamining process with the Eclat algorithm, with a minimum support of 1%, by generating 21 frequent itemset. From the picture above, there are associate patterns produced as follows:

**Figure 2** Transaction of Borrowing more than one book

**Figure 3** pattern results with eclat algorithms
Anatomi Tumbuhan Berbiji → Biologi Sel
Manajemen Pemasaran: Marketing Management 9e Jilid 1 → Akuntansi Biaya
Manajemen Pemasaran Jilid 2 → Manajemen Pemasaran: Marketing Management 9e Jilid 1

By using ECLAT algorithm with minimum support of 1%, by generating 20 frequent itemset with execution time of 0ms.

```
buku54=Y buku85=Y 12
buku270=Y buku268=Y 14
buku20=Y buku27=Y 13
```

From the picture above there is an association pattern produced is book 54 = Y book85 = Y data appearance 12 times, book270 = Y book268 = Y with appearance 14 times, and book20 = Y book27 = Y with appearance 1 time.

4.5. Interpretation

The pattern produced between a priori algorithms and algorithms has the same frequent items as many as 20, and for the results of two candidates itemset also has 2 highest rules with the results of Book 268 → Books 270 and Book 54 → Book 85

Book 20 is a Cost Accounting book
Book 27 is a financial accounting book
Book 54 is the Anatomy of Seed Plants
Book 85 is a Cell Biology book
Book 268 is the book of Marketing Management Volume 2
Book 270 is a book on Marketing Management: Marketing Management 9e Volume 1

Then interpreted the results of the rules into a language that is easier to understand that the resulting association rules is when students borrow the book Anatomy of Seed Plants, students also borrow Cell Biology books at the same time with the appearance of data 12 times. And when students borrow the Volume 2 Marketing Management book, at the same time students also borrow the book Marketing Management: Marketing Management 9e Volume 1 with data appearing 14 times. And when borrowing a book of cost accounting, students will also borrow financial accounting books with 13 appearances.

5. Conclusion

Using an eclat algorithm takes 15 ms in the execution process. Where is the pattern of book borrowing, that is when students borrow the Anatomy of Seed Plants, students also borrow Cell Biology books at the same time, When students borrow the Volume 2 Marketing Management book, at the same time students also borrow the book Marketing Management: Marketing Management 9e Volume 1. And if borrowing cost accounting book, students will also borrow financial accounting books. And the recommendation of the book produced is the book Marketing Management: Marketing Management 9e Volume 1 With 1.39% support with 100% confidence.

6. References

[1] Sinha, G., & Ghosh, S. M. 2014. Identification of best algorithm in association rule mining based on performance. *Int. J. Comput. Sci. Mob. Comput.*, 3(11), 38-45.

[2] Samodra, J.E., Susanto, B. and Raharjo, W.S., 2015. Implementasi Algoritma Eclat Untuk Frequent Pattern Mining Pada Penjualan Barang. Media Teknika Jurnal Teknologi, 10 (2), pp.101-110.
[3] Aprilina, M. N. Q. 2016. Analisa Konsistensi Pola Peminjaman Buku Menggunakan Algoritma FP-Growth (Studi Kasus: UPT Perpustakaan Universitas Sebelas Maret).

[4] Zaki, M. J. 2014. Data Mining And Analysis: fundamental concepts and algorithms . Library of Congress Cataloging.

[5] Anas, A. 2017. Analisa Algorithma Apriori Untuk Mendapatkan Pola Peminjaman Buku Perpustakaan SMPN 3 Batanghari. *Jurnal Ilmiah Media SISFO*, 10(2), 628-641.

[6] Han, J., Pei, J. and Kamber, M., 2011. Data mining: concepts and techniques. Elsevier.

[7] Kusrini, E. T. 2009. *Algoritma Data Mining*. Yogyakarta: ANDI.

[8] Norsyanah, E. 2016. Penerapan Algoritma ECLAT Dalam Menentukan Metode Kontrasepsi Yang Dipilih (Doctoral dissertation, Universitas Islam Negeri Sultan Syarif Kasim Riau).

[9] Wandi, N., Hendrawan, R. A., & Mukhlason, A. 2012. Pengembangan Sistem Rekomendasi Penelusuran Buku dengan Penggalian Association Rule Menggunakan Algoritma Apriori (Studi Kasus Badan Perpustakaan dan Kearsipan Provinsi Jawa Timur). *Jurnal Teknik ITS*, 1(1), A445-A449.

[10] Guntoro, G. and Fikri, M., 2018. Perancangan Aplikasi Single Sign-On Menggunakan Autentikasi Gambar. *Digital Zone: Jurnal Teknologi Informasi dan Komunikasi*, 9(1), pp.12-21.