Implementation of *Association Rule* With Algorithm Apriori On Loan Data Library and Archives Service Book Regency Sukoharjo

Septiana Cahaya Sari[1], Arif Himawan[1], Aris Wahyu Murdiyanto[1]
[1]Information System, Faculty of Engineering and Information Technology, Universitas Jenderal Ahmad Yani, Yogyakarta, Indonesia
E-mail: *[1]**26.septianacs@gmail.com**

Abstract

The library has an important role in improving literacy, education, and facilitating access to information for the community. The Department of Libraries and Archives of Sukoharjo Regency has a high number of collections and visitors every year. An analysis of book borrowing transaction data is necessary to obtain information that can enhance the quality of services in the Sukoharjo Regency Library. This research aims to process book borrowing data at the Sukoharjo District Library and Archives Office by applying the Knowledge Discovery in Databases method. In addition, this also seeks to implement the Apriori algorithm to discover association rules that illustrate the relationships between books that are often borrowed together by library members, as well as to provide recommendations for book management to the library staff. The Knowledge Discovery in Databases method is used because it is a systematic approach that focuses on collecting hidden knowledge from large and complex data. This method consists of five main stages, namely selection, preprocessing, transformation, data mining, and evaluation. This research succeeded in identifying patterns of book borrowing at the Sukoharjo Regency Library and Archives Service based on 1,052 lending transaction data, with a minimum support of 0.005 and a confidence of 0.2 obtained from 64 association rules.

Keywords*: Knowledge Discovery in Databases, Association Rule, Apriori Algorithm, Sukoharjo Regency Library and Archives Service.*

1. INTRODUCTION

The library's own role in supporting education, improving literacy, and expanding access to information for the public [1] [2]. The library keeps innovating by utilizing technology information like digital libraries, RFID (*Radio Frequency Identification*), and *barcode scanners* to improve the quality and quantity of service to visitors and members. [3]The Library and Archives Service Regency Sukoharjo is one of the regional apparatuses authorized in matters of implementing government in the field of libraries and archives. During two years, the library's own collection of books printed as many as 56,067 copies covering 32,750 titles [4].

Based on Statistical Data 2024 number of visitors reached 10,191, with the total number of borrowed books reaching 3,445 [5]. The findings of this data reflect high interest and participation community in utilizing services provided by the library area of Regency Sukoharjo. If the data is not managed properly, the collection of information does not have useful benefits for a government agency [6]. Analyzing the loan data book manually requires a

lot of time. However, with the progress of technology that continues to as technology develop, this process can be done faster and easier, even for large amounts of data. Big [7].

This research was conducted to identify pattern relatedness between books borrowed by members using approach *Association Rule*, Algorithm A priori Algorithm, and the methodology of *Knowledge Discovery in Databases* (KDD). It is expected results of this research can be made into recommendations addition to the collection and optimization compilation racks at the Library and Archives Service, Regency Sukoharjo. In addition, the implementation of results from this research is expected can increase literacy public area of Regency Sukoharjo, as well as support the achievement Golden Indonesia 2045 vision.

2. RESEARCH METHODS

2.1 Selection (Data Collection)

Data was taken from the Inlislite application of the Sukoharjo Regency Library and Archives Service, with the main attributes used being *the member number* as the borrower's identity and *bibliographic data* as the book's identity.

2.2 Pre-Processing (Data Cleaning)

Some processes that can be carried out at the pre-processing stage include checking for null (empty) data, then deleting duplicate data, handling missing values, and, if necessary, adjusting the format of attributes that are not yet in accordance with the analysis process.

2.3 Transformation

Data is transformed become a transaction format with binary numbers (1 and 0), where every transaction is grouped based on the *number of members* and *titles* borrowed *books* To fulfill the implementation algorithm, A priori.

2.4 Data Mining

The data mining process is carried out to find a pattern connection between book titles and stages :

a. Analyzing frequency itemsets tall use mark *support*, namely percentage patterns in the dataset.

$$Support = \frac{Jumlah\ Transaksi\ Mengandung\ A}{Total\ Transaksi} \quad (1)$$

[8]

b. Forming a rule association based on mark *confidence*, namely, a strong connection between items in a pattern association.

$$Confidence = P(B|A) = \frac{Jumlah\ Transaksi\ Mengandung\ A\ dan\ B}{Jumlah\ Transaksi\ Mengandung\ A} \quad (2)$$

[9]

2.5 Evaluation

Analysis results were evaluated to get recommendations pattern borrowing books based on the value of support and confidence, then visualized using:

a. Table (rules association, *support*, *confidence*)
b. Network Graph (relationship) between books based on line thickness)
c. *Heatmap* (strength connection between books on a color scale)

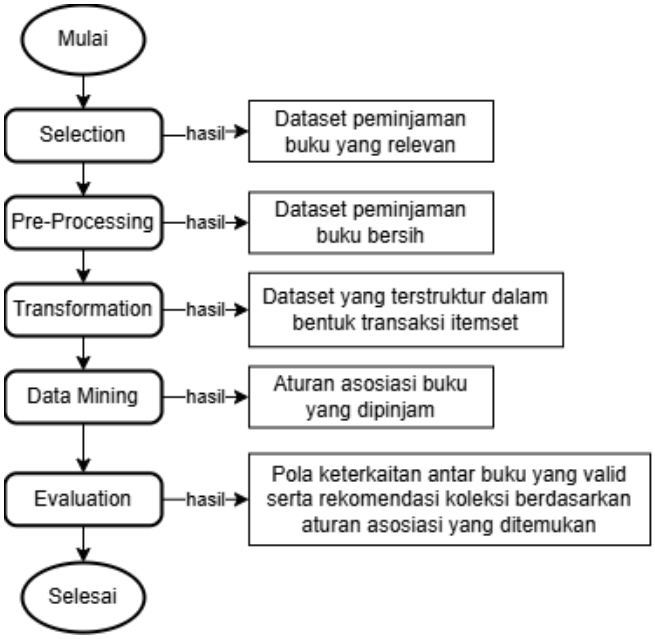


Figure 2.1 *Metode Penelitian*

3. RESULTS AND DISCUSSION

3.1. Selection

Data used is the loan data book 2021 – 2024, consisting of 6,787 entries, with 13 attributes.

No	Tanggal Pinjam	Tanggal Pinjam	Tanggal Dikemba	Jumlah Ha	Nomor Ind	Data Bibli	Nomor Kl	Nomor An	Nama Ang	Jenis Kela	Kelompok	Nama Pet	Nama Petugas Pengambalian
1	10/06/2021 00:00	11/06/2021 00:00	10/06/2021 00:00	11	29888.P.1	Metode	300 - Ilmu	10606.21	ANWAR N	Laki-laki	Masa Rem	Inlislite	Inlislite
2	10/06/2021 00:00	11/06/2021 00:00	17/06/2021 00:00	4	29888.P.1	Metode	300 - Ilmu	10606.21	ANWAR N	Laki-laki	Masa Rem	Inlislite	Inlislite
3	15/06/2021 00:00	16/06/2021 00:00	21/07/2021 00:00	-27	45705/P/2	101	200 - Agan	10609.21	FARHANA	Perempua	Masa Dew	Inlislite	Inlislite
4	16/06/2021 00:00	17/06/2021 00:00	29/06/2021 00:00	-4	44300/H/1	Let's	100 - Filsa	10610.21	SRI NUR H	Perempua	Masa Rem	Inlislite	Inlislite
5	16/06/2021 00:00	17/06/2021 00:00	29/06/2021 00:00	-4	29458.P.1	Supertrik	300 - Ilmu	10610.21	SRI NUR H	Perempua	Masa Rem	Inlislite	Inlislite
6	17/06/2021 00:00	18/06/2021 00:00	06/07/2021 00:00	-8	40740/H/1	Hukum	300 - Ilmu	10606.21	ANWAR N	Laki-laki	Masa Rem	Inlislite	Inlislite
7	21/06/2021 00:00	22/06/2021 00:00	28/06/2022 00:00	-363	26955.P.1	Problem	300 - Ilmu	10389.19	YUDHA AJ	Laki-laki	Masa Dew	Inlislite	Inlislite
8	21/06/2021 00:00	22/06/2021 00:00	28/06/2022 00:00	-363	36688.P.1	Sosiologi	300 - Ilmu	10389.19	YUDHA AJ	Laki-laki	Masa Dew	Inlislite	Inlislite
9	22/06/2021 00:00	23/06/2021 00:00	07/07/2021 00:00	-29	27793.P.1	Manajem	300 - Ilmu	10611.21	MUH. ABD	Laki-laki	Masa Dew	Inlislite	Inlislite
10	29/06/2021 00:00	30/06/2021 00:00	06/07/2021 00:00	2	34716.P.1	Sejarah	200 - Agan	10293.18	ANITA AN	Perempua	Masa Rem	Inlislite	Inlislite
11	30/06/2021 00:00	01/07/2021 00:00	24/09/2021 00:00	-77	45840/P/2	Mommyc	600 - Ilmu	5439.13	ENGGAR E	Perempua	Masa Dew	Inlislite	Inlislite
12	07/07/2021 00:00	08/07/2021 00:00	21/07/2021 00:00	-5	27793.P.1	Manajem	300 - Ilmu	10611.21	MUH. ABD	Laki-laki	Masa Dew	Inlislite	Inlislite
13	21/07/2021 00:00	22/07/2021 00:00	04/08/2021 00:00	2	27793.P.1	Manajem	300 - Ilmu	10611.21	MUH. ABD	Laki-laki	Masa Dew	Inlislite	Inlislite
14	21/07/2021 00:00	22/07/2021 00:00	04/08/2021 00:00	2	42279/H.B	Manajem	600 - Ilmu	10611.21	MUH. ABD	Laki-laki	Masa Dew	Inlislite	Inlislite
15	21/07/2021 00:00	22/07/2021 00:00	21/07/2021 00:00	9	45664/P/2	Data	000 - Kary	10613.21	SUFYAN N	Laki-laki	Masa Rem	Inlislite	Inlislite
16	04/08/2021 00:00	05/08/2021 00:00	25/08/2021 00:00	-12	27793.P.1	Manajem	300 - Ilmu	10611.21	MUH. ABD	Laki-laki	Masa Dew	Inlislite	Inlislite
17	04/08/2021 00:00	05/08/2021 00:00	25/08/2021 00:00	-12	42279/H.B	Manajem	600 - Ilmu	10611.21	MUH. ABD	Laki-laki	Masa Dew	Inlislite	Inlislite
18	25/08/2021 00:00	26/08/2021 00:00	07/09/2021 00:00	2	42213/H.B	Sejarah	200 - Agan	10611.21	MUH. ABD	Laki-laki	Masa Dew	Inlislite	Inlislite
19	25/08/2021 00:00	26/08/2021 00:00	31/08/2021 00:00	3	45695/P/2	Perpusta	700 - Kese	10615.21	REVALINA	Perempua	Masa Rem	Inlislite	Inlislite
20	25/08/2021 00:00	26/08/2021 00:00	31/08/2021 00:00	3	45609/P/2	Yuk,	700 - Kese	10615.21	REVALINA	Perempua	Masa Rem	Inlislite	Inlislite

Figure 3.1 *Dataset Real*

Table 1 Data Description

Attribute	Information
Borrow Date	The date the member borrowed the book

Member Numbe	Bibliographic Data
10606.21	Legal Research Methods Bandung Pustaka Setia 2008 Legal Research Methods 340,072
10606.21	Legal Research Methods Bandung Pustaka Setia 2008 Legal Research Methods 340,072
10609.21	101 Secrets of Educating Pious and Salihah Children cet. 1 Yogyakarta: Mueeza, 2021 Children's Education - Islam 297,577
10610.21	Let's Move Up: because success requires guts!! 1st edition Surakarta: Indiva Media Kreasi, 2014 success 158
10610.21	Supertrik Mathematics for Senior High School/Vocational High School, 1st Edition, Yogyakarta, Jogja Bangkit Publisher 2013 Mathematics - Senior High School 373,238
10606.21	State Financial Law Jakarta: Gramedia Widiasarana Indonesia, 2013 Finance - Law 343,035 98

Due date	Book return deadline
Return Date	Actual date of book return
Number of Days Late	Difference in days between the actual return date and the due date
Registration number	Book identification number
Bibliographic Data	Title and identification of the book borrowed
Class Number	Classification of books by category
Member Number	Unique identification number of library members
Member's name	Full name of the member who made the loan
Gender	Member gender (Male/Female)
Age Group	Library member age groups (Teenagers, Adults, Seniors)
Officer Name	Name of the officer who serves book loans
Name of Return Officer	Name of the officer who serves book returns

Furthermore, combining member numbers and bibliographic data for analysis and application algorithm a priori.

Table 2 Integration result dataset

3.2. Pre-Processing

Removes *null* data on the ' Bibliographic Data' attribute and generates the dataset that was originally totaling 6787 entries, to 6780 data row entries. Next, do an extraction on the attribute ' Bibliographic Data ', borrowing dataset book only has two attributes, namely member number and bibliographic data (book title).

Table 3 Result dataset title extraction

3. 3 . Transformation

Transformation is the process of converting data into a format suitable for processing using data mining techniques. The data collected is data that has a minimum of 2 to 3 transactions. The data will be processed using *Google Collaboratory*, namely in the form of binomial 1 and 0, with a total of 1,052 transaction data.

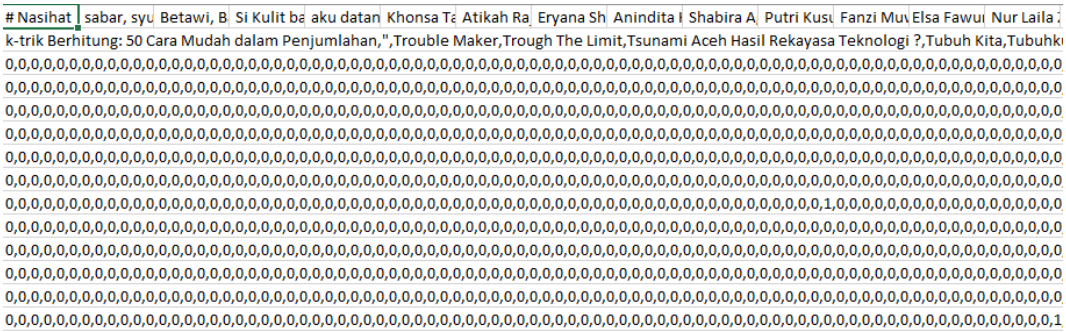


Figure 3.2 *Dataset hasil transformasi*

3. 4. Data Mining

a. Determine mark *support*

Member Number	Bibliographic Data
10606.21	Legal Research Methods
10606.21	Legal Research Methods
10609.21	101 Secrets to Raising Pious and Righteous Children
10609.21	101 Secrets to Raising Pious and Righteous Children
10610.21	Let's Move Up: because success takes guts!!
10610.21	High School Mathematics Supertricks

Count mark *support* based on frequently used book titles appears using the formula in equation (1).

Table 4 3-Itemset Support Value

Book title	Number of Items	*Support*

The Other Side, Shea, Laut Storytelling / Leila S. Chudori	(1/10) * 100%	10%
The Other Side, Shea, Cantik That's Wound	(1/10) * 100%	10%
Sesuk / Tere Liye, Sea Storytelling / Leila S. Chudori, Beautiful That's Wound	(1/10) * 100%	10%
Sesuk / Tere Liye, Cookie Run Sweet Escape Adventure! – Animals, The Other Side	(1/10) * 100%	10%

Based on the results, Table 3.4 shows that the minimum *support* value that meets the 3-itemset requirement is 10% (0.1). A *support value of* 0.1 does not meet the requirement. Rule, therefore, the value *support* is lowered to 0.005, which is obtained from the number of copies of each book title / by the total itemset transactions.

b. Determine mark *confidence*

Table 5 Confidence Value

Itemset	***Support* (%)**	***Confidence* (%)**
The Other Side, Shea	(3/10) * 100 = 30	(3/5) * 100 = 60
The Other Side, Sesuk / Tere Liye	(1/10) * 100 = 10	(1/5) * 100 = 20
The Other Side, Sea Storytelling / Leila S. Chudori	(1/10) * 100 = 10	(1/5) * 100 = 20
The Other Side, Beautiful That's Wound	(2/10) * 100 = 20	(2/5) * 100 = 40
Shea, Sesuk / Tere Liye	(3/10) * 100 = 30	(3/3) * 100 = 100
Shea, Beautiful. That's Wound	(1/10) * 100 = 10	(1/3) * 100 = 33.3
Sesuk / Tere Liye , Sea Storytelling / Leila S. Chudori	(2/10) * 100 = 20	(2/4) * 100 = 50
Sesuk / Tere Liye, Cookie Run Sweet Escape Adventure! - Animal	(1/10) * 100 = 10	(1/4) * 100 = 25
Sea Storytelling / Leila S. Chudori, Beautiful That's Wound	(1/10) * 100 = 10	(1/4) * 100 = 25

Oops, Caught !, Cookie Run Sweet Escape Adventure! - Animals	(1/10) * 100 = 10	(1/1) * 100 = 100

Minimum threshold value for *confidence* determined based on manual calculations in popular books that can be seen in Table 3.5, where the minimum *confidence* value is 20% = 0.2. The next process is to apply the algorithm a priori with value *min_support* 0.5% (0.005) and *min_confidence* 20% (0.2), which results in 64 rules of association.

3. 5 . Evaluation

Based on the results rule association was used to clarify the connection between book items used for visualization, which is visualized in the following table and figure :

a. Visualization Table

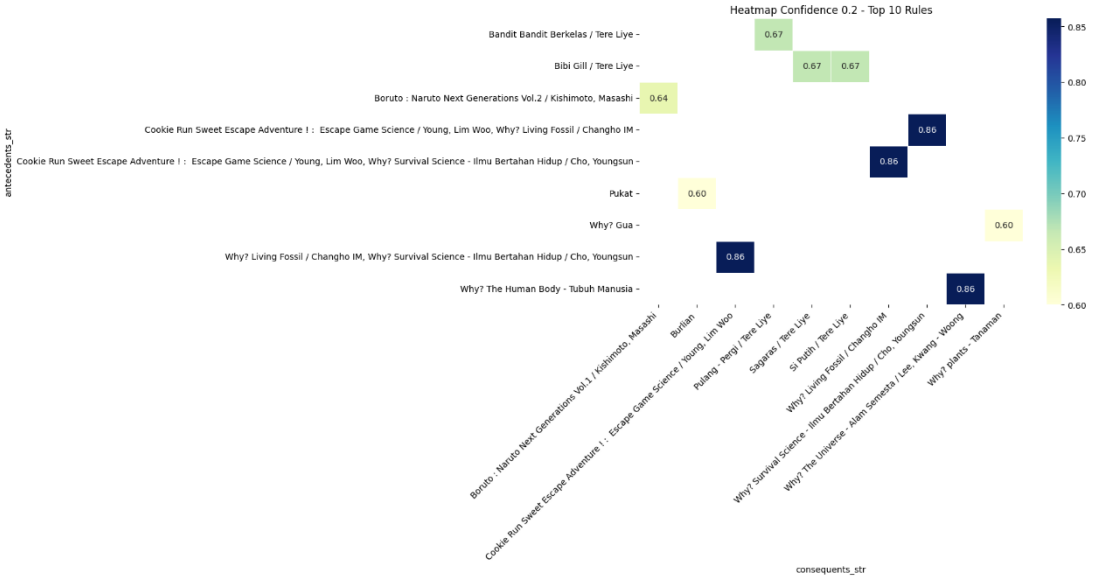
Table 6 Association results *min_support* 0.005 and *min_confidence* 0.2

Antecedents	***Consequences***	***Support***	***Confidence***
Why? Living Fossil / Changho IM, Why? Survival Science – Science Endure Life / Cho, Youngsun	Cookie Run Sweet Escape Adventure!: Escape Game Science / Young, Lim Woo	0.006	0.9
Cookie Run Sweet Escape Adventure! : Escape Game Science / Young, Lim Woo, Why? Living Fossil / Changho IM	Why? Survival Science – Science Endure Life / Cho, Youngsun	0.006	0.9
Cookie Run Sweet Escape Adventure! : Escape Game Science / Young, Lim Woo, Why? Survival Science – Science Endure Life / Cho, Youngsun	Why? Living Fossil / Changho IM	0.006	0.9
Why? The Human Body Man	Why? The Universe – Nature Universe / Lee, Kwang – Woong	0.006	0.9
Bandit Bandit Classy / Tere Liye	Round Trip / Tere Liye	0.006	0.7
Aunt Gill / Tere Liye	Sagaras / Tere Liye	0.006	0.7

Aunt Gill / Tere Liye	The White One / Tere Liye	0.006	0.7
Boruto: Naruto Next Generations Vol 2 / Kishimoto, Masashi	Boruto: Naruto Next Generations Vol . 1 / Kishimoto , Masashi	0.007	0.6
Trawl	Burlian	0.006	0.6
Why? Cave	Why? Plants	0.006	0.6

Based on Table 3.6 of 90% of members borrow library books. Why? Living Fossil / Changho IM, Why? Survival Science - Science Endure Life / Cho, Youngsun, so borrow book Cookie Run Sweet Escape Adventure!: Escape Game Science / Young, Lim WOo.

b. Visualization *Heatmap*



Gambar 3.3 *Visualisasi heatmap Top 10 aturan asosiasi*

Visualization uses a *Heatmap* based on the highest confidence value . Where the y-axis shows *antecedents, the* x-axis depicts *consequents*, and existing values in the heatmap box show mark *confidence* from every rule association. The more dark the color on the box *heatmap*, the higher the mark *confidence* his.

c. Visualization *Network Graph* (Network Diagram)

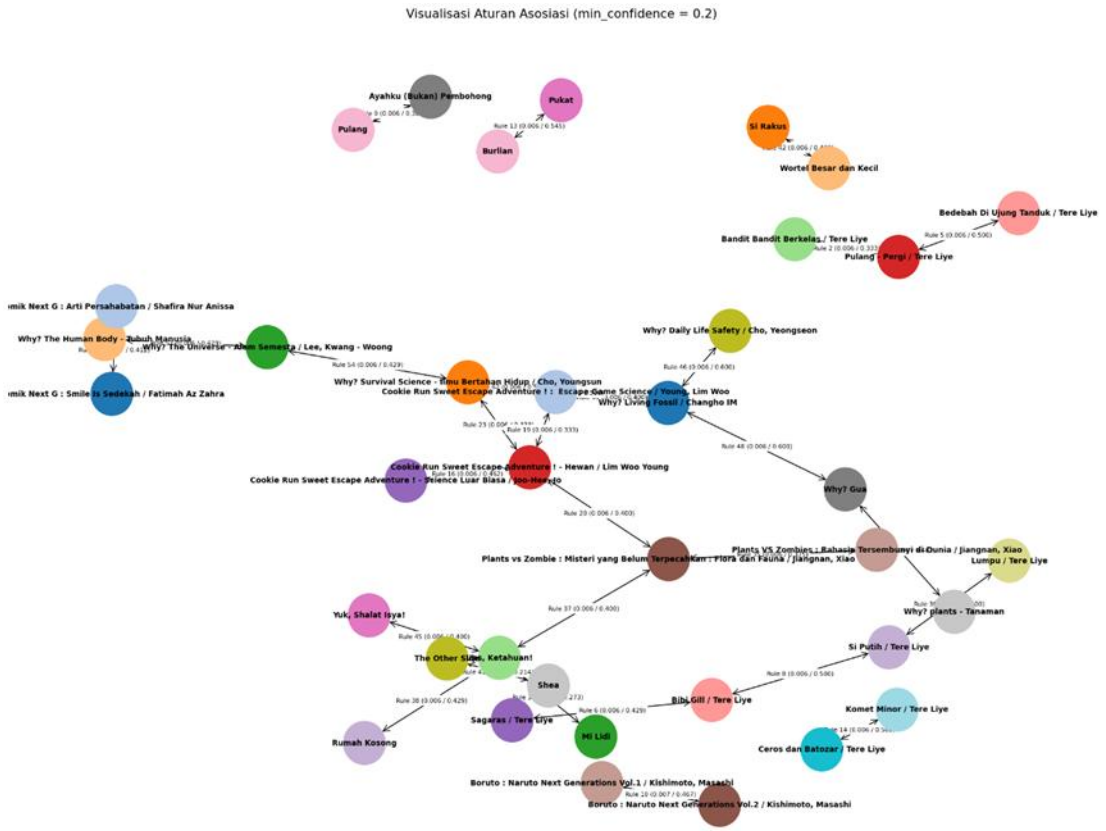


Figure 3.4 *Diagram jaringan hasil asosiasi* min_support *0,005 dan* min_confidence *0.2*

Circle. The colors (*nodes*) indicate the title of the book, the arrow lines (*edges*) indicate direction or relationship *antecedents* and *consequents*. Coloring different *nodes* shows group books that are related to each other from pattern association. Each arrow line displays the order of which rule? Accompanied by valuable *support* and *confidence*.

4. CONCLUSION

This research applies the approach of *Knowledge Discovery in Databases on 6780* loan data Library and Archives Service book Regency Sukoharjo, with data transformation resulting in 1052 transactions. Implementation algorithm A priori done with a minimum *support* value of 0.005 and minimum *confidence* of 0.2, resulting in 64 rules, pattern borrowing, with a value of *confidence* highest reached 90%. The pattern shows that 90% of members borrow the books. Why? Living Fossil and Why? Survival Science - Ilmu Endure Life is also borrowing Cookie Run Sweet Escape Adventure!: Escape Game Science book. Based on these findings, it is recommended to add the number of copies of frequently read books borrowed simultaneously, providing rack special " Rack Popular " rack for books with patterns that connect high, and

develop a system recommendation website or application based on loan data previously used to increase service and management of the collection book.

5. SUGGESTION

There are some suggestions that can be delivered as material considerations for further development, including :

a. Do analysis on the lending dataset books with attribute additions such as class number, group age, and date borrowed to get a comparison of varying results.
b. Compare the implementation algorithm a priori with other algorithms such as FP-Growth or ECLAT to provide a description effectiveness of the model in finding rule association, so that it can made into consideration which algorithm is the most efficient and accurate to apply in managing loan data? Books in the library, in a way, are sustainable.
c. Based on the results implementation algorithm a priori, the party Library Regency Sukoharjo can utilise the results pattern association optimally for procurement, management, and arrangement collection, as well as develop a " recommendation " feature book on the system service circulation in the Library Regency Sukoharjo.
d. Apply research with a similar theme in a larger library to see the effectiveness of Algorithm A priori in more complex collections and transactions .

.

REFERENCE

[1] O. Laily, K. Behesty, and K. Kunci, "Pendidikan Indonesia Bandung," 2023. doi: 10.62668/jimr.v2i02.845.

[2] N. Luh, P. Noviartini, N. Putu, A. Prabawati, and K. Winaya, "Strategi Pengelolaan Perpustakaan Daerah Provinsi Bali dalam Meningkatkan Minat Baca Masyarakat Management Strategy of the Regional Library in Bali Province to Enhance Community Reading Interest," Online, 2024. doi: 10.61292/birev.93.

[3] A. Octaviani and P. Dewi, "Big Data di Perpustakaan dengan Memanfaatkan Data Mining," *Anuva*, vol. 4, no. 2, pp. 223–230, 2020, doi: 10.14710/anuva.4.2.223-230.

[4] Tunardi, "2024, Banyaknya Koleksi Cetak Perpustakaan berdasarkan Jenis," https://data.sukoharjokab.go.id/dataset/2024-banyaknya-koleksi-cetak-perpustakaan-berdasarkan-jenis.

[5] Proboningsih, "2024, Data Pengunjung (Pengguna Layanan) Perpust Konvensional," https://data.sukoharjokab.go.id/dataset/2024-data-pengunjung-pengguna-layanan-perpust-konvensional.

[6] F. Widya Andhika *et al.*, "Penerapan Algoritma Apriori pada Sistem Peminjaman Buku di Perpustakaan SMK Pemuda 3 Kesamben," 2022. doi: 10.36040/jati.v6i2.5779.

[7] Y. Andini, J. Tata Hardinata, Y. Pranayama Purba, and P. A. Studi Sistem Informasi STIKOM Tunas Bangsa JlJend Sudirman Blok No, "Penerapan Data Mining Terhadap Tata Letak Buku Di Perpustakaan Sintong Bingei Pematangsiantar Menggunakan Metode Apriori," 2022. doi: 10.51351/jtm.11.1.2022661.

[8] I. Fitri Polorida Ginting and D. Saripurna, "Penerapan Data Mining Dalam Menentukan Pola Ketersediaan Stok Barang Berdasarkan Permintaan Konsumen Di Chykes Minimarket Menggunakan Algoritma Apriori," *Jurnal Sains Manajemen Informatika dan Komputer*, vol. 20, no. 1, pp. 28–37, 2021, doi: 10.53513/jis.v20i1.2504.

[9] A. Ibezato Zalukhu, D. Sartika, and S. Wahyuni, "Bulletin of Information Technology (BIT) Penerapan Algoritma Apriori untuk Optimasi Strategi Penjualan Berdasarkan Analisis Pola Pembelian di Torsa Cafe," vol. 5, no. 4, pp. 295–304, 2024, doi: 10.47065/bit.v5i2.1715.