Implementation of Data Mining on Sales Data of Bambu Ungu Cafe to Find out Consumer Purchasing Patterns Using the Apriori Algorithm

Rahmad Fadli [1], Fahmi Ruziq [2], Chairul Imam
[123] Universitas Battuta, Medan, Indonesia
[123] Fakultas Teknologi, Universitas Battuta, Medan, Indonesia
[1] rahmadfadli614@gmail.com [2] fahmiruziq89@gmail.com [3] chairulimam313@gmail.com

Article Info	ABSTRACT
	In today's digital age, restaurants or cafes face increasing competition and complex challenges. To stay relevant and succeed in this business, the use of data mining is crucial, as it becomes a valuable tool in optimizing sales, increasing customer satisfaction, and achieving long-term success in the restaurant and café industry. Data mining helps in effective marketing strategies. By analyzing customer data, purchase information, and preferences restaurants and cafes can identify different customer segments and create customized marketing. With so much sales transaction data, it will certainly be difficult if the data is analyzed manually, therefore information will be obtained if there is processing with the help of a system to get sales patterns. The results of this processing will produce transaction information to support product transaction decisions. To solve this problem, the researcher designed a calculation information system. In this case the author uses the waterfall method in the research process. For system design the author uses the PHP programming language with a database format using MySQL. Finally, with this information system, the calculation process can be done automatically without the need to calculate manually is appropriate, provided that all data inputted is valid. *This is an open-access article under the license.*

Corresponding Author:

Rahmad Fadli
Universitas Battuta
Email: rahmadfadli614@gmail.com

1. INTRODUCTION

In today's digital era, restaurants or cafes face increasingly fierce competition and complex challenges [1]. The emergence of technology in replacing the role of previous activities that use traditional methods or have not used databases is one example of how technology can revolutionize various aspects of life, including in education and business [2]. To remain relevant and successful in this business, the use of data mining becomes very important [3], as data mining becomes a valuable tool in optimizing sales, increasing customer satisfaction, and achieving long-term success in the restaurant and cafe industry. For example, diagnosis of thyroid disease is difficult because the symptoms can vary and resemble other diseases, so proper interpretation of clinical data is needed for accurate diagnosis using data mining [4].

Data mining helps in effective marketing strategies. By analyzing customer data, purchase information, and preferences restaurants and cafes can identify different customer segments and create customized marketing. Data mining also helps in stock and inventory management, so that restaurant and cafe owners can estimate the demand for certain products in a certain period of time.

Bambu Ungu is a business entity engaged in the sale of light dishes to relax for the general public. Where this cafe itself is one of the hundreds of cafes in the city of Medan. In the competitive business world, especially the restaurant industry such as this cafe, attracting consumers to be interested in food and beverage products is not an easy thing. This cafe includes quite a lot of visitors, but often experiences problems such as unavailability of food or stock, even though the order is classified as in demand.

Basically, the owner of Bambu Ungu has not analyzed the data specifically, for example, to determine what food and drinks customers often order or the relationship between one item and another, causing problems such as running out of food and beverage ingredients. With so much sales transaction data, of course it will be difficult if the data is analyzed manually, therefore information will be obtained if there is processing with the help of a system to get sales patterns. The results of this processing will produce transaction information to support product transaction decisions. Thus, these results help determine the most desirable sales patterns or products.

The rapid development of information technology has a significant impact in various aspects of life, including in decision making [5], such as a study conducted to predict liver cancer [6] and a new member admission selection system [7]. The a priori algorithm can be used in the sales process, by providing a relationship between items in sales data, in this case food or drinks ordered so that consumer purchasing patterns will be obtained. The purchase pattern is processed to produce information which can later become the basis for making procurement decisions. According to Ummaerah Safriaty Juwaeni Jibril's research entitled "Market Basket Analysis with the Apriori Algorithm Using Association Rule for Sales of Heavy Equipment Spare Parts at Cv. Sentosa Sukses Abadi" concluded that using data analysis using the apriori algorithm can help the company in making decisions to determine the marketing strategy of products to be sold against purchasing patterns [8].

Based on the root of the problem, research is carried out that can produce an information system that is able to provide customer purchasing patterns and provide feedback in order to evaluate decision making for procurement of goods in a certain period. The information system created also goes through the stages of the system development life cycle (SDLC) with the waterfall method where SDLC is an important approach in software development to help reduce the risk of errors and ensure the resulting system is in accordance with user needs [9]. And the information system is made web-based, where the website is a collection of publicly accessible and interrelated document pages that share one domain name [10].

With this information system, it is hoped that it can overcome the problems that exist in Bambu Ungu, therefore the researcher made a final project with the title "Implementation of Data Mining on Sales Data of Bambu Ungu Cafe to Find out Consumer Purchasing Patterns Using the Apriori Algorithm".

2.   METHOD

2.1. Data Source
There are several techniques that can be used in data collection, including the following:
a. Observation, researchers collect sales transaction data by direct observation to the cafe.
b. Interview, researchers dig up clear information from several people who hold the power of the shop and ask directly to customers and managers in the cafe to get information and data - the data we need from the research we do at the cafe restaurant.

The following are the results of observations and interviews which are material to be tested according to research needs, namely as follows:

Table I. Food Menu

Product Code	Product Name	Food Type	Initials
0001	Plain Fried Rice	Main Menu	NGB
0002	Village Fried Rice	Main Menu	NGK
0003	Chicken Fried Rice	Main Menu	NGA
0004	Indomie Kuah	Main Menu	IK
0005	Fried Indomie	Main Menu	IG
0006	Fried Potato	Main Menu	KG
0007	Fried Sausage	Main Menu	SG
0008	Meat Burger	Main Menu	BD
0019	Fried Tofu	Main Menu	TG

0010	Chocolate Milk	Main Menu	SC
0011	Sweet Tea	Main Menu	TM
0012	Capuccino	Main Menu	C
0013	Lemon Tea	Main Menu	LT
0014	Aqua	Main Menu	A

From the research results, this study also obtained transaction data that will be used for the research process with the Apriori method. This data is obtained from the results of observations with sales data collection techniques. The following is the data that has been obtained:

Table II. Transaction Data

Date	NGB	NGA	NGK	IK	IG	KG	SG	BD	TG	SC	TM	C	LT	A
June/1	2	3	1	3	6	5	2	2	3	5	6	6	2	15
June/2	4	6	3	7	7	5	2	4	3	7	5	8	4	20
June/3	2	2	1	3	2	2	1	0	0	2	3	3	6	14
June/4	0	1	1	3	3	5	2	1	2	3	4	5	2	7
June/5	1	1	2	2	1	3	2	1	0	3	3	5	0	4
June/6	2	3	2	4	4	3	2	1	2	4	4	5	4	8
June/7	0	1	0	2	3	3	1	0	0	3	4	6	3	3
June/8	1	2	2	2	2	1	2	3	1	5	2	5	2	4
June/9	3	4	3	7	9	3	2	1	2	5	5	8	4	15
June/10	1	3	2	4	4	2	3	3	3	5	6	5	4	14
June/11	1	3	4	3	4	5	2	1	1	5	4	7	3	6
June/12	0	1	1	2	6	1	0	0	1	3	3	3	3	8
June/13	1	2	2	6	6	5	2	0	2	5	5	5	5	8
June/14	2	0	1	3	5	2	3	2	2	6	8	3	5	4
June/15	3	3	3	7	4	2	1	4	2	7	9	3	4	8
June/16	3	4	3	5	8	3	2	1	2	5	5	8	4	13
June/17	1	3	2	8	3	2	3	3	3	5	7	5	4	10
June/18	2	1	2	4	4	2	2	4	4	6	7	4	1	8
June/19	0	2	0	4	2	2	1	3	1	8	8	4	3	4
June/20	1	3	3	5	7	1	0	0	1	5	5	4	2	6
June/21	1	2	0	2	3	6	1	1	7	8	6	8	2	6
June/22	2	2	2	3	5	2	1	1	5	5	7	3	3	8
June/23	2	2	1	3	2	2	1	0	0	2	3	3	6	14
June/24	4	6	3	7	7	5	2	4	3	7	5	8	4	20
June/25	2	2	1	2	2	1	2	1	0	4	5	3	2	4
June/26	0	0	2	3	4	1	0	0	2	3	3	2	3	4
June/27	1	2	0	2	2	1	1	0	0	3	2	2	2	3
June/28	2	3	1	1	5	1	1	2	1	4	4	4	1	8
June/29	2	2	2	4	7	2	2	2	3	2	7	8	3	9
June/30	2	4	2	3	4	4	2	1	4	5	6	5	5	8
	48	**73**	**52**	**114**	**131**	**82**	**48**	**46**	**60**	**140**	**151**	**148**	**96**	**263**

2.2. Problem Topic Analysis

The method used by researchers in this study is using the waterfall method. This method is a method for the development stage or better known as the traditional method or classic method.
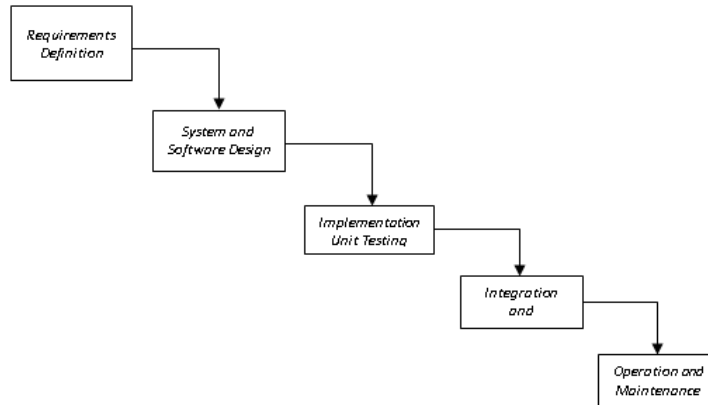
Figure I. Waterfall Method

a. Analysis: Bambu Ungu restaurant transaction data was analyzed to understand customer buying patterns.
b. Design: A sales prediction system was designed using various diagrams and database structures.
c. Implementation: The sales prediction system was programmed and implemented at Bambu Ungu restaurant.
d. Testing: The system was tested to ensure its functionality and accuracy.
e. Maintenance: The system is monitored and updated regularly to ensure its smooth operation.

3. RESULTS AND DISCUSSION

a. Data Mining Calculation

Steps to create a data mining model using the apriori algorithm: First, determine what data you want to process. Second, determine the minimum support and minimum confidence values. Third, compile the association rules formed.

Table III. Transaction Table

No	Items
1	Plain Fried Rice, Aqua, Sweet Tea
2	Chicken Fried Rice, Aqua,
3	Plain Fried Rice, Aqua
4	Village Fried Rice, Aqua
5	Chicken Fried Rice, Aqua, Sweet Tea
6	Chicken Fried Rice, Aqua, Sweet Tea
7	Indomie soup, Aqua
8	Indomie soup, Aqua, sweet tea, fried tofu
9	Indomie soup, Aqua
10	Fried Indomie, Aqua, Sweet Tea
11	Fried Indomie, Aqua
12	Fried Indomie, Aqua, Capuccino
13	Fried Indomie, Aqua
14	Fried Indomie, Aqua, Meat Burger
15	Fried Indomie, Aqua, Chocolate Milk, Fried Tofu
16	Lemon Tea, French Fries
17	Lemon Tea, Fries, Chocolate Milk
18	Capuccino, Fries, Chocolate Milk, Meat Burger
19	Chocolate Milk, Capuccino
20	Capuccino, , Fried Sausage, Chocolate Milk, Fries
21	Capuccino, Fried Sausage
22	Sweet Tea, Fried Tofu
23	French Fries, Chocolate

From the above transactions, the value of the Product Occurrence Frequency Itemset is obtained in the following table:

Table IV. Tabular Transaction Table

No	Items	Frekuensi Kemunculan
1	Aqua	15
2	Capuccino	6
3	Sweet Tea	6
4	Fried Indomie	6
5	Chocolate Milk	5
6	Fried Potato	5
7	Fried Tofu	3
8	Chicken Fried Rice	3
9	Indomie Kuah	3
10	Fried Sausage	2
11	Meat Burger	2
12	Plain Fried Rice	2
13	Lemon Tea	2
14	Village Fried Rice	1

Then given the minimum support = 20% of 23 sales transactions based on the table above, and will search for the support value of 1 itemset with the following formula:

$$\text{Support (A)} = \frac{\text{number of transactions containing Aqua}}{\text{Total transaction}} \times 100\%$$

$$\text{Support (Aqua)} = \frac{\text{number of transactions containing Aqua}}{\text{Total transaction}} \times 100\%$$

$$\text{Support (Aqua)} = \frac{15}{23} \times 100\% = 65,21\%$$

Then the support value of 1 itemset will be obtained as in the table below:

Table V. Iteration Table 1

No	Items	Support
1	Aqua	15/23x100%= 65,21
2	Capuccino	6/23x100%= 26,08
3	Sweet Tea	6/23x100%= 26,08
4	Fried Indomie	6/23 x100%= 26,08
5	Chocolate Milk	5/23 x100%= 21,73
6	Fried Potato	5/23 x100%= 21,73
7	~~Fried Tofu~~	~~3/23 x100%= 13,04~~
8	~~Chicken Fried Rice~~	~~3/23 x100%= 13,04~~
9	~~Indomie Kuah~~	~~3/23 x100%= 13,04~~
10	~~Fried Sausage~~	~~2/23 x100%= 0,86~~
11	~~Meat Burger~~	~~2/23 x100%= 0,86~~
12	~~Plain Fried Rice~~	~~2/23 x100%= 0,86~~
13	~~Lemon Tea~~	~~2/23 x100%= 0,86~~
14	~~Village Fried Rice~~	~~1/23 x100%= 0,43~~

Items with a support value obtained by setting the minimum support = 20%, the items that meet the minimum support value = 20% can be seen in the following figure:

Table VI. Iteration 1 That Meets the Minimum Support Value

No	Items	Support
1	Aqua	65,21%
2	Capuccino	26,08%
3	Sweet Tea	26,08%
4	Fried Indomie	26,08%
5	Chocolate Milk	21,73%
6	Fried Potato	21,73%

Then look for the Support 2 Itemset value using the following formula:

$$\text{Support (A)} = \frac{\text{number of transactions containing AB}}{\text{Total transaction}} \times 100\ \%$$

As an example of applying formulas to data:

$$\text{Support (Aqua, Capucino)} = \frac{\text{number of transactions containing Aqua,Capucino}}{\text{Total transaction}} \times 100\ \%$$

$$\text{Support (Aqua, Capucino)} = \frac{1}{23} \times 100\ \% = 0{,}4\ \%$$

For the calculation results of each 2 itemset can be seen in the table below:

Table VII. Iteration 2

No	Items	FK	Support
1	~~Aqua, Capuccino~~	~~1~~	~~1/23 x 100 = 0,4 %~~
2	Aqua, Sweet Tea	5	5/23 x 100 = 21,7 %
3	Aqua, Fried Indomie	6	6/23 x 100 = 26 %
4	~~Aqua, Chocolate Milk~~	~~1~~	~~1/23 x 100 = 0,4 %~~
5	~~Aqua, French Fries~~	~~0~~	~~0/23 x 100 = 0 %~~
6	~~Capuccino, Sweet Tea~~	~~0~~	~~0/23 x 100 = 0 %~~
7	~~Capuccino, Fried Indomie~~	~~1~~	~~1/23 x 100 =0,4 %~~
8	~~Capuccino, Chocolate Milk~~	~~3~~	~~3/23 x 100 = 13 %~~
9	~~Capuccino, French Fries~~	~~3~~	~~3/23 x 100 = 13 %~~
10	~~Sweet Tea, Fried Indomie~~	~~1~~	~~1/23 x 100 =0,4 %~~
11	~~Sweet Tea, Chocolate Milk~~	~~0~~	~~0/23 x 100 = 0 %~~
12	~~Sweet Tea, French Fries~~	~~0~~	~~0/23 x 100 = 0 %~~
13	~~Fried Indomie, Chocolate Milk~~	~~1~~	~~1/23 x 100 =0,4 %~~
14	~~Fried Indomie, Fried Potato~~	~~0~~	~~0/23 x 100 = 0 %~~
15	~~Chocolate Milk, French Fries~~	~~3~~	~~3/23 x 100 = 13 %~~

Then the items that meet the minimum support value = 20% can be seen in the following figure:

Table VIII. Iteration 2 That Meet the Minimum Support Value

No	Items	Support
1	Aqua, Sweet Tea	21,7 %
2	Aqua, Fried Indomie	26,08%

Next, find the Support 3 Itemset value using the following formula:

$$\text{Support (Aqua, Sweet Tea, Fried Indomie)} = \frac{1}{23} \times 100\ \% = 0{,}4\ \%$$

For the calculation results of each 3 itemset can be seen in the table below:

Table IX. Iteration 3 That Does Not Meet the Minimum Support Value

No	Items	Support
1	~~Aqua, Sweet Tea, Fried Indomie~~	~~1/23x100% = 0,4 %~~

Next, find the confidance value of each combination of the 2 itemsets using the following formula:

$$\text{Support (A∩B)} = \frac{\text{number of transactions containing A∩B}}{\text{Total transactionA}} \times 100\ \%$$

Table X. Confidence Value 2 Itemset

No	Items	Confidence
1	~~Aqua → Sweet Tea~~	~~21,7/65,21 x 100 = 33,2%~~
	Sweet Tea→ Aqua	21,7/26,08 x 100= 83,2 %
2	~~Aqua → Fried Indomie~~	~~26,08/65,21 x 100= 39,9 %~~
	Fried Indomie → Aqua	26,08/26,08 x 100= 100%

Furthermore, by setting the minimum confidance value = 80% and support value = 20%, the association rules formed are as follows:

Table XI. Aturan Asosiasi (Association Rule)

No	Items	Support	Confident
1	Sweet Tea → Aqua	21,7 %	83,2 %
2	Fried Indomie → Aqua	26%	100%

Next, find the lift ratio value using the following formula:

$$\text{Lift Ratio} = \frac{\text{Support Sweet Tea, Aqua}}{\text{Support Sweet Tea} * \text{Support Aqua}} \times 100\% = \frac{21,7}{26,08 * 65,21} \times 100\% = 1,27$$

$$\text{Lift Ratio} = \frac{\text{Support Fried Indomie, Aqua}}{\text{Support Fried Indomie} * \text{Support Aqua}} \times 100\% = \frac{26,08}{26,08 * 65,21} \times 100\% = 1,5$$

Based on the results obtained, the following rules are formed: If consumers buy {Sweet Tea} then they will buy {Aqua} with a confidence level of 83%. If consumers buy {Indomie Goreng} then they will buy {Aqua} with a confidence level of → 100%.
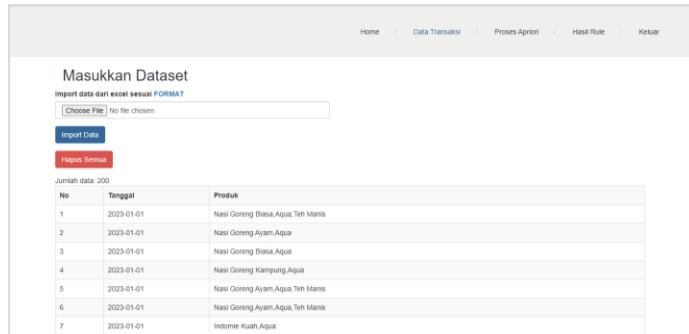
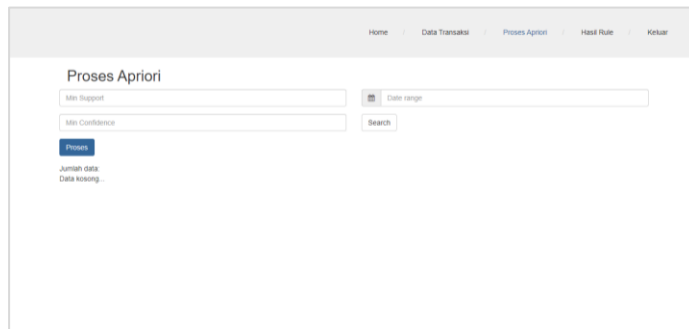b. Implementation

Figure II. Transaction Data Display

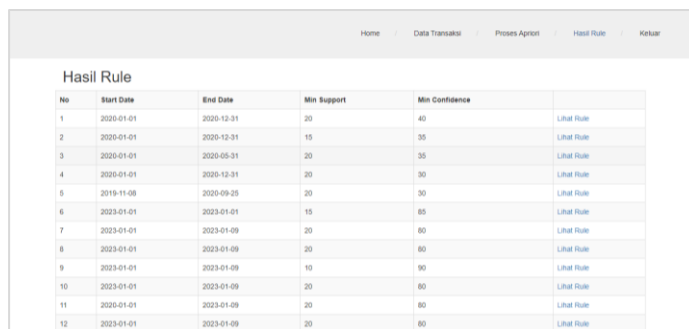**Figure III. Apriori Process View**

Figure IV. Rule Result Display

4. CONCLUSION

After carrying out all stages from design, implementation and testing, the researcher provides the following conclusions:

1. That the Apriori Algorithm data mining method can find food combination patterns with association rules by calculating the Support and Confidence values on the food menu.
2. Application of data mining with the Apriori algorithm in analyzing goods or products to find sales patterns for interconnected food menus in the form of rules / rules. and can find out the percentage of sales on each product item.
3. By getting the percentage of sales on each product item using the apriori algorithm, it will be able to determine which menu items will be promoted.

REFERENCES

[1] E. A. Pyanikova, A. E. Kovaleva, S. I. Galchenko, S. N. Kobchenko, E. V. Ovchinnikova, and M. B. Pikalova, "Digitalization of Restaurant Business as a Factor of Competitiveness Increase," Aug. 2020, doi: 10.2991/AEBMR.K.200730.100.

[2] R. Syuhada, F. Ruziq, and E. Purwawijaya, "Pelatihan Aplikasi Pembuatan Ujian Online Berbasis Android untuk Guru Sekolah Menengah Pertama," *Jurnal Abdidas*, vol. 2, no. 3, pp. 465–471, Jun. 2021, doi: 10.31004/ABDIDAS.V2I3.298.

[3] M. Sholeh, D. Andayati, R. Yuliana Rachmawati, P. Studi Informatika, and F. Teknologi Informasi dan Bisnis, "DATA MINING MODEL KLASIFIKASI MENGGUNAKAN ALGORITMA K-NEAREST NEIGHBOR DENGAN NORMALISASI UNTUK PREDIKSI PENYAKIT DIABETES," *TeIKa*, vol. 12, no. 02, pp. 77–87, Oct. 2022, doi: 10.36342/teika.v12i02.2911.

[4] E. Purwawijaya, F. Soufitri, F. Ruziq, and D. Syahputra, "Application Of The Expert System Of Certainty Factor Methods In Diagnosing Thyroid Disease | INFOKUM," Infokum. Accessed: Jun. 29, 2024. [Online]. Available: https://infor.seaninstitute.org/index.php/infokum/article/view/282

[5] F. Ruziq, M. Rhifky Wayahdi, P. Studi, S. Informasi, F. Teknologi, and U. Battuta, "Sistem Pendukung Keputusan Seleksi Karyawan Baru dengan Simple Additive Weighting pada PT. Technology Laboratories Indonesia," *Jurnal Minfo Polgan*, vol. 11, no. 2, pp. 153–159, Aug. 2022, doi: 10.33395/JMP.V11I2.13506.

[6] M. Rhifky Wayahdi, F. Ruziq, and M. R. Wayahdi, "KNN and XGBoost Algorithms for Lung Cancer Prediction," *Journal of Science Technology (JoSTec)*, vol. 4, no. 1, pp. 179–186, Dec. 2022, doi: 10.55299/JOSTEC.V4I1.251.

[7] M. R. Wayahdi and F. Ruziq, "Implementasi Metode WASPAS pada Sistem Penerimaan Anggota Baru," *Jurnal Minfo Polgan*, vol. 13, no. 1, pp. 164–171, Feb. 2024, doi: 10.33395/JMP.V13I1.13504.

[8] U. S. J. Jibril, "MARKET BASKET ANALYSIS DENGAN ALGORITMA APRIORI MENGGUNAKAN ASSOCIATION RULE UNTUK PENJUALAN SPARE PART ALAT BERAT PADA CV. SENTOSA SUKSES ABADI = MARKET BASKET ANALYSIS WITH APRIORI ALGORITHM USING ASSOCIATION RULE FOR HEAVY EQUIPMENT SPARE PART SALES AT CV. SENTOSA SUKSES ABADI," Feb. 2022.

[9] M. R. Wayahdi and F. Ruziq, "Pemodelan Sistem Penerimaan Anggota Baru dengan Unified Modeling Language (UML) (Studi Kasus: Programmer Association of Battuta)," *Jurnal Minfo Polgan*, vol. 12, no. 1, pp. 1514–1521, Aug. 2023, doi: 10.33395/JMP.V12I1.12870.

[10] M. R. Wayahdi, S. H. N. Ginting, and F. Ruziq, "Pelatihan Membangun Website Portofolio Menggunakan Bootstrap V5.3 Pada Siswa/I SMK Swasta Jambi Medan," *PRAXIS: Jurnal Pengabdian kepada Masyarakat*, vol. 2, no. 1, pp. 86–94, Sep. 2023, doi: 10.47776/PRAXIS.V2I1.715.