

FP_Tree Patterns of Using Social Media (Social Media) in E-Commerce Transactions

¹Lismardiana Sembiring, ²Anjar Pinem, ³Jimmy Perangiangin

^{1, 2, 3}Teknik Informatika Universitas Mandiri Bina Prestasi Medan, Indonesia

Email: ¹lismardianasembiring@gmail.com, ²Anjar.pinem@gmail.com, ³jimmy.mbp@gmail.com

Keywords

FP_Growth, Rule Mining Association, ecommerce, Social Media.

Abstract. The needed for the general public to shop online through e-commerce has become a trend since the start of Covid-19. Social media has become a daily meal for the general public, from teenagers to parents and even the elderly, on average they already have an account and are fluent in social media, even an individual can have two or three social media accounts such as; Instagram, Facebook, Tiktok and Twitter. Of all people in Indonesia who have used social media, on average they have shopped online. Several research have also done the FP_Growth method and the Rule Mining Association in increasing sales turnover and marketing products using social media. The FP_Growth concept and the Rule mining association will produce a prequent item or set item in online shopping transactions using social media accounts. From the results of the prequent item or set item in online shopping transactions, the purpose to determine the mapping pattern of using social media in online shopping. This is useful as a reference for B2B, B2C, C2C, C2B e-commerce to increase the sales turnover of their e-commerce products by developing an electronic marketing strategy so that they are able to compete with other e-commerce actors. The mapping pattern has a min support group of 22% (A4, A5, D4, D6); 30% (A4, A5, D4, D6, A7, B4, B7), a group of 60% min support (A6, B5, B6, C7, D7, C6) and min group conf 27% (A6, B5, B6, C7, D7); 38% (C6); conf 16% (A7, B4, B7)

1. INTRODUCTION

The growth of e-commerce is growing rapidly in line with people's need to use social media for online shopping. Since the start of Covid-19, people have been required to stay at home to carry out all activities such as: WFH (Work From Home), shopping for kitchen and household needs, [1] fashion shopping from home and online learning for elementary, middle and high school children. High school to college level is also done from home. Even though Covid-19 has ended in Indonesia everywhere, people are no longer encouraged to wear masks, and are free to gather in crowds, but the habit of shopping online still cannot be eradicated, as it continues to this day. Likewise, in the use of social media facilities, people are very fluent in social media, both young people, teenagers, parents and the elderly, on average each of them has a social media account. There are even individuals who have more than one social media account, two or even three, such as: Facebook, Instagram, TikTok, Twitter. The use of social media is also used by them when shopping online. So we increasingly see that social media has been used as an online business venue. E-commerce players are increasingly active in marketing their products on social media. According to tempo.co data, this online shopping trend is being exploited by e-commerce parties who project that e-commerce growth in Indonesia will increase by more than 40% in 2021, as shown in the graph below:



Figure 1. 2021 ecommerce transactions

<https://data.tempo.co/data/1070/proyeksi-transaksi-e-commerce-2021>

Based on Figure 1, it states that the growth of e-commerce transactions in 2021 will increase, of course, with this, e-commerce is the main driver of increasing Indonesia's digital economic turnover. With the increase in digital economic turnover, the use of social media and social e-commerce has joined in to fully support the increasing growth of online transactions, such as the use of social media: Instagram, Facebook, Twitter, opening up opportunities for B2B, B2C, C2B, C2C e-commerce players in promoting products. they.

From the above background, there are several previous studies on e-commerce research such as: Srisadono and Wahyu [2] conducted research on the use of social media in increasing sales turnover in e-commerce; Hartawan et al[3] researched the influence of IG social media advertising on people's buying interest in e-commerce; This supports the statement that with the increasing use of social media in online shopping, new e-commerce models have emerged which have become competition for e-commerce business actors.

The aim of this research is to map [4] the relationship between the number of employees tasked with serving consumers in online transactions, and the use of social media in online shopping which is most preferred by the e-commerce community. The results of this mapping are used as a benchmark for B2B (Business to Business), B2C (Business to Customer), C2B (Customer to Business), C2C (Customer to Customer) e-commerce players in designing future marketing strategies and as a reference for them. to maintain online business so that they are able to compete with other e-commerce business actors in increasing sales of e-commerce products.[5]

2. METHOD

The research stages were carried out using data collection techniques first, then continued with data pre-processing, after that straight to the data transformation stage as well as data selection and application of the FP_Growth algorithm accompanied by FI_Tree and finally to the Association Rule Mining stage which can be depicted as in Figure 1 below:

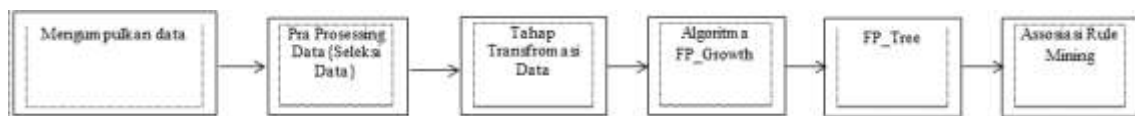


Figure 2. Research Process Flow

The Data Mining process is carried out in 3 stages, that is:

- Data collection is taken from : <https://selular.id/2021/11/top-10-marketplace-di-indonesia-q3-2021/>
- Data selection and data transformation: Determining Data. The criteria used are; Online Store Name, Twitter, Facebook, Instagram and Number of employees

FP_Growth algorithm

The FP_Growth algorithm is used to calculate data sets that frequently appear (Frequent ItemSet) using 3 steps, that is:

- Generation of Conditional Pattern Base
- Generation of Conditional FP_Tree
- Frequent Itemset Search

[5] The tree-based shopping cart analysis algorithm is very effective in completing the mining task only requiring two database scans. The first scan calculates the support for each item. This scan also creates a table header, recording the item name and the corresponding support value as well as the node-link that connects to the first node in the FP_Tree with the same item name. The support values in the table header are sorted from largest value to smallest value. Items that have a support value below the threshold are removed (filtered). Scans of the two remaining items are sorted based on their support values and then entered into the FP_Tree. The FP_Tree structure contains a root node labeled as null, a set of sub tree item-prefixes as the children of root, and a header table.

Structure this node with the same path from the root, and node-link is a pointer that connects to this node with the same path from the root, and node-link is a pointer that connects to the next node in

the FP_Tree. Fp_Tree is a data storage structure built to map each transaction data into each specific path. FP_tree is used to search for Prequent Patterns with a minimum threshold of 60% and minimum support and minimum confidence of 70%. FP_Growth construction is as shown in Figure 3 below:

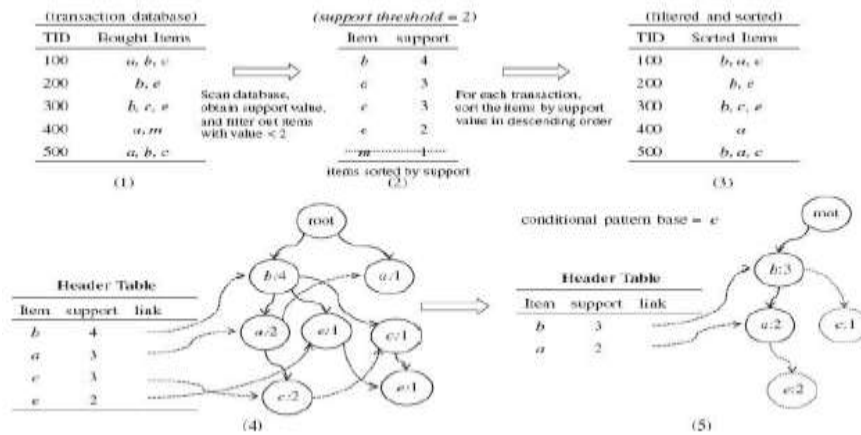


Figure 4: FP_Tree construction [5]

Rule Mining Association

Association Rule mining is a technique for finding relationships between items in a specified data set. When determining a relationship, there is a measure of trust (interestingness measure), namely [6] Support is a measure that shows how much level of dominance an item or itemset has over the entire transaction. This measure is used to determine whether an item or itemset is worthy of seeking confidence.[4] Confidence is a measure that shows the relationship between two items or itemsets conditionally with the support and confidence value formula as follows:

$$\text{Support (A} \rightarrow \text{B)} = \frac{\text{Count of Transactions Contains A and B}}{\text{Amount of Transactions}}$$

$$\text{Confidence (A} \rightarrow \text{B)} = \frac{\text{Count of Transactions Contains A and B}}{\text{Count of Transactions Contains A}}$$

Gambar 4. Assosiasi Rule Mining[5]

3. RESULTS AND DISCUSSION

Data Collection From the results of data collection <https://selular.id/2021/11/top-10-marketplace-di-indonesia-q3-2021> In 2021, the number of visitors will be 1,020 respondents aged 18-60 years who will make online transactions using social media from all e-commerce players in Indonesia based on the data obtained in the table below:

Table 1: Display of e-commerce transaction competition data Using Social Media

NO	Online Store Name	Monthly Web Visitors	Ranking AppStroe	Ranking Playstore	Twitter	Instagram	Facebook	Number of employees
1	Tokopedia	147.790.000	#2	#4	853.000	3.828.300	6.525.650	4.944
2	Shopee	126.996.700	#1	#1	603.800	7.757.940	21.855.970	12.192
3	Bukalapak	29.460.000	#6	#5	215.600	1.661.140	2.518.990	2.316
4	Lazada	27.670.000	#3	#2	430.000	2.975.370	31.364.410	4.126
5	Blibli	18.440.000	#8	#7	529.600	1.622.480	8.598.260	1.979
6	Bhineka	6.996.700	#21	#17	67.100	42.280	1.036.230	487

7	Orami	6.260.000	n/a	n/a	5.820	6.040	351.770	211
8	Ralali	5.123.300	#26	n/a	2.880	41.160	91.390	176
9	JD.ID	3.026.700	#7	#6	54.000	641.740	999.050	1.330
10	Zalora	3.366.700	#5	#8	6.850	1.028.750	18.050	580
11	Sociolla	2486700	#4	#3	5.560	1.008.850	16.570	561
12	Matahari	1175800	#10	n/a	92.800	1.700.000	1.571.370	512
13	Jakarta Notebook	928600	#16	n/a	10.300	43.790	45.510	70
14	Jakmall	835800	#17	#10	3.690	53.000	99.780	71
15	Fabelio	650700	n/a	n/a	740	293.060	87.560	287
16	Mapemall	589600	#14	#19	9.360	647.000	204.660	3.500
17	Elevenia	461200	#24	#15	114.000	105.030	1.115.820	156
18	Otten Coffee	410100	#23	#22	7.820	450.320	663.620	87
19	Jam Tangan	408000	#13	#11	1.800	n/a	362.380	n/a
20	iStyle	391100	#15	n/a	2.590	220.000	58.610	n/a
21	Alfacart	374100	n/a	#9	7.820	59.500	n/a	86
22	Shepora	264200	#9	#13	3370	467.720	19.482.930	118
23	Laku6	239700	#18	#16	n/a	48.220	117.930	n/a
24	My Hartono	231500	#11	#12	380	186.160	68.430	n/a
25	Orori	202900	n/a	n/a	2400	46.120	225830	23
26	PlazaKamera	165000	n/a	n/a	n/a	144.490	43.130	n/a
27	Hijup	156600	#27	#20	54.900	1.184.540	309.210	90
28	Mothercare	152100	n/a	n/a	26.800	662.010	154.870	20
29	Favo	143000	#29	#24	430	34.790	n/a	25
30	Bro.do	106200	n/a	n/a	21.300	871.660	1.217.610	74
31	Pemmz	101500	n/a	n/a	1.320	29.660	30.110	17
32	Dinomarket	101500	#25	#23	33.400	44.880	41.110	32
33	Berrybenka	89100	#12	#18	14.800	634.220	920.160	184
34	Electronic City	87400	n/a	n/a	44.100	39.790	210.880	578
35	Bobobobo	83800	n/a	n/a	3.350	147.100	219.450	37
36	Qoo10	44000	#22	#21	n/a	2.100	532.050	36
37	Tees	26300	n/a	n/a	9.160	7.090	54.240	n/a
38	Hjabenka	19700	#20	#25	2.410	626.670	744.730	184
39	Sorabel	14800	#19	#14	13.400	1.001.060	4.276.500	81

From table 1 above there are fields displayed: online shop name, Twitter social media; Instagram social media; [7] Facebook social media and number of employees. After that, we will continue with preprocessing the itemset data. The algorithm that will be used is fp-growth. This algorithm has a basic knowledge of previously known frequent itemsets to process further information. In the Fp-growth algorithm there are three stages carried out, but before entering this stage, to make it easier to form frequent itemsets, the fields will be coded first. This coding process is based on the initials of each field code contained in the transaction. The following is the coding of each field which can be seen in the table below:

Table 2. Social media field coding results

Twitter	IG	FB	Number of employees
A1= >=603.000	B1= >=7.000.000	C1= >=31.000.000	D1= .->12.000
A2= >=529.000	B2= >=3.000.000	C2= >=21.000.000	D2= >=3000
A3= >=430.000	B3= >=2.000.000	C3= >=18.000.000	D3= >=2000
A4= >=114.000	B4= >=1.000.000	C4= >=8.000.000	D4= >=1000

Twitter	IG	FB	Number of employees
A5= >=67.000	B5= >=600.000	C5= .>=6.000.000	D5= >=500
A6= >=20.000	B6= <=600.000	C6= >=1.000.000	D6 <=500
A7= <=20.000	B7= <=60.000	C7= <=1.000.000	D7= <=100

The Fp-growth

Algorithm is an algorithm that will determine frequent itemsets. The advantage of Fp-growth is that it only requires two transaction data scans and is very efficient in scanning data. The data used in this stage is data on the use of social media in e-commerce transactions. The data cleaning process uses Weka Explorer. can be seen in the following table display: [8]

Table 3. Display of transaction data preprocessing results
 ecommerce uses social media

NO.	Online name store	Twitter	Instagram	Facebook	Number of employees
1	Tokopedia	853.000	3.828.300	6.525.650	4.944
2	Shopee	603.800	7.757.940	21.855.970	12.192
3	Bukalapak	215.600	1.661.140	2.518.990	2.316
4	Lazada	430.000	2.975.370	31.364.410	4.126
5	Blibli	529.600	1.622.480	8.598.260	1.979
6	Bhineka	67.100	42.280	1.036.230	487
7	Orami	5.820	6.040	351.770	211
8	Ralali	2.880	41.160	91.390	176
9	JD.ID	54.000	641.740	999.050	1.330
10	Zalora	6.850	1.028.750	18.050	580
11	Sociolla	5.560	1.008.850	16.570	561
12	Matahari	92.800	1.700.000	1.571.370	512
13	Jakarta Notebook	10.300	43.790	45.510	70
14	Jakmall	3.690	53.000	99.780	71
15	Fabelio	740	293.060	87.560	287
16	Mapemall	9.360	647.000	204.660	3.500
17	Elevenia	114.000	105.030	1.115.820	156
18	Otten Coffee	7.820	450.320	663.620	87
22	Shepora	3370	467.720	19.482.930	118
25	Orori	2400	46.120	225830	23
27	Hijup	54.900	1.184.540	309.210	90
28	Mothercare	26.800	662.010	154.870	20
30	Bro.do	21.300	871.660	1.217.610	74
31	Pemzm	1.320	29.660	30.110	17
32	Dinomarket	33.400	44.880	41.110	32
33	Berrybenka	14.800	634.220	920.160	184
34	Electronic City	44.100	39.790	210.880	578
35	Bobobobo	3.350	147.100	219.450	37
38	Hjabenka	2.410	626.670	744.730	184
39	Sorabel	13.400	1.001.060	4.276.500	81

This data mining process [9] will be used to produce frequent itemsets from each social media usage transaction from each e-commerce actor, where the results of the highest frequency of occurrence of each field code from the data on the number of sales transactions for each e-commerce actor can be seen for any variable. which can be processed in the next stage, [10] as in the data transformation i the table below:

Table.4 Data Transformation

TID1-5	=	A1,A2,A3,A4,B1,B2,B3,B4,C1,C2,C4,C5,C6,D1,D2,D3,D4
TID5-10	=	A1,A2,A5,A6,A7,B4,B5,B7,C6,C7,D4,D6
TID10-15	=	A5,A6,A7,B4,B6,B7,C6,C7,D5,D7
TID15-20	=	A4,A7,B5,B6,B7,C3,C6,D6,D7
TID20-25	=	A6,A7,B4,B5,B6,B7,C6,C7,D7
TID25-30	=	A6,A7,B4,B5,B6,B7,C6,C7,D7

To search for patterns from the selected transaction data, the Fp_Growth Algorithm stage is carried out. Constructs a conditional FP-Tree and TID List, then extracts a combination of items (frequent itemsets). From the FP_Tree above, there are several items/itemsets that can be seen on the FP_Tree node line which [11] is the results of generating a Frequent Itemset in TID records are in table 3. From the FP_Tree above, it is concluded that the items/itemsets that produce Frequent Items/Itemsets can be seen in table 4 below:

Table 5. Frequent Items

Twitter	IG	FB	Number of employees
A1	2	B1	1
A2	2	B2	1
A3	1	B3	1
A4	2	B4	6
A5	2	B5	5
A6	5	B6	5
A7	6	B7	6

From table 4 above, it is explained that the results of frequent items: A1, B1, B2, B3, C1, C2, C3, C4, C5, D1, D2, D3 and D5 were declared disqualified and were not included in the nomination because the frequent items were worth 1 frequent occurrence in each transaction. [10] To see the results of Frequent Items which have been grouped based on the lowest order and up to the highest order, they are as follows:

Table 6. Hasil Frequent Priority

ITEM	FREQUENT
A4	2
A5	2
A6	5
A7	6
B4	6
B5	5
B6	5
B7	6
C6	7
C7	5
D4	2
D6	2
D7	5

In table 5 above, the results of Frequent Item Priority Sorting Results in the FP_Growth process, the results of generating the Frequent Itemset in table 5 above which have a high number of frequents are: C6, B7, B4, A7. So after sorting [12] as in table 6 below:

Table 7. Ordering Frequent Item Priority Results

ITEM	FREQUENT
A4	2
A5	2
D4	2
D6	2

ITEM	FREQUENT
A6	5
B5	5
B6	5
C7	5
D7	5
A7	6
B4	6
B7	6
C6	7

In table 6 above, the Frequent Item Priority Sorting Results in the FP_Growth process are applied to the FP_Tree process. The FP_Tree process [12] is shown in Figure 5 below:

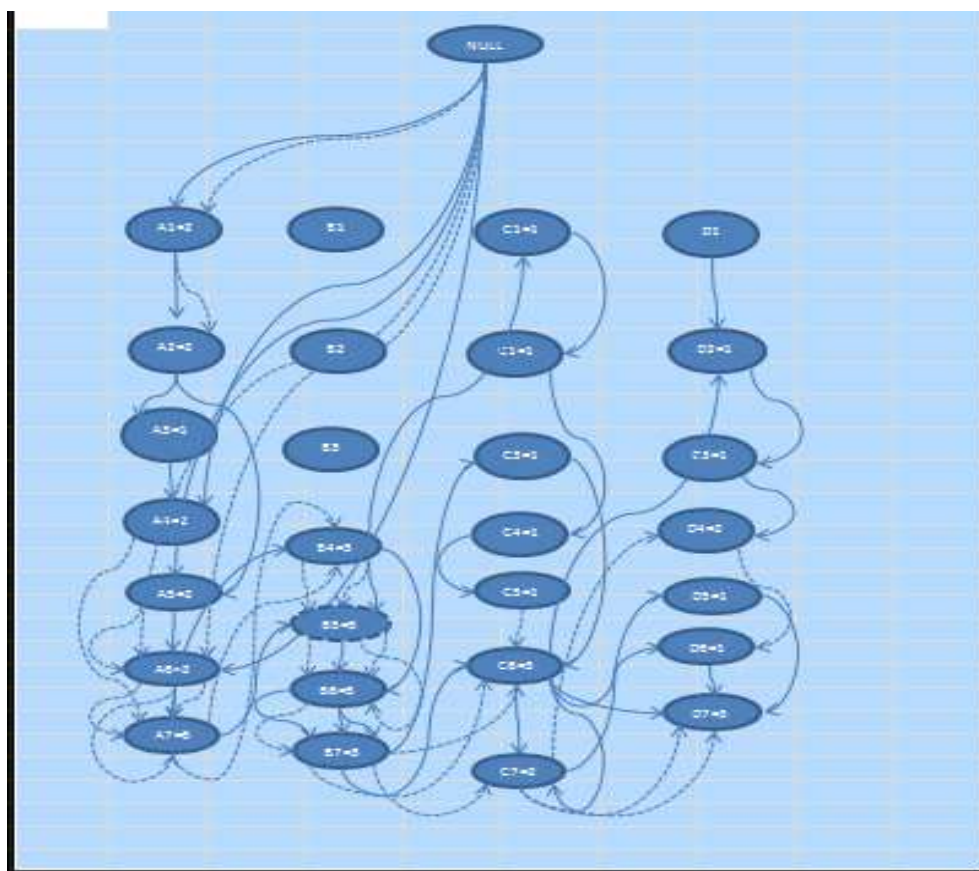


Figure 6. FP_Tree

Rule Mining Association Analysis

Association analysis or association data mining is a method in data mining to find associative rules in combinations of items or relationships between attributes. The results of association rule mining of 60% min support and 70% min confidence are in table 7 below:

Table 8. Rule Mining Association

ITEM	FREQUENT	Min Support 60%	Min Confidence 70%
A4	2	30%	22%
A5	2	30%	22%
D4	2	30%	22%
D6	2	30%	22%
A6	5	60%	27%
B5	5	60%	27%

B6	5	60%	27%
C7	5	60%	27%
D7	5	60%	27%
A7	6	30%	16%
B4	6	30%	16%
B7	6	30%	16%
C6	7	60%	38%

The results of association rule mining in table 8 above show several pattern variables that are linked to the number of employees in serving consumers in online shopping transactions, there are several groups, that is:

1. Min.Support section
 - a. For Min. Support 30% is in groups (A4, A5, D4, D6, A7, B4, B7)
 - b. For Min.Support 60% there are groups (A6, B5, B6, C7, D7, C6)
2. Min.Confidence Section
 - a. For Min.Support 22% (A4,A5,D4,D6)
 - b. For Min. Conf. 27% have groups (A6, B5, B6, C7, D7)
 - c. For Conf. 38% are C6
 - d. For Min. Conf. 16% (A7,B4,B7)



Figure 7. Graphic pattern of the relationship between the number of employees and use of social media in e-commerce transactions

In graph 5 above it is explained that the number of competing online transactions is in the data {A4, A5, D4, D6} which has the same percentage number where A is the symbol for Twitter and D4, D6 is the number of employees and {A6, B5, B6, C7} has the same percentage number where A6 = Twitter, B = Instagram and C6 is Facebook with the same number and the highest is item C6, that is the number of employees ≤ 500 employees who provide online sales. [13]

Table 9. Pattern of relationship between number of employees and use of social media in e-commerce transactions

Code	Min Support	Min Confidence	Amount employees	Online Store Name
A4	30%	22%	Twitter>14000	Sociolla,Matahari
A5	30%	22%	Twitter>67000	
D4	30%	22%	Number of employees >1000	Sociolla,Matahari,Febelio
D6	30%	22%	Number of employees <500	
A6	60%	27%	Twitter>20000	Tokopedia,Shopee,Bukalapak,Lazada,Blibli
B5	60%	27%	Instagram>600000	JD.JD,Mapemall
C7	60%	27%	Facebook<1000000	JD.JD,Sociolla,Mapemall,otten coffe,orori,bro.do,berrybenka
D7	60%	27%	Number of employees <100	jakarta notebook,jakmall,otten coffe,alfacart,Orori,Favo,Bobobobo,Qoo10,S

				orabel
A7	30%	16%	Twitter<20000	Fabelio,My Hartono,Jam Tangan,Pemmz
B4	30%	16%	Instagram >1000000	Bukalapak,Blibli,Zalora,Sociolla,Matahari,H ijup,Sorabel
B7	30%	16%	Instagram < 60000	Bhineka,orami, ralali, Jakarta Notebook, Jakmall,Alfacart,laku6,orori,favo,pemmz,din omarket,electronic city,Qoo10,tees
C6	60%	38%	Facebook>=100000 0	bhineka, elevania,Matahari,Bro.do

From the information in table 8, an explanation of the output of e-commerce competition can be seen in the fields of number and name of online stores where the most dominant competition is from the results of min support of 30%, 60% and min confidence of 22%, 27% which compete to drive e-commerce.

4. CONCLUSION

From the mapping results for support values of 30%, 60% and confidence values of 27% and 38%, that the increasing transactions of e-commerce players via social media, influenced by the number of employees whose job is to serve customers when shopping online. For e-commerce players with a support value of 30% and a confidence value of 16%, 22% is recommended to increase the number of employees to serve online transactions on social media,so you can increase turnover Product sale .

REFERENCES

- [1] A. Maulana, N. R. Arjun, F. Akbar, N. Suryanti Ayu, and H. Firmansyah, "Peran E-Commerce di Tengah Pandemi terhadap Gaya Hidup Masyarakat Indonesia Masa Kini," *J. Educ. Technol.*, vol. 1, no. 1, pp. 55–61, 2021, [Online]. Available: <http://jurnalilmiah.org/journal/index.php/jet>
- [2] W. Srisadono, "Strategi Perusahaan E-Commerce Membangun Brand Community Di Media Sosial Dalam Meningkatkan Omset Penjualan," *J. Pustaka Komun.*, vol. 1, no. 1, pp. 167–179, 2018.
- [3] E. Hartawan, D. Liu, M. R. Handoko, G. Evan, and H. Widjojo, "Pengaruh Iklan Di Media Sosial Instagram Terhadap Minat Beli Masyarakat Pada E-Commerce," *JMBI UNSRAT (Jurnal Ilm. Manaj. Bisnis dan Inov. Univ. Sam Ratulangi)*, vol. 8, no. 1, pp. 217–228, 2021, doi: 10.35794/jmbi.v8i1.33853.
- [4] T. Marzuqah, I. Permana, and M. Afdal, "Penerapan Algoritma FP-Growth Pada Pencarian Hubungan Waktu Pembelian Dan Barang yang Dibeli Untuk Strategi Promosi Penjualan," vol. 10, no. 3, 2023, doi: 10.30865/jurikom.v10i3.6347.
- [5] A. Nugroho Susanto Putro and R. Indra Gunawan, "Implementasi Algoritma FP-Growth untuk Strategi Pemasaran Ritel Hidroponik (Studi Kasus : PT. HAB)," *J. Buana Inform.*, vol. 10, no. 1, pp. 11–18, 2019.
- [6] A. R. Wibowo and A. Jananto, "Implementasi Data Mining Metode Asosiasi Algoritma FP-Growth Pada Perusahaan Ritel," *Inspir. J. Teknol. Inf. dan Komun.*, vol. 10, no. 2, p. 200, 2020, doi: 10.35585/inspir.v10i2.2585.
- [7] M. J. Zuroni and H. L. Goh, "Factors Influencing Consumers' Attitude Towards E-Commerce Purchases Through Online Shopping," *Int. J. Humanit. Soc. Sci.*, vol. 2, no. 4, pp. 223–230, 2012.
- [8] S. P. Tamba, A. W. Tan, Y. Gunawan, and A. Andreas, "Penerapan Data Mining Untuk Pembuatan Paket Promosi Penjualan Menggunakan Kombinasi Fp-Tree Dan Tid-List," *J. Tek. Inf. dan Komput.*, vol. 4, no. 2, p. 201, 2021, doi: 10.37600/tekinkom.v4i2.309.
- [9] W. W. Ariestya, W. Supriyatin, and I. Astuti, "Marketing Strategy for the Determination of Staple Consumer Products Using Fp-Growth and Apriori Algorithm," *J. Ilm. Ekon. Bisnis*, vol. 24, no. 3, pp. 225–235, 2019, doi: 10.35760/eb.2019.v24i3.2229.

- [10] S. Lismardiana, T. Thamrin, and T. Wanra, *Data Mining Implementasi Algoritma Apriori dan FP_Growth*. Eureka Media Aksara, 2022.
- [11] B. Mulyawan, Vionelsy, and T. Sutrisno, "Product recommendation system on building materials shopping using FP-Growth algorithm," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 1007, no. 1, 2020, doi: 10.1088/1757-899X/1007/1/012144.
- [12] L. Yuan, W. Itwee, and S. Wei, "Association Rule Mining Technique with Optimized FP-Tree Algorithm," 2017.
- [13] P. Veleva, "Internet Advertising - a Marketing Tool Supporting E-Commerce," *Trakia J. Sci.*, vol. 17, no. Suppl.1, pp. 883–890, 2019, doi: 10.15547/tjs.2019.s.01.145.