International Journal Of Artificial Intelegence Research
Vol 6, No 1, June 2022
ISSN: 2579-7298

Determination Potential Experts by Application The Apriori Algorithm and the K-Means Algorithm

Rini Sovia[a,1*], Sarjon Defit[a,2]

[a] *Doctoral Program of Information Technology, Faculty of Computer Science, Putra Indonesia University "YPTK" Padang*
[1]*rini_sovia@upiyptk.ac.id.* [2] *sarjon_defit@upiyptk.ac.id*
** corresponding author*

ARTICLE INFO

Article history:
Received 22 May 2021
Revised 25 Ags 2021
Accepted 28 Jan 2022

Keywords:
Prediction
Experts
Projects
Apriori
K-Means

ABSTRACT

Experts are people who have special expertise who provide services based on their expertise. The company has experts in handling projects that will be carried out for the progress of the company. The importance of the quality of experts in the company can improve the quality of human resources. The Apriori algorithm is a data mining method that has the aim of looking for association patterns based on the project being carried out so that they can be identified by experts who are often used in handling projects. Furthermore, a data mining approach is needed to classify experts with the K-means algorithm used. This study combines the Apriori and K-means algorithms, by grouping experts based on the handling of the project they are working on.

I. Introduction

Experts are needed in handling a project. In increasing the pace of development in Indonesia, potential experts are needed. The need for quality human resources is urgent to produce a better future creation in development activities at the regional and national levels. The company applies the concept of selecting experts manually without any support in selecting experts in handling projects that will be carried out in the future. This makes it difficult for companies to choose the right experts to handle the project. On the other hand, the company wants to find experts in the right position in managing the project.

To find out the experts who are often used in consulting services in handling projects will be very difficult if done by manual observation, so the problem that arises is how the company can identify potential experts in the company. To overcome this problem, a more effective data arrangement was carried out by taking information from expert data while working on the project. From existing expert data, a method is needed that can assist in determining potential experts in handling projects.

The method used in this research is the a priori method and K-Means in dealing with this problem. K-means is part of the Clustering method in Data Mining, while Apriori is a method of association rules. With the K-Means method, you can classify data that will produce information, while a priori the data will be selected so that it can produce a conclusion about the ranking of experts who are often used in handling projects. The aim of this research is:

1. Make it easier for the company to analyze and group data in knowing the quality of the performance of experts based on the handling of the project being carried out.
2. Understanding the concept of merging an algorithm Apriori Algorithm K - Means for data on the company's experts.
3. Implement data analysis using the Apriori algorithm and K-Means Algorithm in optimizing the data expert in handling projects using Rapidminer software.
4. Divide data into the cluster / to group together, where the same data characteristics will be grouped into the same cluster, while the data has different characteristics to the other group.
5. Helping managers in choosing the right experts.

Research related to data mining that the author did already much studied by researchers before, including research on the prediction value of the eyes of college students using Algorithm K-Apriori to the value of currency courses that will be taken to consider the value of the eyes of

college and the need to stage the establishment of the rule[1]. Data mining is a procedure by taking techniques statistics, mathematics, artificial intelligence, and machine learning to extract and identify useful information as well as the knowledge that is relevant from a variety of databases[2][3][4][5].

Knowledge Discovery in Database (KDD) is a method used to extract patterns from data consisting of Data integration, Data Selection, Data Transformation, Data Mining, Pattern Evaluation, and Knowledge Presentation [6][7][3]. At the KDD stage, it focuses on the process of processing data analysis patterns, mining archeology of data, extracting information, and integrating business development processes[8][9]. Data mining is also defined as a process to explore additional value from a collection to gain new knowledge that cannot be done manually. Types of methods in data mining can be classified as follows[10].

a. Association Rules
   Association Rules have a function to find associations between variables, correlations, or structures among items in the database. The association rule is a data mining technique that can analyze the market basket of data so that the data is found to be a combination of items that are often used [11].
b. Decision Tree
   The decision tree is a flowchart structure that resembles a tree, which is a decision tree that leads to the resulting solution.
c. Clustering
   Clustering is a method that is often used in heuristic data mining. Clustering is used to categorize different data.

The development of science and technology can encourage institutions to take advantage of technological developments in all fields. One sat u of development that is the system of processing of data very quickly and accurately[12].

Apriori is an algorithm that is known for performing the frequent itemset search process using the association rule technique. The apriori algorithm uses known frequent itemset knowledge to process further information[13][14][15]. A frequent itemset is a set of items that appear simultaneously. The importance of the association process in data mining can be seen in two benchmarks, namely support and confidence [16][17]. The a priori algorithm performs an iterative approach process which is also called level-wise search, the k-itemset is used to explore or find (k + 1) itemset. The a priori property is that each frequent-itemset subset must be a frequent-itemset [18]. The Apriori Algorithm is a well-known algorithm for finding high-frequency patterns[16].

The data mining method can be used to classify the data or Clustering one of them is the algorithm of K-means [19][20]. Algorithm K-means is a category in partitioning clustering that each of the data that exist must be entered in a cluster specified at a stage of the process so that the stages the following switch to the other cluster. cluster. The K - Means algorithm is also known as an algorithm that can classify big data and outlier very fast [2][21]. The K-Means algorithm said also an algorithm of non-hierarchy that is derived from the method of clustering. The Objective of grouping the data that is to minimize the objective which has been set in the process of grouping [22][21][23].

The data is split into two categories in the clustering process, with the data that has similar characteristics clustered together in a cluster, and the data that is distinct in a group that is more[24][25][26]. The method of clustering is a practical method that can be applied in many fields of science, such as intelligence artificial and recognition patterns, marketing, science policy, and others [27][28]. Clustering is a division of data into the group that has the object and the characteristics are the same [29][30].

II. Methods

A. Apriori Process Rules

On research has a hierarchy that is done at the start of the hierarchy of data analysis to results in getting. The hierarchy of the a priori process is shown in Figure 1.

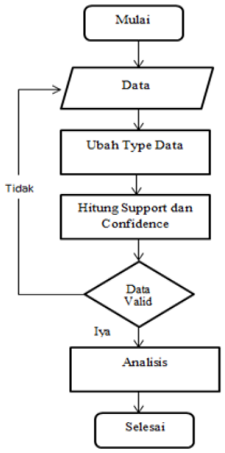

Figure 1 Apriori Process Rules

The process will begin from hierarchy analysis of patterns of frequency high, which is looking for a combination of items to meet the requirements minimum of the value of the support in the database, by using the formula:

$$Support(A) = \frac{\sum transaksi\ yang\ mengandung\ A}{\sum transaksi}\ x\ 100\%$$

If the support value of a combination of more than 1 set item, the value can be obtained using the formula:

$$Support(A,B) = \frac{\sum transaksi\ yang\ mengandung\ A\ dan\ B}{\sum transaksi}\ x\ 100\%$$

The associative rule is formed, after determining the high-frequency pattern is found, then the process of finding the associative rule is met to meet the minimum requirements for the confidence value by calculating the associative confidence rule A → B.

$$Confidence = P(A|B) = \frac{\sum transaksi\ yang\ mengandung\ A\ dan\ B}{\sum transaksi\ A}\ x\ 100\%$$

B. K-Means Process Rules

After the a priori process rule is complete, the next process is to classify the experts based on the clusters that have been determined using the K-means algorithm, so the k-means process is shown in Figure 2

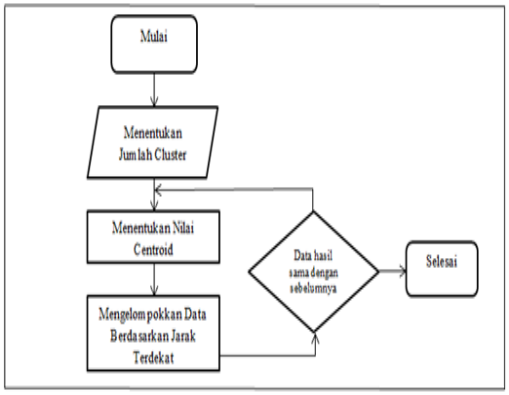

Figure 2 K-Means Process Rules

1) Determine the number of clusters

Determine the number of K clusters that want to be formed to be processed into the system.

2) Calculates distance with Euclidean Distance

Each cluster has each centroid. Calculate the distance of each object with the centroid for each cluster using the Euclidean Distance. The object will be in the group late in the cluster based on the minimum distance to the centroid. With the following formula:

$$d_{ij} = \sqrt{\sum_{k-1}^{p} \{x_{ik} - x_{jk}\}^2}$$

3) Data Clusters

The process on the data cluster will stop if there are no objects that will move to each cluster, Then it will produce a data cluster of experts.

III. Result and Discussion

1) Expert Data

Analysis of patterns of frequency higher than the combination on the data power experts, Data energy experts can visible in Table I below it

Table 1 Expert personel data

Profession	Power Expert
1	Masri zal, Irvan, Budaya, Akhriyansyah, Hafni, Efri zon, Alek , Sudirman , Wilton , Afrizal, Ahmad, Dedi , Nopid , Badrinus , Yandri , Kurniawan, Yusdar , Supriadi nur , M. Iqbal , Yan , Nur
2	Dasmayzal , Sudirman , Yan , Badrinus , Dedi , Dedyanto , Syafrinal , Yandri , Dedi , Ikhsan , M. Agung , Hafni , Akhriyansyah , Zulfadli , Kurniawan, Asnan , Desy , Masrizal
3	Mohammad , Kurniawan , Mohammad , Akhriyansyah
4	Hafni , Wilton, Efrizon , Dasmayzal , Erwin, Effendi
5	Hafni , Isralsyah

2) Expert Data Represents

Data experts in table 1 can represented in Table 2.

Table 2 Expert data representation

Code Power Expert	Power Expert
A	Masrizal
B	Irvan
C	Hafni
D	Efrizon
E	Sudirman
F	Dasmayzal
G.	Wilton
H	Afrizal
I	Dedi
J	Mohammad
K	Yandri
L	Dedyanto

3) Expert Data Tabulation

Expert data in table 1 is formed from tabular tables which will make it easier to find out how many existing items will be used as shown in table 3.

Table 3 Expert tabular data format

Job	Irvan	Isralsyah	Zulfadli	Dasmayzal	Desy
P1	1	0	0	0	0
P2	0	0	1	1	1
P3	0	0	0	0	0
P4	0	0	0	1	0
P5	0	1	0	0	0
P6	0	0	0	0	0
P7	0	1	0	1	0
P8	1	0	0	1	0
P9	0	0	0	0	0

4) Itemset Formation

a) 1 Itemset

Completion of the process that is based on data that exist in table 3 experts tabular format. Based on the high-frequency pattern analysis formula.

1. Support A = 3/9 = 33.33%
2. Support B = 2/9 = 22:22%
3. Support C = 5/9 = 55.55%
4. Support D = 3/9 = 33.33%
5. Support E = 2/9 = 22:22%
6. Support F = 4/9 = 44.44%
7. Support G = 2/9 = 22:22%
8. Support H = 2/9 = 22:22%
9. Support I = 3/9 = 33.33%
10. Support J = 2/9 = 22:22%

b) Combination of 2 Itemset

The process with the formation of 2 itemsets can be seen with tabulated data in table 3 in the analysis below.

1. Support AN 1/9 = 11.11%
2. Support AX 2/9 = 22.22%
3. Support A-Bb 2/9 = 22.22%
4. Support CD 3/9 = 33.33 %
5. Support CF 3/9 = 33.33%
6. Support CI 3/9 = 33.33%
7. Support CK 2/9 = 22.22%
8. Support CM 3/9 = 33.33%
9. Support A-Bb 2/9 = 22.22%
10. Support CD 3/9 = 33.33%

The conclusion of the search ap riori seen in Table 4 below this:

Table 4 Conclusion of the search for apriory algorithm

Item Set 1	Item 2	Support
Hafni		0.5556
Akhriyansyah		0.4444
Dasmayzal		0.4444
Ahmad		0.3333
Dedi		0.3333
Efrizon		0.3333

Kurniawan		0.3333
Masrizal		0.3333
Yandri		0.3333
Hafni	Akhriyansyah	0.3333
Hafni	Efrizon	0.3333
Hafni	Dasmayzal	0.3333
Hafni	Dedi	0.3333
Hafni	Ahmad	0.3333

c) Establishment of Association Rules

After all frequency patterns are found, then look for association rules that meet the requirements. A minimum confidence value of 50% so that it can be formulated by forming association rules.

1. Confidence C- IMBb 33.33% / 55.55% = 0.6 (60%)
2. Confidence I- MBbC 33.33% / 33.33% = 1 (100%)
3. Confidence M- BbCI 33.33% / 33.33% = 1 (100%)
4. Confidence Bb-CIM 33.33% / 44.44% = 0.75 (75%)
5. Confidence IMBb -C 33.33% / 33.33% = 1 (100%)

The results of the search algorithm is a priori Issuer entered into the search algorithm to perform K-Means clustering seen in Table 5

Table 5 Management data of expert that has been grouped

power Expert	Number of E xperts	Support Value
A	1	0.3333
B	1	0.2222
C	1	0.5555
D	1	0.3333
E	1	0.2222
F	1	0.44444
G.	1	0.2222
H	1	0.2222
I	1	0.3333
J	1	0.2222

Table 6 Distance between c1 and c2

C0	C1	C2	CLUSTER
0.2222	2,00308	1,006153	C0
0.1111	2	1	C0
0.2222	2,00308	1,006153	C0
0.2222	2,00308	1,006153	C0
1,006153	1	0	C2
1,006153	1	0	C2
1,006153	1	0	C2
1,006153	1	0	C2
2.003083	0	1	C1
2.003083	0	1	C1

Next do know to categorize an average value based on the shortest distance to the C1 and C2 to obtain an average value group as follows:

$$C_k = \frac{1}{C_K} \sum d_1$$

C0 = {1, 0.2357}
C1 = {3, 0.2121}
C2 = {2, 0.242}

Once the value of the centroid is newly acquired, then compare who value centroid new with the value centroid long, if different, do the calculation distance back by using the new centroid. Results calculated Ungan can be seen in Table 7.

Table 7 cluster results

Cluster	Members Of The Group	Number Of Group Members
C0	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33	33
C1	79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109, 110, 111	33
C2	34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78	45

From the data, it can be concluded that cluster C0 has 33 members in the very good category, in cluster C1 there are 33 people in the normal category, and cluster C2 there are 45 in the good category.

B. Application Implementation to Rapidminer

In this study, the rapidminer 9.2 software was used for the apriori algorithm and k-means clustering processes as shown in Figure 3.

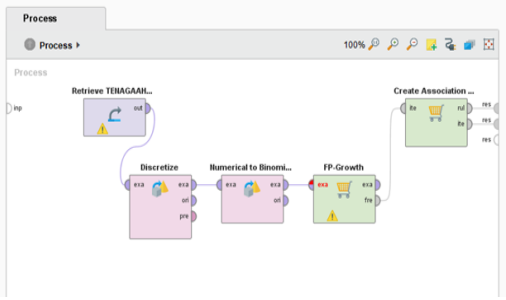

Figure 3 A priori operator model of the association rule algorithm

The results obtained after running the Apriori Algorithm Association Rule operator model are shown in Figure 4.

AssociationRules

```
Association Rules
[Akhriyansyah] --> [DASMAYZAL] (confidence: 0.500)
[DASMAYZAL] --> [Akhriyansyah] (confidence: 0.500)
[Akhriyansyah] --> [EFRIZON_DEDI] (confidence: 0.500)
[Akhriyansyah] --> [MASRIZAL_BASRI] (confidence: 0.500)
[Akhriyansyah] --> [Yandri_Ardi] (confidence: 0.500)
[Akhriyansyah] --> [Afrizal] (confidence: 0.500)
[Akhriyansyah] --> [Badrinus_Nur] (confidence: 0.500)
[Akhriyansyah] --> [Budaya_Rusjdi] (confidence: 0.500)
[Akhriyansyah] --> [Irvan_Afdal] (confidence: 0.500)
[Akhriyansyah] --> [M_AGUNG_RIDIANSYAH] (confidence: 0.500)
[Akhriyansyah] --> [Mohammad_Yusdar] (confidence: 0.500)
[Akhriyansyah] --> [Sudirman] (confidence: 0.500)
[Akhriyansyah] --> [Supriadinur] (confidence: 0.500)
[Akhriyansyah] --> [Yan_Hendrawan] (confidence: 0.500)
[DASMAYZAL] --> [Ahmad_Zaki] (confidence: 0.500)
[DASMAYZAL] --> [Dedi_Iskandar] (confidence: 0.500)
[DASMAYZAL] --> [EFRIZON_DEDI] (confidence: 0.500)
[DASMAYZAL] --> [MASRIZAL_BASRI] (confidence: 0.500)
[DASMAYZAL] --> [ERWIN] (confidence: 0.500)
[DASMAYZAL] --> [M_AGUNG_RIDIANSYAH] (confidence: 0.500)
[Akhriyansyah] --> [Hafni_Hasan, DASMAYZAL] (confidence: 0.500)
[DASMAYZAL] --> [Hafni_Hasan, Akhriyansyah] (confidence: 0.500)
[Akhriyansyah] --> [Hafni_Hasan, EFRIZON_DEDI] (confidence: 0.500)
[Akhriyansyah] --> [Hafni_Hasan, KURNIAWAN DESTIKA] (confidence: 0.500)
```

Figure 4 The results of the a priori operator model association rule algorithm

As a result of the process of association rules, 1607 patterns are connected to experts to handle projects. After the results of the a priori algorithm have been obtained, import the results that have been done to the a priori algorithm, add the K-Means operator. The data is connected with clustering

which is useful for knowing the output towards the result. If it is connected, then determine the number of clusters and the maximum number of runs that will be given, and then click the play button in Figure 5:

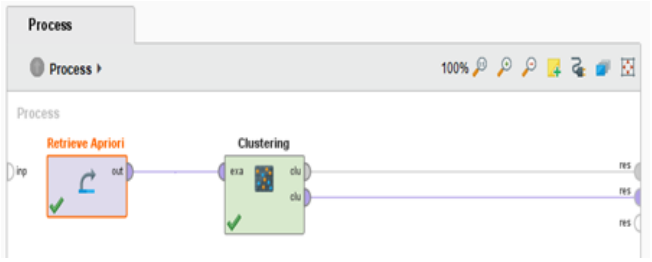

Figure 5 Adding K-Means Operators

The operator at the top will produce the data view to forming clusters of the overall data that conformed to the n data that has been entered can be seen in Figure 6 below it.

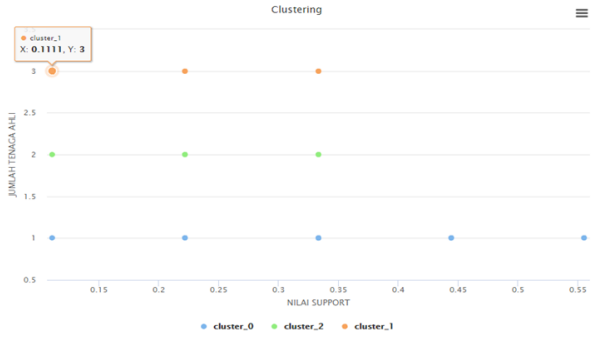

Figure 6 Scatter Display on Chart

In Cluster models, visible are the displays for clusters that have been classified seen in pictures and videos ar 7 below this.

Cluster Model

```
Cluster 0: 33 items
Cluster 1: 33 items
Cluster 2: 45 items
Total number of items: 111
```

Figure 7 Cluster Model View

The results of the analysis both manually and using Rapidminer software have the same analysis values as follows:

1. It has 3 clusters, namely C0 Very Good 33 items, C1 Good 33 items, and C2 Regular 45 items.
2. The data used can be applied to the Apriori algorithm and K-Means.

The application of identification of the relationship between work variables and experts can run well in calculating that is formed using a priori and k-means can obtain information on variables that affect project handling that will be provided to experts by the company.

IV. Conclusion

From the research that has been done on the data of experts in managing projects at the company as follows :

1. Using the Apriori Algorithm and the K-Means Algorithm can group experts based on the handling of the project being carried out.
2. By using company data can understand the application of the a priori algorithm and K-Means.
3. Can implement data analysis using Rapid Miner using the Apriori algorithm and K-Means.

4. Using clustering can classify data based on the characteristics of the existing data.
5. By conducting this research, managers can choose the right personnel to handle projects in the future

References

[1] L. Muflikhah, W. L. Yunita, and M. T. Furqon, "Prediksi Nilai Mata Kuliah Mahasiswa Menggunakan Algoritma K-Apriori," *Sisfo*, vol. 06, no. 02, pp. 157–172, 2017, doi: 10.24089/j.sisfo.2017.01.001.

[2] S. Setiawan, "Analisis Cluster Menggunakan Algoritma K-Means Untuk Mengetahui Kemampuan Pegawai Dibidang It Pada Cv . Roxed Ltd," *J. Pelita Inform.*, vol. 18, pp. 80–86, 2019.

[3] A. L. S. Saabith, E. Sundararajan, and A. A. Bakar, "Parallel implementation of Apriori algorithms on the Hadoop-MapReduce platform - An evaluation of literature," *J. Theor. Appl. Inf. Technol.*, vol. 85, no. 3, pp. 321–351, 2016.

[4] A. K. Dubey, U. Gupta, and S. Jain, "Comparative study of K-means and fuzzy C-means algorithms on the breast cancer data," *Int. J. Adv. Sci. Eng. Inf. Technol.*, vol. 8, no. 1, pp. 18–29, 2018, doi: 10.18517/ijaseit.8.1.3490.

[5] S. Panjaitan *et al.*, "Implementation of Apriori Algorithm for Analysis of Consumer Purchase Patterns," *J. Phys. Conf. Ser.*, vol. 1255, no. 1, 2019, doi: 10.1088/1742-6596/1255/1/012057.

[6] J. Han and M. Kamber, *DataMiningConceptsAndTechniques*. .

[7] X. Yuan, "An improved Apriori algorithm for mining association rules," *AIP Conf. Proc.*, vol. 1820, no. March 2017, 2017, doi: 10.1063/1.4977361.

[8] R. Sovia, E. P. W. Mandala, and S. Mardhiah, "Algoritma K-Means dalam Pemilihan Siswa Berprestasi dan Metode SAW untuk Prediksi Penerima Beasiswa Berprestasi," *J. Edukasi dan Penelit. Inform.*, vol. 6, no. 2, p. 181, 2020, doi: 10.26418/jp.v6i2.37759.

[9] F. A. Rahman, S. M. Shamsuddin, S. Hasan, and N. A. Haris, "A review of KDD-data mining framework and its application in logistics and transportation," *Int. J. Supply Chain Manag.*, vol. 5, no. 2, pp. 77–84, 2016.

[10] Aprilla Dennis, "Belajar Data Mining dengan RapidMiner," *Innov. Knowl. Manag. Bus. Glob. Theory Pract. Vols 1 2*, vol. 5, no. 4, pp. 1–5, 2013, doi: 10.1007/s13398-014-0173-7.2.

[11] A. Salam, J. Zeniarja, W. Wicaksono, and L. Kharisma, "Pencarian Pola Asosiasi Untuk Penataan Barang Dengan Menggunakan Perbandingan Algoritma Apriori Dan Fp-Growth (Study Kasus Distro Epo Store Pemalang)," *Dinamik*, vol. 23, no. 2, pp. 57–65, 2019, doi: 10.35315/dinamik.v23i2.7178.

[12] I. Vhallah, S. Sumijan, and J. Santony, "Pengelompokan Mahasiswa Potensial Drop Out Menggunakan Metode Clustering K-Means," *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 2, no. 2, pp. 572–577, 2018, doi: 10.29207/resti.v2i2.308.

[13] D. Sepri, M. Afdal, and S. Riau, "Analisa Dan Perbandingan Metode Algoritma Apriori Dan Fp-Growth Untuk Mencari Pola Daerah Strategis Pengenalan Kampus Studi Kasus Di Stkip Adzkia Padang," *J. Sist. Inf. Kaputama*, vol. 1, no. 1, 2017.

[14] R. M. Simanjorang, "Implementation of Apriori Algorithm in Determining the Level of Printing Needs," *Infokum*, vol. 8, no. 2, Juni, pp. 43–48, 2020.

[15] Y. Nur, A. Triayudi, and I. Diana, "Implementation of Data Mining to Predict Food Sales

Rate Method using Apriori," *Int. J. Comput. Appl.*, vol. 178, no. 35, pp. 22–28, 2019, doi: 10.5120/ijca2019919228.

[16] E. Elisa, "Market Basket Analysis Pada Mini Market Ayu Dengan Algoritma Apriori Jurnal Edik Informatika," *J. Edik Inform.*, vol. 4, pp. 29–38, 2017.

[17] P. Anitha and M. M. Patil, "RFM model for customer purchase behavior using K-Means algorithm," *J. King Saud Univ. - Comput. Inf. Sci.*, 2020, doi: 10.1016/j.jksuci.2019.12.011.

[18] C. Dengan and M. Algoritma, "Pengembangan Aplikasi Penentuan Tingkat Keuntungan Pada E-," vol. 2016, no. Sentika, pp. 18–19, 2016.

[19] R. K. Dinata, H. Novriando, N. Hasdyna, and S. Retno, "Reduksi Atribut Menggunakan Information Gain untuk Optimasi Cluster Algoritma K-Means," *J. Edukasi dan Penelit. Inform.*, vol. 6, no. 1, p. 48, 2020, doi: 10.26418/jp.v6i1.37606.

[20] M. Z. Hossain, M. N. Akhtar, R. B. Ahmad, and M. Rahman, "A dynamic K-means clustering for data mining," *Indones. J. Electr. Eng. Comput. Sci.*, vol. 13, no. 2, pp. 521–526, 2019, doi: 10.11591/ijeecs.v13.i2.pp521-526.

[21] T. H. Sardar and Z. Ansari, "An analysis of MapReduce efficiency in document clustering using parallel K-means algorithm," *Futur. Comput. Informatics J.*, vol. 3, no. 2, pp. 200–209, 2018, doi: 10.1016/j.fcij.2018.03.003.

[22] M. Mardalius, "Pemanfaatan Rapid Miner Studio 8.2 Untuk Pengelompokan Data Penjualan Aksesoris Menggunakan Algoritma K-Means," *Jurteksi*, vol. 4, no. 2, pp. 123–132, 2018, doi: 10.33330/jurteksi.v4i2.36.

[23] D. Q. Zeebaree, H. Haron, A. M. Abdulazeez, and S. R. M. Zeebaree, "Combination of k-means clustering with genetic algorithm: A review," *Int. J. Appl. Eng. Res.*, vol. 12, no. 24, pp. 14238–14245, 2017.

[24] H. Priyatman, F. Sajid, and D. Haldivany, "Klasterisasi Menggunakan Algoritma K-Means Clustering untuk Memprediksi Waktu Kelulusan Mahasiswa," *J. Edukasi dan Penelit. Inform.*, vol. 5, no. 1, p. 62, 2019, doi: 10.26418/jp.v5i1.29611.

[25] J. Jooa, S. Bangb, and G. Parka, "Implementation of a Recommendation System Using Association Rules and Collaborative Filtering," *Procedia Comput. Sci.*, vol. 91, no. Itqm 2016, pp. 944–952, 2016, doi: 10.1016/j.procs.2016.07.115.

[26] Y. Farhang, "Face Extraction from Image based on K-Means Clustering Algorithms," *Int. J. Adv. Comput. Sci. Appl.*, vol. 8, no. 9, pp. 96–107, 2017, doi: 10.14569/ijacsa.2017.080914.

[27] A. Z. Y. Ridho Ananda, "JURNAL RESTI Penentuan Centroid Awal K-means pada proses Clustering Data Evaluasi," *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 1, no. 10, pp. 544–550, 2021.

[28] M. Kaur and S. Kang, "Market Basket Analysis: Identify the Changing Trends of Market Data Using Association Rule Mining," *Procedia Comput. Sci.*, vol. 85, no. Cms, pp. 78–85, 2016, doi: 10.1016/j.procs.2016.05.180.

[29] L. Hakim and H. Seruni, "Indikasi Penyimpangan Laporan Keuangan Akademik Universitas XYZ Menggunakan Algoritma Greedy dan K-Means," *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 2, no. 1, pp. 301–306, 2018, doi: 10.29207/resti.v2i1.261.

[30] A. F. N. Alrammahi and K. B. S. Aljanabi, "A new approach for improving clustering algorithms performance," *Indones. J. Electr. Eng. Comput. Sci.*, vol. 20, no. 3, pp. 1569–1575, 2020, doi: 10.11591/ijeecs.v20.i3.pp1569-1575.