Comparison Between Davies-Bouldin Index and Silhouette Coefficient Evaluation Methods in Retail Store Sales Transaction Data Clusterization Using K-Medoids Algorithm

Kholiq Amrulloh
Informatics Department
Faculty of of Science and Informatics
Universitas Jenderal Achmad Yani
Cimahi, Indonesia
kholiqa18@if.unjani.ac.id

Tacbir Hendro Pudjiantoro
Informatics Department
Faculty of Science and Informatics
Universitas Jenderal Achmad Yani
Cimahi, Indonesia
thp@if.unjani.ac.id

Puspita Nurul Sabrina
Informatics Department
Faculty of Science and Informatics
Universitas Jenderal Achmad Yani
Cimahi, Indonesia
puspita.sabrina@lecture.unjani.ac.id

Asep Id Hadiana
Informatics Department
Faculty of Science and Informatics
Universitas Jenderal Achmad Yani
Cimahi, Indonesia
asep.hadiana@lecture.unjani.ac.id

Abstract

Retail business is the business of selling goods or services to consumers in units or retail. This retail business is part of the distribution channel that plays a vital role in a series of marketing activities as well as a liaison between the interests of producers and consumers. Based on sales transaction data in retail stores in 2020 obtained from www.kaggle.com, the inventory of goods is not proportional to the sales of goods. Excessive inventory and low sales levels resulted in goods accumulation in retail stores. When the sales cycle of goods is down, the stock must be prepared according to the level of sales. It takes a grouping of data to schedule an inventory of interests following the status of the purchase of goods. The data grouping used in this study uses the K-Medoids algorithm. K-Medoids is a method of partitioning clustering to group a set of (n) objects into several (k) clusters. Based on the elbow method, the optimal cluster number is 2 clusters. From the clustering process, the results obtained are cluster 1 has 320 data and cluster 2 has 765 data. The accuracy level of the cluster formed using the Davies-Bouldin Index method is 0.662748, and the Silhouette Coefficient is 0.276353.

Keywords
Retail Business, K-Medoids, Davies-Bouldin Index, Silhouette Coefficient.

1. Introduction
As a country that has the fourth density level in the world, Indonesia is an attractive country for retail business. Retail is a distribution channel that plays an important role in many market activities and acts as an intermediary between producers and consumers. Retail business is an activity related to the sale of goods and services carried out by end users for personal and non-commercial purposes (Susilawati and Nugraha 2016). Sales are activities carried out by trading companies to ensure their survival in order to make a profit (Aminah 2017).

The problems experienced include the inventory of goods which results in the accumulation of goods in the warehouse. The accumulation of goods occurs because consumer purchases are not proportional to the inventory of goods prepared by the store owner. In this case, expenses will increase to fund stockpiling of goods, maintenance, and the bad things that can occur due to a long shelf life. Regular warehouse management certainly guarantees the quality of its products. Because when the demand for a product is known, the company prepares the product as needed so that consumers are fully served and can continue to be committed to the product. Consumers can decide to switch if the retailer or distributor cannot provide the core service for the product they are looking for (Triatmojo 2013).

In a previous study entitled K-Medoids Algorithm for Product Marketing Strategy Determination. The main purpose of grouping sales data can determine the right marketing strategy. The data is sales data at a supermarket, which is as many as 1360 sales transaction data. This research has 5 clusters. From 909 transaction records, cluster 1 has characteristics of about 1-2 purchased goods and 301-302 items not purchased, cluster 2 has characteristics of 166 transactions. Records and clusters containing transactions with approximately 3-7 purchased and approximately 296-300 not purchased, cluster 3 has characteristics consisting of 66 transaction records, 8-14 transactions for goods purchased and 289 to 295 not purchased, and cluster 4 consists of of 132 transaction records, contains transactions with approximately 15-48 purchased, and approximately 255 to 288 not purchased. Cluster 5 is characterized by 87 transaction records and contains 49-303 bought quantities, and unbought quantities around 0-254 (Triyanto 2015).

The previous research was entitled Analysis of the K-Medoids Clustering Algorithm in Clustering the Covid-19 Outbreak in Indonesia. The main goal is to use K-Medoids technology to find the number of clusters of Covid-19 cases in Indonesia. The data includes variables obtained from the Ministry of Health of the Republic of Indonesia: cases of positive patients, cases of recovered patients, and patients who died. Based on this research, K-Medoids can group Covid-19 data in all infected areas for optimal grouping into 3 clusters. Of the 34 records received, one record is in the first cluster, two records are in the second cluster, and 31 records are in the third cluster. (Sindi et al. 2020).

Research entitled K-Medoids Algorithm for Grouping Villages with School Facilities in Indonesia. The aim is to propose that the government pay more attention to states with few school facilities. From the calculation of 34 states, there are 3 clusters, the highest cluster in 3 states, the middle cluster in 16 states, and the lowest cluster in 15 states (Damanik et al. 2019). In a study entitled Grouping Sales Transaction Data Using the K-Means Clustering Algorithm at A Nur's Store. The goal is to create a clustering system to group commodities using K-means at A Nur's store as a strategy for determining stock of goods. The results of this study are 1 product that is not selling well, 9 products that are selling well and 40 products that are very selling (Jeklin 2016).

From the several studies above related to the clustering method, based on the research entitled Comparison of K-Means and K-Medoids Algorithms for Grouping Loading and Unloading Transaction Data in Riau Province, the K-Medoids algorithm is a solution for K-Means which is sensitive to data that has outliers (Kamila, Khairunnisa, and Mustakim 2019), therefore the author will use the K-Medoids algorithm in this study. In previous studies, there has also been no research comparing evaluation methods from the results of data clustering so that this can be a gap for authors to conduct research. The evaluation methods to be compared are the Davies-Bouldin Index and the Silhouette Coefficient.

The author will process sales transaction data at retail stores for certain purposes, where the data will be processed to produce groups of goods that can be used as references in planning inventory. Transaction data grouping is done with K-Medoids. This method is a method used to get k-clusters among the data closest to the object in data grouping (Sindi

et al. 2020). The K-Medoids algorithm is an update of the K-Means method based on using a medoid instead of observing the average held by each cluster with the aim of reducing the sensitivity level of the resulting partition to the extremes contained in the data (Triyanto 2015). K-Medoids algorithm is a clustering technique to group a collection of (n) objects into several (k) clusters. K-Medoids uses objects in a set of objects as representatives of the cluster. The object selected for cluster representation is called the medoid (Nurlaela, Primajaya, and Padilah 2020).

1.1 Objectives

The purpose of this study is to compare the evaluation method of the Davies Bouldin Index and the Silhouette Coefficient on the results of grouping retail store sales transaction data using the K-Medoids algorithm.

2. Literature Review

Forest/land fires are one of the annual disasters in several countries in the world. This incident has attracted the attention of the government because it has caused a lot of losses in the economy, environment and society. Indonesia is a country with a high incidence of forest/land fires. Indonesia suffered losses of up to IDR 209 trillion in 2015. Early management is needed because of the losses. One of the steps you can take is to use hotspot data to classify areas of potential forest/land fires. Forest/land fires are indicated by hotspot detection satellites and displayed as hotspots. This study uses hotspot data which includes latitude, longitude, magnitude, frp (fire radiation power), and reliability parameters using the K-Medoids method. The K-Medoids method is a clustering method designed to divide the dataset into several groups. The advantages of this method can overcome the weakness of the K-Means method which is sensitive to outliers. The results of this study indicate that the K-Medoids method for the hotspot data clustering process produces the best silhouette coefficient of 0.56745 when using two clusters with an increase in the amount of data to 7352. The analysis of clustering results shows that the use of two clusters produces potential data clusters. That is, cluster 1 with an average brightness of 344,470 K includes high potential with an average reliability of 87.08%, and cluster 2 includes medium potential. The average brightness is 318,800K and the average reliability is 58.73%. (Pramesti et al. 2017).

A customer-centered marketing strategy plays an important role in managing good customer relationships. For your marketing strategy to work, you can use customer segmentation to group customers according to the same characteristics. When preparing a marketing strategy, IT can be used in areas of data processing, including data mining. The use of computer technology to process data that is less than optimal has led to the accumulation of information-poor data. We apply the clustering technique using the K-Medoids algorithm on the sales transaction dataset to determine customer segmentation. Your marketing strategy is based on the types and characteristics of your customers in each formed customer cluster or segment. Cluster validation testing was carried out using the Silhouette index and the Davies-Bouldin index to determine the optimal number of clusters. The results of this study indicate that the optimal number of clusters is 3 clusters, with a maximum silhouette index value of 0.375 and a minimum Davis-Dolbin index value of 1.030. The customer segments created by the survey are lost customers, loyal customers, and new customers (Sulistyawati and Sadikin 2021).

3. Methods

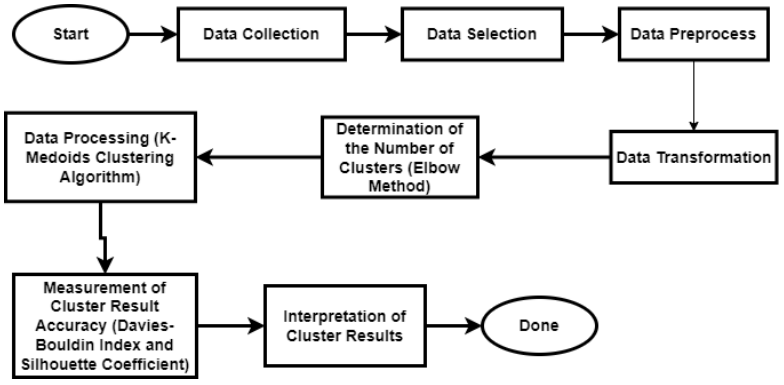

Figure 1 Research Method

3.1 Data Collection

The author conducted a data search to support the research process. Below is the data collection procedure:

a. Literature review
   This technique is done by examining reports related to the use of the K-Medoids Algorithm and the inventory of goods in retail stores.
b. Statistics
   In this statistical data look for public datasets regarding sales transaction data in retail stores. The dataset was obtained from the public dataset site, namely Kaggle.com. The data has eleven variables, namely goods_id, name_of_goods, total_stock_of_initial_goods, transaction, goods_sold, remainder_of_goods, purchase_price, selling_price, profit, average sales (one year) and year.

3.2 Data Selection

This selection is doing data selection. The selection of the data in question is the variable of use used in the study. The variables used are total_stock_of_initial_goods, transaction, goods_sold, remainder_of_goods, profit, and average sales in one year.

3.3 Data Preprocess

Perform data cleaning, with remov all duplicate data, checking data inconsistent, empty data, and correct data errors. Handling empty data will be done by deleting data objects.

3.4 Data Transformation

The transformation is carried out in order to equalize the data in the range 0 to 1 to avoid outliers by normalizing the data. Min-max normalization was chosen as the normalization method. The transformation is done to equalize the data in the range 0 to 1 to avoid outliers by normalizing the data (Goldman, Ian. and Pabari 2021). Min-max normalization was chosen as the normalization method. The data normalization stage is carried out using the formula in equation 1.

$$X_{new} = \frac{X - \min A}{\max A - \min A}(newMax - newMin) + newMin \quad (1)$$

Where:
Xnew : normalized value
X : the data will be normalized
minA : smallest value of data
minB : largest value of data
newMax : largest value desired
newMin : smallest value desired

3.5 Determination Number of Clusters

Determine the number of clusters based on the data you have using the Elbow Method. The resulting number of clusters can be used to initialize the value of k at the beginning of the K-Medoids process. The elbow method is used to determine the number of clusters based on the percentage comparison of the number of clusters that form an angle at a point. When the first cluster to the second cluster on the graph gives the angle or the resulting value decreases drastically, then it is the optimal number of clusters. Comparisons are obtained by calculating the number of Sum of Square Errors (SSE) from each cluster value. The larger the number of clusters K, the smaller the resulting SSE value will be (Dewi and Pramita 2019). Determination of the value of k using the elbow method with the SSE formula which can be seen in equation 2.

1. Initialize K value
2. Start
3. Increase in K value
4. Measuring SSE

$$SE = \sum_{K=1}^{K} \sum_{Xi=1} |Xi - Ck|^2 \quad (2)$$

Information:
K : c cluster

Xi : object data distance k-i
Ck : center of cluster i

5. If the SSE point drops drastically, it indicates the correct value of k
6. Done

3.6 Data Processing

Perform data processing with the K-Medoids algorithm to look for certain patterns so that the results of grouping the data obtained can be used as knowledge for reference in determining inventory. K-Medoids is a clustering technique that groups a collection of (n) objects into several (k) clusters. K-Medoids uses objects in a collection of objects as representatives of the cluster. The object selected to represent the cluster is called the medoid (Nurlaela, Primajaya, and Padilah 2020).

K-Medoids is almost the same as K-Means clustering. The difference between the two is that K-Medoids use the object for the medoid as the center of each cluster, while K-Means uses the mean for the center of the cluster. K-Medoids have the advantage of being immune to noise and outliers. Another advantage is that the results of clustering do not depend on the order in which the data is entered (Pramesti et al. 2017). The following is the flow of the K-Medoids algorithm:

1. Determine the value of k as the center point of the initial medoid
2. Calculate the distance of the data (object) to the medoid with the Euclidean Distance equation as in equation 3.

$$d_{\text{Eucledian}}(x, y) = \sqrt{\sum_{i=1}^{n}(xi - yi)^2} \quad (3)$$

Where:
d_{xy} : object distance between x to y
n : the amount of data
xi : variable data to i
yi : data variable medoid to i

3. Select a random (object) from the members of each cluster as a candidate for a new medoid.
4. Calculate the distance of each object in each cluster using the new candidate medoid.
5. Calculate the total deviation (S) using the sum of the new distances -the sum of the old distances. If the total deviation is less than 0, swap objects with clusters to form k new objects as medoids.
6. Repeat steps c to e until there is no change.

3.7 Measurement of Cluster Result Accuracy

3.7.1 Davies-Bouldin Index

Accuracy was measured using the Davies-Bouldin index and the silhouette factor. This is a good cluster sign, because the Davies Bouldin Index (DBI) value is close to 0 and is not negative. (Jumadi and USU 2018). The method introduced by David L. Davies and Donald W. Bouldin. This name uses the second name, the Davies-Bouldin Index (DBI) of the evaluation cluster. DBI has an internal cluster scoring scheme. Cluster scores are considered not good or good because of the number and proximity of cluster scores (Herviany et al. 2021).

a. Sum of square within cluster (SSW)

Used to find the cohesion value in the i-th cluster which can be seen in equation 4.

$$\text{SSW} = \frac{1}{m_i}\sum_{j=1}^{m_i} \text{d}(x_j, c_i) \quad (4)$$

Information:

m_i : the amount of data in the i-th cluster

x : data in cluster

x_j : data in the cluster

c_i : medoid cluster i

d (x_j, c_i) : Euclidean distance of each data to medoid

b. Sum of Square Between-cluster (SSB)

Used to find the separation between clusters which can be seen in equation 5.

$$SSB_{i,j} = d(C_i, C_j) \quad (5)$$

Information:

Ci = cluster One

Cj = another cluster

D (Ci, Cj) = distance from one centroid to another

c. Ratio

Used to get the comparison value between the i-th cluster and j-th cluster. The ratio calculation can be seen in equation 6.

$$R_{i,j} = \frac{SSW_i + SSW_j}{SSB_{i,j}} \quad (6)$$

Information:

Ri,j = ratio between clusters

SSWi = cluster 1

SSWj = cluster 2

SSBi,j = separation from clusters 1 and 2

d. Davies-Bouldin Index (DBI)

Used to get the Davies-Bouldin Index (DBI) value which can be seen in equation 7.

$$DBI = \frac{1}{k} \sum_{i=1}^{k} max_{i \neq j}(R_{i,j}) \quad (7)$$

Information:

k = existing cluster

Ri,j = Ratio between cluster i and j

Max = find the largest inter-cluster ratio

3.7.2 Silhouette Coefficient

Is a cluster assessment method to check the quality of objects in a cluster. Procedure for calculating silhouette coefficient (Nurlaela, Primajaya, and Padilah 2020), of them :

a. Calculate the average distance from an i-th object with each other object in one cluster which can be seen in equation 8.

$$\text{a(i)} = \frac{1}{[A]-1} \sum J \in_{A, j \neq 1} d(i, j) \quad (8)$$

b. Calculate the average distance from the i-th object with all objects in the other clusters, then take the smallest value that can be seen in equation 9.

$$\mathrm{b(i)} = \frac{1}{[A]} \sum J \in C \;\; d(i,j) \tag{9}$$

c. Silhouette Coefficient value is obtained by a(i) combined with b(i). So it can be formulated as in equation 10.

$$s(i) = \frac{b(i)-a(i)}{\max\,(a(i),b(i))} \tag{10}$$

Where:

s(i) : Silhouette Value
a(i) : Average distance between i and all objects in its cluster.
b(i) : Average distance between data i on all objects in other clusters.
d(i,j) : Distance between object i and j.

3.8 Interpretation of Cluster Result

Interpretation of cluster results is a stage that is carried out as knowledge of the results of data grouping.

4. Results and Discussion

4.1 Data Collection

The data obtained from Kaggle.com is data on sales transactions at retail stores in 2020. The data is 1085 data with 11 variables. Retail store sales transaction data can be seen in Table 1.

Table 1 Sales Transaction Data

goods_ id	name_of_ goods	total_ stock_of_ initial_ goods	transaction	goods _sold	remainder_ of_goods	purchase_ price	selling_ price	profit	average_ sales_ in_one_ year	years
228327	Glutinous Rice 2 kg	6859	26	6751	108	42000	44400	2400	260	2020
270422	Sugar ½ kg	6926	30	6358	568	6000	7400	300	2	2020
251405	Bodrex 1 sheet	6764	3619	6233	531	5000	5300	1400	212	2020
274214	Sugar 1 kg	6781	2280	5937	844	12000	13500	1500	3	2020
…	…	…	…	…	…	…	…	…	…	…
299637	White Coffe Mangoose	6921	1	1	6920			2000	1	2020

4.2 Data Selection

From the data that has been obtained, then the data selection for use in this study. Table 2 is a form of data that has been carried out in the selection stage.

Table 2 Selection Result Data

total_ stock_of_ initial_ goods	transaction	goods _sold	remainder_ of_goods	profit	average_ sales_ in_one_ year

6859	26	6751	108	2400	260
6926	30	6358	568	300	2
6764	3619	6233	531	1400	212
6781	2280	5937	844	1500	3
…	…	…	…	…	…
6921	1	1	6920	2000	1

4.3 Data Preprocess

The data preprocessing stage is carried out to prepare the data so that it can be processed by the system by overcoming incomplete data and converting the data into the same format.

4.4 Data Transformation

The form of the data is transformed into a format that can be processed. Because this study uses the K-Medoids method, in order to avoid outliers, data uniformity is needed, namely the data is changed using min-max normalization which will be in the form of a range of 0 to 1. The data before being transformed is as in Table 3 and after being transformed as in Table 4.

Table 3 Data Before Transformation

total_ stock_of_ initial_ goods	**transaction**	**goods _sold**	**remainder_ of_goods**	**profit**	**average_ sales_ in_one_ year**
6859	26	6751	108	2400	260
6926	30	6358	568	300	2
6764	3619	6233	531	1400	212
6781	2280	5937	844	1500	3
…	…	…	…	…	…
6921	1	1	6920	2000	1

Table 4 Data After Transformation

total_ stock_of_ initial_ goods	**transaction**	**goods _sold**	**remainder_ of_goods**	**profit**	**average_ sales_ in_one_ year**
0,4337	0,0069	1,0000	0,0000	0,7931	1,0000
0,7028	0,0080	0,9418	0,0668	0,0690	0,0039
0,0522	1,0000	0,9233	0,0614	0,4483	0,8147
0,1205	0,6299	0,8794	0,1068	0,4828	0,0077
…	…	…	…	…	…
0,6827	0,0000	0,0000	0,9887	0,6552	0,0000

4.5 Determination Number of Cluster

The number of clusters is determined using the elbow method. The number of clusters is used as the basis for the process of grouping k-medoids. In determining the number of clusters, it is best to use the R language and the Rstudio application. From the results obtained that the optimal value of k is as much as 2 as can be seen in Figure 2.

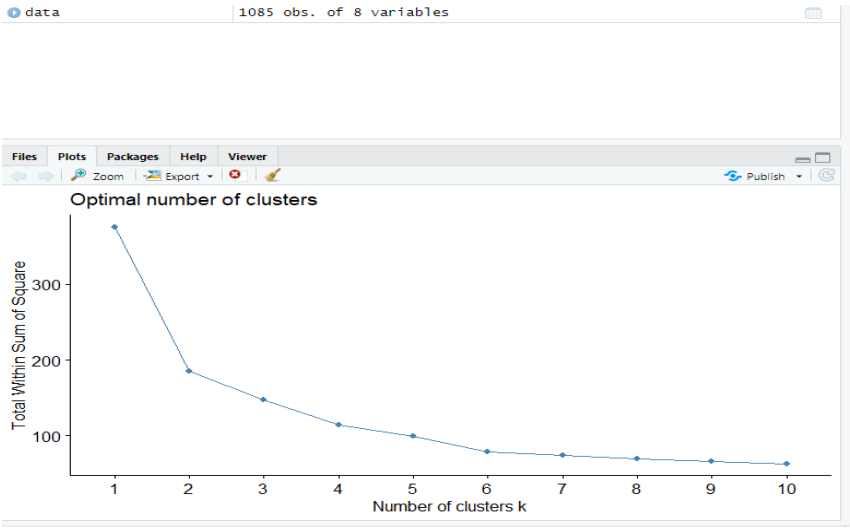

Figure 2 Elbow Method Cluster Value

4.6 Data Processing

The iteration results that occur when the data clustering process occurs 5 times and the results obtained can be seen in Table 5.

Table 5 K-Medoids Clustering Result

total_ stock_of_ initial_ goods	transaction	goods _sold	remainder_ of_goods	profit	average_ sales_ in_one_ year	C1	C2	Cluster
0,4337	0,0069	1,0000	0,0000	0,7931	1,0000	1,909345	1,779628	2
0,7028	0,0080	0,9418	0,0668	0,0690	0,0039	1,362063	1,56692	1
0,0522	1,0000	0,9233	0,0614	0,4483	0,8147	2,079129	2,054461	2
0,1205	0,6299	0,8794	0,1068	0,4828	0,0077	1,669985	1,632214	2
…	…	…	…	…	…	…	…	…
0,6827	0,0000	0,0000	0,9887	0,6552	0,0000	0,540197	0,294595	2

The total distance in this clustering result is 493,99747.

4.7 Measurement of Cluster Result Accuracy

In the evaluation of the Davies-Bouldin index, each calculation has results, namely SSW 1 0.490288, SSW 2 0.44066, ratio 1.325497, with the DBI value obtained is 0.662748.
In the evaluation of the silhouette coefficient, the value of s(i) is 0.27635.

5. Conclusion

From this research, the k-medoids algorithm can group the sales transaction data obtained. From the elbow method, determining the number of clusters as many as 2 clusters, then the results obtained are clusters of 320 data in cluster 1 and 765 data in cluster 2.

Based on the results of the evaluation of cluster results using the Davies-Bouldin index method, the result is 0.662748 which indicates that the results of the cluster formed are quite good, while the results of the cluster evaluation using the Silhouette Coefficient method are 0.276353 which are included in the cluster category with low structure.

References

Aminah, Surti. "Dagangan Dengan Menggunakan Metode Masuk Pertama Keluar Pertama (Mpkp) Perpetual Pada UD Aneka Diesel Banjarmasin." 2017.

Damanik, Ivana Indrini Putri et al. "Algoritma K-Medoids Untuk Mengelompokkan Desa Yang Memiliki Fasilitas Sekolah Di Indonesia." *Prosiding Seminar Nasional Riset Information Science (SENARIS)* 1(September): 520. 2019.

Dewi, Dewa Ayu Indah Cahya, and Dewa Ayu Kadek Pramita. "Analisis Perbandingan Metode Elbow Dan Silhouette Pada Algoritma Clustering K-Medoids Dalam Pengelompokan Produksi Kerajinan Bali." *Matrix : Jurnal Manajemen Teknologi dan Informatika* 9(3): 102–9. 2019.

Goldman, Ian. and Pabari, Mine. "Implementasi Algoritma K-Medoids Dalam Menentukan Tingkat Penyebaran Pandemi Covid-19 Di Indonesia." 2021.

Herviany, Mufidah, Saleha Putri Delima, Triyana Nurhidayah, and Kasini. "Perbandingan Algoritma K-Means Dan K-Medoids Untuk Pengelompokkan Daerah Rawan Tanah Longsor Di Provinsi Jawa Barat." *MALCOM: Indonesian Journal of Machine Learning and Computer Science* 1(1): 34–40. 2021.

Jeklin, Andrew. "Pengelompokan Data Transaksi Penjualan Menggunakan Algoritma Clustering K-Means Pada Toko An Nur." (July): 1–23. 2016.

Jumadi, Bernard, and USU. "Peningkatan Hasil Evaluasi Clustering Davies Bouldin Dengan Penentuan Titik Pusat Cluster Awal K Means." 2018.

Kamila, Insanul, Ulya Khairunnisa, and Mustakim Mustakim. "Perbandingan Algoritma K-Means Dan K-Medoids Untuk Pengelompokan Data Transaksi Bongkar Muat Di Provinsi Riau." *Jurnal Ilmiah Rekayasa dan Manajemen Sistem Informasi* 5(1): 119. 2019.

Nurlaela, Siti, Aji Primajaya, and Tesa Nur Padilah. "Algoritma K-Medoids Untuk Clustering Penyakit Maag Di Kabupaten Karawang." *I N F O R M a T I K a* 12(2): 56. 2020.

Pramesti, Dyang Falila, Lahan, M Tanzil Furqon, and Candra Dewi. "Implementasi Metode K-Medoids Clustering Untuk Pengelompokan Data Potensi Kebakaran Hutan/Lahan Berdasarkan Persebaran Titik Panas (Hotspot)." *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer* 1(9): 723–32. 2017.http://j-ptiik.ub.ac.id.

Sindi, Sukma et al. "Analisis Algoritma K-Medoids Clustering Dalam Pengelompokan Penyebaran Covid-19 Di Indonesia." *Jti (Jurnal Teknologi Informasi)* 4(1): 166–73. 2020.

Sulistyawati, Anggi Ayu Dwi, and Mujiono Sadikin. "Penerapan Algoritma K-Medoids Untuk Menentukan Segmentasi Pelanggan." *Sistemasi* 10(3): 516. 2021.

Susilawati, Wati, and Didit Aditia Nugraha. "Pengaruh Kualitas Pelayanan Ritel Dan Atmosfir Gerai Terhadap Loyalitas Pelanggan Pada Minimarket Yomart Wanaraja 2 Garut." *Jurnal Wacana Ekonomi* 15(02): 009–019. 2016.www.jurnal.uniga.ac.id.

Triatmojo, Marif Panji. "Analisis Penerapan Manajemen Persediaan Di Toko Retail Mikro Di Kabupaten Sleman Yogyakarta." : 1–5. 2013.

Triyanto, Wiwit Agus. "Algoritma K-Medoids Untuk Penentuan Strategi Pemasaran Produk." *Simetris : Jurnal Teknik Mesin, Elektro dan Ilmu Komputer* 6(1): 183. 2015.

Biographies

Kholiq Amrulloh is a final year undergraduate student in the department of informatics, University Jenderal Achmad Yani, Cimahi, Indonesia. Him primary interests are Data Mining, Systems Analysis, and software engineering.

Tacbir Hendro Pudjiantoro is an Associate Professor. Doctoral Candidate from the Indonesian University of Education. Researcher in the field of Knowledge Management and handles several Information Systems projects.

Puspita Nurul Sabrina is an Associate Lecturer. Received her master's degree in Informatics from Insitute Techonology of Bandung. Amongst her Researcher interest are Data mining, Software Quality, Business Intelegent.

Asep Id Hadiana received his master's degree in Enterprise Information System from Indonesian Computer Univerity and a Doctor of Philosophy from Universiti Teknikal Malaysia Melaka (UTeM). He is a lecturer in the Informatics Department, Faculty of Science and Informatics, Universitas Jenderal Achmad Yani. Amongst his research interest are Cyber Security, Data Mining, Spatial Analysis, Location Based Services and Geographic Information Systems.