ISSN 2085-4579

Use of the K-Medoids Algorithm for Food Clustering Using Nutritional Value and Evaluation of the Elbow Method and the Davies Bouldin Index Method

Wildani Eko Nugroho[1], Safar Dwi Kurniawan[1], Yerry Febrian Sabanise[1], Prayoga Alga Vredison[2]
[1] D3 Computer Enginering Program Study, Politeknik Harapan Bersama, Tegal, Indonesia
wild4n1@gmail.com, safar.kurniawan45@gmail.com, yerryfebrian@gmail.com
[2]Informatika, Universitas Pancasakti, Tegal, Indonesia
yogavredizon23@gmail.com

Accepted on June 07, 2025
Approved on June 23, 2025

Abstract—**The four categories of vital nutrients—calories, carbs, protein, and fat—must be present in the food that people eat on a daily basis. Humans require nutrition since it will enable them to do everyday duties and maintain their health. This study applies K-Medoids optimization to the clustering approach. The purpose of this study is to classify foods with comparable nutritional values. Foods with high, medium, and low nutritional levels are grouped into three clusters. Finding the ideal number of clusters is one of the difficulties in data clustering. For this, the Davies-Bouldin Index (DBI) and the elbow technique are frequently employed. The DBI evaluates the quality of clusters by taking into account the distance between clusters and the proximity between points inside the cluster, whereas the Elbow technique plots the intra-cluster variance against the number of clusters to assist in determining the number of clusters. Better clusters are indicated by lower DBI values. The outcomes of the K-Medoids algorithm using the Elbow approach showed a DBI of 0.631 and a cluster value (k) = 3 of 0.046. The better of these two approaches will be used for clustering. According to the comparative results of the two algorithms, the DBI technique yielded a value of 0.631 for the best cluster, while the Elbow approach yielded a number of 0.046 for the best number of clusters. The study's clustering results can be utilized to choose and consume foods that will meet nutritional needs and help delay the onset of food related disorders. For instance, if you wish to gain weight, you can choose foods in cluster 0. Cluster 2 foods can be picked if you wish to diet or lose weight, while Cluster 1 meals can serve as a benchmark if taken in excess, as this can lead to obesity.**

Index Terms*—*K-Means, K-Medoids, Nutrition, and Clustering

I. INTRODUCTION

One of the primary resources for human survival is food. There are many different types of food from different parts of Indonesia. For instance, the Moluccas have bagea, Makassar has palubasa, Semarang has spring rolls, Medan has bacang, and there are many more varieties. The flavors, textures, and attributes of these dishes vary [1].

A variety of vegetable and animal ingredients make up the actual food. Consequently, food is a rich source of calories, protein, fat, and carbohydrates, among other nutrients. Humans require nutrition in order to perform daily tasks and lead healthy lives [2].

A pertinent strategy in the fields of nutrition research and food technology is the use of the K-Medoids algorithm for food grouping according to nutritional value. Food clustering based on nutritional content might help people choose foods that meet their nutritional demands as the necessity of a balanced diet becomes more widely recognized [3]. This approach not only aids individuals in making informed food choices but also supports researchers and dietitians in developing tailored meal plans that cater to specific dietary needs. By leveraging the K-Medoids algorithm, we can create comprehensive food databases that enhance our understanding of nutritional relationships and promote healthier eating habits [4].

The K-Medoids algorithm has been used in a number of research studies to cluster foods according to their nutritional value. For instance, Prayoga Alga Vredizon et al.'s study divided foods into three clusters high, medium, and low based on their nutritional worth using K-Means [1]. Nevertheless, K-Means and K-Medoids were not specifically compared in this study [5].

Furthermore, a study by Tias Rahmawati et al. used K-Medoids to compare the Silhouette Coefficient and Elbow techniques in clustering Indonesian provinces according to the Human Development Index (HDI) [5]. This study demonstrates the significance of selecting the appropriate evaluation method in figuring out the ideal number of clusters, despite the fact that food is not the main focus.

Thus, the purpose of this work is to use the K-Medoids algorithm to cluster foods according to their nutritional content and assess the clustering outcomes using the Elbow and Davies-Bouldin Index methodologies. It is intended that the findings of this study will serve as a guide for future research of a similar nature and aid in the selection of foods that meet the nutritional requirements of each individual.

II. METHODS

A. *Clustering*

Clustering is a technique used to separate a set of data into groups based on a desired match. Clustering is the technique of organizing objects or data into clusters in data mining so that the resulting data is nearly equal to the original and can be identified from items in other clusters [6].

Clustering is used to separate data into regions with similar features such that each group has unique attributes. Without the use of preset labels or rules, clustering is a directed, or "unsupervised," data mining technique that arranges and searches data according to similarities between the data [7].

.

B. *K-Medoids Algorithm*

A partition clustering technique for forming clusters from a set of n objects is the K-Medoids algorithm. With this method, a specific item from the data set serves as a cluster center or representative. This algorithm's primary idea is to reduce the disparity between the objects and the selected cluster centers and make sure that the reference points in each cluster match the data that is provided [8].

The steps in the K-Medoids algorithm method are as follows [3]:

1. Calculate how many clusters (k) need to form.
2. Choose at random the initial cluster centers (medoids) for every cluster.
3. Utilizing the Euclidean distance range calculation in Eq., assign each observation to the nearest cluster.

$$d(x,y)\sqrt{\sum_{i-1}^{\pi}(x_i - y_i)^2; i = 1,2,3,\ldots n} \quad (1).$$

4. Choose items from each cluster at random to serve as new medoids.
5. Determine the separation between the new medoids and every object in each cluster.
6. Determine the distance value between the new and old total distances to determine the total deviation (S). To create a new set of k objects as medoids, swap the cluster data objects if S is less than 0.
7. Continue Steps 4 through 6 until the medoids remain unchanged, producing a cluster and each member of the cluster.

C. *Elbow Methods*

One popular technique for figuring out how many clusters should be formed during the clustering process is the elbow approach. To maximize or ascertain the ideal number of clusters for the clustering process to be executed, the elbow approach is utilized. Cluster cohesiveness and separation are calculated using the elbow approach [9]. Cluster separation gauges how distinct or disjointed one cluster is from another, and cluster cohesiveness gauges how tightly related the data in the cluster is. The sum of squares error (SSE) can be used to quantify cluster cohesiveness and cluster separation [10].

The following formula can be used to calculate SSE :

$$SSE = \sum_{k=1}^{k} \sum_{xi} |x_i - C_k|^2 \quad (2)$$

Description :
K = Cluster in C
X_i = Distance of the data on the i-th object
Ck = The i-th cluster's center

D. *Davies Bouldin Index (DBI) Methods*

One technique to guarantee the optimal number of clusters following the clustering process is Davies Bouldin Index (DBI) [11]. The DBI technique aims to minimize the distance between objects within a cluster and maximize the difference across clusters. Additionally, the DBI measure is used to evaluate the effectiveness of various clustering techniques; values near zero imply better clusters. The DBI algorithm divides the distance between the cluster centers by the average of the compactness values for each place [12]. An outline of the procedures to determine the DBI value is provided below [13]:

1. Finding the Sum of Squares The attachment between members of a cluster, or how similar they are to one another, is measured by the SSW; the smaller the cluster, the more similar the members are to one another. To find the matrix, cohesion, and homogeneity, SSW is computed. According to the equation formula, cohesion is the relationship between group members in a single cluster.

   $$\text{SSW} = \frac{1}{m_i}\sum_{j=1}^{m_i} d(x_j \,.\, c_i) \quad (3)$$

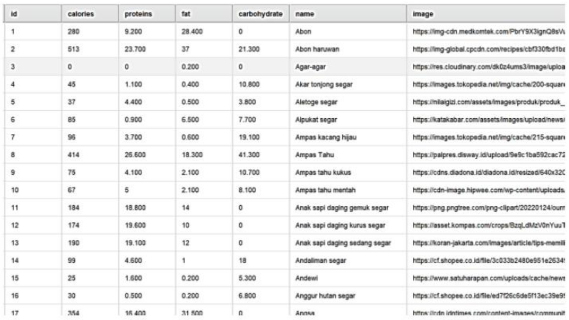
   Description :
   m_i = quantity of information in clusterke-i
   x = information within the cluster
   d(x,c) = data's distance from the centroid
   x_i = information within the cluster
   ci = center of cluster i centroid
2. To ascertain separation or heterogeneity, the Sum of Square Between Clusters (SSB), which is a sufficiently high distance between groups to keep them apart from one another, is calculated. The distinction between two groups is separation.
3. The purpose of calculating the ratio is to determine how well one cluster compares to

other clusters. There must be more separations than cohesiveness.

4. DBI calculation

III. RESULT AND DISCUSSION

Pre-processing, also known as preparation, is the first step in the research process. This involves a number of steps, such as data preprocessing, which includes attribute selection, removing duplicate data, normalizing data, and lowering the dimensions of the dataset used. After undergoing data preprocessing, the final dataset consists of food nutritional values. The K-Medoids technique is used on the final dataset in order to find the optimal clusterization performance value [14]. The K-Medoids Algorithm is used to conduct the evaluation procedure during the testing phase. A screenshot of the food nutrition dataset from the Indonesian Food and Drink Nutrition Dataset, https://www.kaggle.com/datasets/anasfikrihanif/indonesian-food-and-drink-nutrition-dataset, On November 1, 2023, Anas Fikri Hanis submitted 1346 nutrition lists for 100 grams of Indonesian food website is shown in Figure 1.



id	calories	proteins	fat	carbohydrate	name	image
1	280	9.200	28.400	0	Abon	https://img-cdn.medkomtek.com/PbrY9X3ignQ8sV
2	513	23.700	37	21.300	Abon haruwan	https://img-global.cpcdn.com/recipes/cbf330fbd1b
3	0	0	0.200	0	Agar-agar	https://res.cloudinary.com/dk0z4ums3/image/uploa
4	45	1.100	0.400	10.800	Akar tonjong segar	https://images.tokopedia.net/img/cache/200-square
5	37	4.400	0.500	3.800	Aletoge segar	https://nilaigizi.com/assets/images/produk/produk_
6	85	0.900	6.500	7.700	Alpukat segar	https://katakabar.com/assets/images/upload/news/
7	96	3.700	0.600	19.100	Ampas kacang hijau	https://images.tokopedia.net/img/cache/215-square
8	414	26.600	18.300	41.300	Ampas Tahu	https://palpres.disway.id/upload/9e9c1ba592cac72
9	75	4.100	2.100	10.700	Ampas tahu kukus	https://cdns.diadona.id/diadona.id/resized/640x32
10	67	5	2.100	8.100	Ampas tahu mentah	https://cdn-image.hipwee.com/wp-content/uploads
11	184	18.800	14	0	Anak sapi daging gemuk segar	https://png.pngtree.com/png-clipart/20220124/our
12	174	19.600	10	0	Anak sapi daging kurus segar	https://asset.kompas.com/crops/BzqLdMzV0nYuuT
13	190	19.100	12	0	Anak sapi daging sedang segar	https://koran-jakarta.com/images/article/tips-memili
14	99	4.600	1	18	Andaliman segar	https://cf.shopee.co.id/file/3c033b2480e951e2634
15	25	1.600	0.200	5.300	Andewi	https://www.satuharapan.com/uploads/cache/news
16	30	0.500	0.200	6.800	Anggur hutan segar	https://cf.shopee.co.id/file/ed7f26c6de5f13ec39e9
17	354	16.400	31.500	0	Angsa	https://cdn.idntimes.com/content-images/communit

Fig 1. Food Nutrition Dataset on the Kaggle Site

Several preparation stages must be completed before the K-Medoids algorithm is implemented. This is done in order to maximize the outcomes of the K-Medoids algorithm's implementation. Among the preprocessing procedures used in this study are the following ones:

a. Atribut selection

The purpose of attribute selection is to eliminate unnecessary qualities and choose those that will be utilized in the K-Medoids algorithm. Since the K-Medoids method can only accept characteristics with numeric data types, unneeded attributes are removed. Therefore, non-numerical attributes must be chosen first. This is the outcome of choosing the traits that will be utilized later in table 1.

Table 1. Result Of Attribute Selection

Name of the Attribute	Description
Calories	Calories 100 grams/food (cal)
Fat	100 grams of fat every meal (grams)
Protein	100 grams of protein each meal (grams)
Carbohidrat	100 grams of carbohydrates per meal (grams)

An overview of datasets that might be able to give an overview to assess the amount of nutrients that enter the body when ingesting typical Indonesian foods and drinks can be found in the attribute selection table above.

b. Remove Duplicate

By comparing each piece of data one at a time, the Remove Duplicates function eliminates duplicates. One data set from the comparison results will remain after all data with the same value are eliminated one at a time. Since the nutritional value of each product is typically the same, this research removes redundant data to increase the efficiency of data processing.

c. Outlier Detection

Finding data that behaves differently from other typical data requires outlier detection, which will have an impact on the K-Medoids algorithm's data processing procedure. Distance-based outlier identification is the outlier identification method employed in this study. Two points are determined by selecting a distance-based outlier, which is subsequently verified. It will be regarded as an outlier if the neighbor points are near together and then separated. Once this study has been tested, the parameters' findings are derived utilizing the distances of ten neighbors, nine outliers, and the Euclidean distance.

d. Normalization of Data

One of the data normalizing techniques employed in this study is min-max normalization. The original data values are transformed into various values between 0 and 1 using this min-max normalization. Forming data into numbers with the same range is the goal of the normalization process, which speeds up data processing.

e. The reduction of dimensions

Principal Component Analysis (PCA) is a technique used to minimize the dimensions of data containing numerous variables. PCA shrinks the data's dimensionality without appreciably lowering its information content. Based on the data, PCA creates a new set of dimensions that are rated. The principal components of PCA are derived from the breakdown of the covariance matrix's eigenvalues and eigenvectors. Two components and a fixed number of dimensions are the parameters utilized for PCA implementation. Two traits remained after the previous four were eliminate

1. K-Medoids Algorithm

A number of tasks or procedures need to be completed in experiments utilizing the K-Medoids algorithm, including dimension reduction, outlier detection, attribute selection, data standardization, and the elimination of duplicate data. The Elbow method and the Davies-Bouldin Index method are the two techniques used to test the K-Medoids algorithm. A perspective of the K-Medoids experiment is shown in Figure 2.

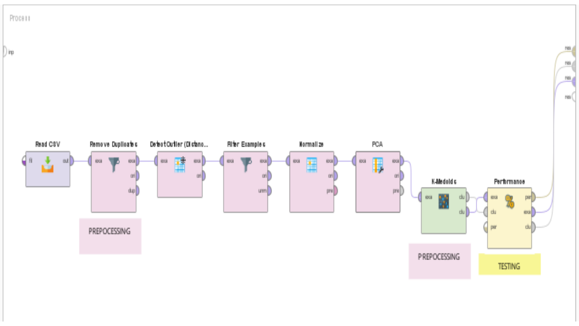


Fig 2. Algorithm Testing Scenario for K-Medoids

The K-Medoids method is tested in the image above. Both the DBI approach and the elbow method are employed in some of these exams. The K-Medoids algorithm is tested to determine its accuracy and success rate. The DBI approach and the elbow method are two methods that can be used to determine the number of clusters.

A. Elbow Method

In order to determine the number of clusters in the K-Medoids algorithm, the foundation of this phase is to establish an ideal K value. Each value in the Cluster Sum of Squares (WCSS) will be computed using the Elbow technique, which runs cluster values from 2 to n. The WCSS represents the sum of squares between each cluster's centroid value and average value [15]. The average distance of the sum of squares between the cluster center values is the sort of average calculation that is computed. Next, using a chartline, choose the spot that creates the most elbow. The outcomes of applying the Elbow approach to the K-Medoids algorithm, as displayed in Table 2, are as follow. Within-cluster sum of squares, K value, and average Figure 3. Graph Calculating How Many Clusters the K-Medoids Algorithm Needs

Table 2. Avg. Within-Cluster Sum Of Square K-Medoids

K	WCSS
2	0,091
3	0,046
4	0,036
5	0,027
6	0,023
7	0,019

The results of the Avg. Within-cluster Sum of Square K-Medoids, which represents the sum of these squared distances for every data point within each cluster, are shown in the table above. In K-Medoids, the ideal number of clusters is found using WCSS. Data within the cluster is more homogeneous, and the cluster is more compact when the WCSS is lower. Three to seven clusters are tested in the above table to determine the ideal number of clusters. Figure 3 shows the Elbow Method cluster experiments in Figures 3-7.

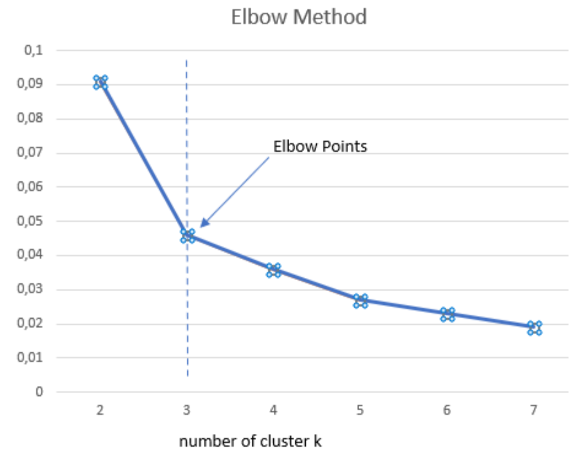


Fig 3. Calculating the K-Medoids Algorithm's Cluster Count

Figure 2 shows the outcome of an experiment to find the ideal number of clusters based on the K-Medoids algorithm's SSE (Sum of Square Error) value using the Elbow approach, which has significantly dropped in this instance. The cluster's quality decreases with increasing SSE value and vice versa. The cluster quality improves with a reduced SSE value. As can be seen in the above image, the SSE value is at its peak when there are two clusters (k = 2), and it dramatically decreases but forms an elbow when there are three clusters (k = 3). The SSE value dropped once more when there were four clusters (k = 4), and so on until there were seven clusters (k = 7). The elbow is found at the number of clusters k = 4 by calculating the average value. This is because the graph shows that the number of clusters that form an elbow is clearly visible when the number of clusters k = 3, and the number of clusters k = 4 to k = 7 is seen to start decreasing as well. sum of squares within the cluster of 0.046.

B. Davies Bouldin Index (DBI) Method

A cluster validation that incorporates separation and cohesion data is the Davies-Bouldin Index (DBI). Cohesion is the degree of similarity between the data and the cluster centroids, whereas separation is the distance between the cluster centroids. The quantity and proximity of the clustered data determine the quality of the clustering results. The cluster outcomes are better when the DBI value is smaller (non-negative => 0) [16]. The outcomes of applying the Davies-Bouldin Index approach to the K-Medoids algorithm, as displayed in table 3. are as follows.

Table 3. Value Of The Davies Bouldin Index K-Medoids

K	DBI
2	0,810
3	0,631
4	0,996
5	1,009
6	0,996
7	1,144

According to the above table 3., cluster k = 3 has the minimum DBI value, which is 0.631. This demonstrates that the outcomes match expectations.

The optimal number of clusters is three, according to the Elbow and Davies-Bouldin Index techniques that have been used. Table 4 below displays these findings:

Table 4. Value Of Wcss K-Medoids

Cluster	Avg. Within Centroid Distance
0	0,091
1	0,046
2	0,036

As can be seen from the above table, the cluster's data are reasonably close to one another, as indicated by the Average Within Centroid Distance value being near zero. The scatter/bubble in the image can be used to help visualize the image.

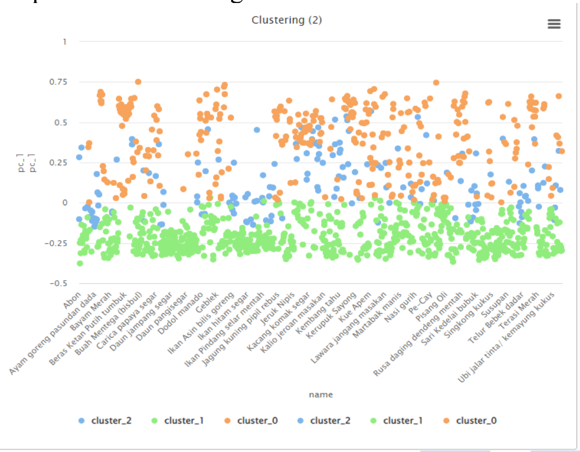


Fig 4. K-Medoids Algorithm Scatter/Bubble Outcome with Three Clusters

Three clusters were created by testing the K-medoids technique, as seen in the image above. Cluster 0 is represented by the color orange, Cluster 1 by the color green, and Cluster 2 by the color blue. Table 6 below displays the detailed clusterization results. However, Table 5 displays the quantity of data for every cluster.

Table 5. The Quantity Of Information For Every K-Medoids Cluster

Cluster	Jumlah Data
0	347
1	744
2	175

The best or most optimal number of clusters is at number three, according to the K-Medoids algorithm's implementation with multiple preprocessing steps, including data selection, duplicate removal, data normalization, and dimensional reduction, as well as cluster optimization using the Elbow Method and Davies-Bouldin Index. Clusters 0, 1, and 2 data are acquired by evaluating K-Medoids using these two techniques. There are 347 food data points in cluster 0, 744 data points in cluster 1, and 175 food data points in cluster 2. Table 5 illustrates this, and Table 6 provides the specific findings for the list of items that fall into each cluster.

Table 6. Some Food Information From The Clustering Findings Of K-Medoids

K	List Of Food
0	Asam payak segar, Bagea kelapa asin, Bagea kelapa manis, Bagea kenari asin, Bagea kenari manis, Bagea kw 1, Bagea kw 2, Bakpia kue, Bakwan, Bantal, Batatas tali ubi rebus, Beef burger, Belitung talas kukus, Beras ganyong, Beras Giling, Beras Giling masak (nasi), Beras ladang mentah, Beras Menir, Beras menir mentah, Beras Merah tumbuk, Bihun, Bihun goreng instan, Bihun Jagung mentah, Bihun mentah, Biji nangka/Biji salak, Bika ambon, Bika Ambon, Biskuit, Bonggol Pisang kering, Brem, Brondong, Buah kom segar, Bubur sagu, Bulung Jajan, Bulung Sangu, Bungkil Biji Karet, Bungkil Kelapa, Burung sarang segar, Cabai merah kering, Cake tape, Cantel mentah, Cengkeh kering, Centel, Ceriping getuk singkong, Coklat bubuk, Coklat Manis batang, Coklat Susu batang, Combro, Daun ndusuk segar, Daun salam bubuk, Deblo, Djibokum masakan, Dodol, Dodol bali, Dodol BanjarmasinEmping komak, Enting-enting gepuk hello kity, Es Krim (Coconut milk), Es Mambo, Gadeng/Gadung kering, Gaplek, Gatep segar, Gatot, Geblek, Gelatine, Gemblong, Gula putih, Gurandil, Havermout, Intip goreng, Jagung grontol, Jagung Kuning giling, Jagung Kuning pipil baru, Jagung uning pipil kering mentah, Jagung pipil var. harapan kering, Jagung pipil var. metro kering, Jagung Rebus, Jagung titi, Jali, Jam selai, Jampang huma mentah, Jamur encik, Jamur kuping kering, Japilus, Jawawut, Jenang, Jengkol segar, Kabuto, Kacang Andong, Kacang Arab, Kacang babi kering, Kacang Bogor goreng, Kacang tunis kering, Kacang uci kering, Kacang urei kering, Kambose, Katul Beras, Katul Jagung, Kecimpring singkong goreng, Kelepon, Kentang Hitam, Keremes, Keripik gadung, Kerupuk Aci, Kerupuk cumi goreng, Kerupuk Ikan berpati, Kerupuk kemplang (ikan) mentah, Kerupuk kemplang goreng, Ketumbar, Ketupat ketan, Kopi bagian yang larut, Kopi bubuk instant, Koro Benguk biji, Koro Roway biji, Koro Wedus biji, Koya, Koya mirasa, Kranji segar, Kue ali, Kue Apem, Kue bangket, Lopis, Lupis ketan, Madu, Maizena tepung, Makaroni, Markisa squash, Martabak manis, Martabak mesir, Martabak Telur, Masekat, Melase, Merica, Mie Bendo, Mie Goreng, Mie kering, Mie Sagu, Misoa, Misro, Nangka biji, Nasi, Nasi beras merah, Nasi Goreng, Nasi jagung, Nasi minyak, Nopia special, Oncom ampas kacang hijau, Onde-onde, Oyek, Oyek (dari singkong), Pastel, Pati Singkong (tapioka), Pempek belida, Pempek kulit, Pempek tenggiri, Pepaya lodeh, Permen, Petis Udang, Petis udang kering, Petis udang pasta, Pisang ketip segar, Pisang Mas, Pisang Oli, Pulut, Putu Mayang, Rarawuan, Rasbi (Beras Ubi), Rasi (Beras Singkong), Rebung laut mentah, Rempeyek kacang uci, Renggi goreng, Risoles, Roti boong, Roti Gambang, Roti Putih, Roti warna sawo matang, Sagu Ambon, Sale pisang siam. Santan (kelapa saja), Sapi abon, Sarimuka, Semur Jengkol, Sente talas segar, Serbuk Coklat, Serimping talas kebumen, Setrup sirup, Singkong Goreng, Singkong kukus, Singkong tape, Srikaya, Susu asam untuk bayi bubuk, Susu Kental Manis, Susu skim bubuk, Suwir-suwir, Talas pontianak segar, Tapai singkong, Tauco cap DAS cake, Tauji cap singa, Teh, Teh hijau daun kering, Teh melati daun kering, Tekokak kering, Tempe Sayur, Wajit camilan, Widaran, Wingko babat, Yangko
1	Bekasam, Bekasang, Belibis daging segar, Belimbing, Belut segar, Belutlaut segar, Bengkuang, Bentul (Komba) talas segar, Bentul talas kukus, Betok wadi masakan, Bihun Goreng, Bir (4% alkohol), Bit, Biwah segar, Buah ruruhi segar, Buah tuppa segar, Bubur tinotuan (Manado), Bulgogi masakan, Buncis, Buncis asam, Buncis rebus, Bunga pepaya segar, Bunga turi segar, Buntil, Buntil daun talas, Buras, Cabai gembor merah segar, Cakalang asar (asap papua), Cammetutu,

K	List Of Food
	Cap cai sayur, Carica papaya segar, Cempedak, Dideh darah ayam, Dideh darah sapi, Dodonkol, Domba daging kurus segar, Domba ginjal segar, Duku, Durian, Duwet segar, Eceng Gondok, Embacang, Encung asam segar, Erbis, Es krim, Es Sirup, Fillet o-fish, Gado-gado, Gadung kukus, Gadung mentah, Gadung ubi kukus, Gadung ubi segar, Gambas (Oyong), Gambas lodeh, Gandaria, Ganyong kukus, Ganyong mentah, Ganyong segar, Gembili, Gembili ubi segar, Genjer segar, Gete kuah asam masakan, Ginjal Babi, Ginjal Domba, Ginjal Sapi, Gudeg, Gudeg sayur, Gulai asam keueung masakan, Hangop, Hati Babi, Hati Sapi, Hofa/Ubi hutan segar, Ikan Patin segar, Kentang, Keong, Kepala Susu (Krim), Kepiting, Kerang, Kerbau daging segar, Keribang ubi segar, Kerokot segar, Kesemek, Ketimun, Ketimun madura segar, Ketoprak, Ketupat Tahu, Kokosan, Komak polong segar, Koro wedus polong, Kotiu hinela tawang nggole, Krokot, Kucai, Kucai Muda (Lokio), Kuda daging segar, Kue Koya, Kue sus, Kulit melinjo, Kundur segar, Kunyit, Kura-kura, Langsat, Lantar segar, Lawar babi masakan, Lawar penyu masakan, Lawara jangang masakan, Lawara penjah masakan, Lema/ Rebung asam, Lemon segar, Lemon Squasih, Lemonade, Leunca buah, Lilin bungkus gedi, Lobak, Lokan segar, Lontar segar, Lumai/Leunca segar, Mangga benggala segar, Oramu ninahu ndawa olaho masakan, Otak, Paku hinela wulelenggapaya, Pala daging, Papais, Papeda, Parede baleh masakan, Paria (Pare), Paria Putih kukus, Paria putih segar, Pe-Cay, Peda Ikan Banjar, Pelecing kangkong, Pelepah manuk masakan, Pempek adaan, Pempek kapal selam, Pempek kelesan, Penyu serapah masakan, Pepare ular segar, Pepaya Muda, Pepaya segar, Pucuk lumai/daun leunca segar, Punai daging goreng, Pundut nasi, Purundawa, Purut segar, Putri malu segar, Putri selat, Ragi, Rajungan segar, Rambutan, Rambutan Aceh, Rambutan binjai segar, Rambutan sinyonya, Rawon masakan, Rebon (udang kecil segar), Rebung, Rendang sapi masakan, Rimbang segar, Rujak cingur, Rumput laut, Rusip, RW anjing masakan, Uceng/ bunga melinjo segar, Udang besar segar, Udang galah segar, Udang segar, Ulat sagu segar, Umbut rotan, Usus Sapi, Uwi, Waluh balamak, Woku ubi, Wortel Segar, Wortel kukus, Wortel rebus, Yoghurt.
2	Abon, Abon haruwan, Ampas Tahu, Angsa, Arwan sirsir, Ayam, Ayam goreng Kentucky sayap, Babi hutan masak rica masakan, Bebek (itik), Bebek alabio daging segar, Bebek daging goreng, Belut goreng, Buaya daging dendeng mentah, Bubur, Bungkil Kacang Tanah, Cassavastick, Coklat Pahit batang, Corned Beef, Cumi-cumi goreng, Daging Babi Gemuk, Daging Babi Kurus, Dendeng belut goreng, Dendeng Daging Sapi, Dendeng mujahir goreng, Domba daging gemuk segar, Empal (daging) Goreng masakan, Empal Goreng, Enting-enting wijen, Gendar goreng, Ham, Itik alabio daging dendeng mentah, Jambal goreng, Kacang atom, Kacang belimbing (kecipir) kering, Kacang goyang, Kacang Kedelai basah, Kaholeo masakan, Kecipir biji, Keju, Keju Kacang Tanah, Kelapa hutan kering, Kerupuk urat, Kluwek, Kripik Tempe Goreng, Kue kelapa, Kwaci, Kwark (Quark), Lamtoro var. gung tanpa kulit, Pala biji, Paniki masak santan masakan, Pencok lele masakan, Pisang Siam goreng, Rebon kering, Rusa daging dendeng mentah, Saga biji tanpa kulit, Santan murni, Sapi daging dendeng mentah, Sapi daging gemuk segar, Sardines dalam kaleng, Sari Kedelai bubuk, Sate Bandeng, Sate pusut masakan, Sie reuboh masakan, Susu bubuk, Tahu telur, Teri balado masakan, Teripang dendeng mentah, Terong + Oncom makanan, Tinoransak masakan, Toge -Tahu makanan, Udang kering, Udang kering mentah, Udang papay mentah, Ulat sagu panggang, Wijen, Worst (sosis daging).

Table 7. An Explanation Of Every Cluster

Cluster	Description
0	From 100 grams of food, this cluster found that the food had a high calorie content (60–720 calories), low protein and fat content (average <50 mg), and high carbohydrate content (>300 mg)
1	This cluster discovered that, out of 80 grams of food, the average protein and fat content was between 2 and 50 mg, the calorie content was between 100 and 500 kcal, and the carbohydrate level was less than 5 mg.
2	This cluster discovered that the dietary data had enough carbohydrate levels (average of over 1 mg) from 90 grams of food, protein levels below 60 mg, fat levels below 40 mg, and calorie levels between 4 kcal and 300.

Three clusters with disparate outcomes are derived from the K-Medoids algorithm's implementation results.

Foods with high calorie and carbohydrate content are found in Cluster 0. They have minimal quantities of fat and protein. According to this cluster, food data has a high calorie content (60–720 calories), low protein and fat content (average <50 mg), and high carbohydrate content (>300 mg) per 100 grams of food.

Foods having a high calorie content and relatively low protein, fat, and carbohydrate contents are also abundant in cluster 1. According to this cluster, 80 grams of food had a high calorie level (above 100 kcal to 500 kcal), an average protein and fat content (above 2 mg to 50 mg), and a carbohydrate content (below 5 mg).

Foods with very high calorie content and relatively low levels of protein, fat, and carbohydrates make up the final cluster, cluster 2. According to this cluster, dietary data had nutritional levels ranging from 4 to 300 calories, less than 60 mg of protein, less than 40 mg of fat, and more than 1 mg of carbs on average from 90 grams of food.

In K-Medoids clustering, the elbow approach visualizes an "elbow" graph made from the SSE (sum of squared errors) values against different numbers of potential clusters in order to discover the ideal number of clusters. The elbow points on this graph show how many clusters are best suited to the data. Alternatively, the K-Medoids algorithm can be used to assess the quality of the clustering it produces by using the DBI (Davies-Bouldin Index) approach. DBI is used to assess how well data lies within a cluster (intra-cluster) and how well data separates clusters (inter-cluster). The higher the quality of clustering generated by K-Medoids, the lower the DBI score.

IV. CONCLUSION

The ideal number of clusters is three, according to research using K-Medoids to categorize meals according to nutritional values from the selection phase to testing. The Davies-Bouldin Index Cluster and the Elbow Method are used in the cluster optimization process to calculate this number. Foods with high calorie and carbohydrate content are found in Cluster 0.

They have minimal quantities of fat and protein. Foods with high calorie content and relatively low protein, fat, and carbohydrate content are also prevalent in cluster 1. Foods with a high calorie content and relatively low proportions of protein, fat, and carbohydrates make up the final cluster, cluster 2.

The Elbow and Davies-Bouldin Index methods are both employed in the K-Medoids algorithm. The Elbow method yields a value of 0.046, which is utilized as an elbow point, while the Davies-Bouldin Index approach yields the smallest value, 0.631.

REFERENCES

[1] P. Alga Vredizon, H. Firmansyah, N. Shafira Salsabila, and W. Eko Nugroho, "Penerapan Algoritma K-Means Untuk Mengelompokkan Makanan Berdasarkan Nilai Nutrisi," *J. Technol. Informatics*, vol. 5, no. 2, pp. 108–115, 2024, doi: 10.37802/joti.v5i2.577.

[2] A. Pratama, "Noodle Grouping Based On Nutritional Similarity With Hierarchical Cluster Analysis Method," *Anal. Kekuatan Kontruksi Rangka pada Peranc. Belt Conveyor Menggunakan Ansys Work.*, vol. 2, no. 2, pp. 58–71, 2023.

[3] I. W. Septiani, A. C. Fauzan, and M. M. Huda, "Implementasi Algoritma K-Medoids Dengan Evaluasi Davies-Bouldin-Index Untuk Klasterisasi Harapan Hidup Pasca Operasi Pada Pasien Penderita Kanker Paru-Paru," *J. Sist. Komput. dan Inform.*, vol. 3, no. 4, p. 556, 2022, doi: 10.30865/json.v3i4.4055.

[4] A. Hoerunnisa, G. Dwilestari, F. Dikananda, H. Sunana, and D. Pratama, "Komparasi Algoritma K-Means Dan K-Medoids Dalam Analisis Pengelompokan Daerah Rawan Kriminalitas Di Indonesia," *JATI (Jurnal Mhs. Tek. Inform.*, vol. 8, no. 1, pp. 103–110, 2024, doi: 10.36040/jati.v8i1.8249.

[5] B. Indikator and I. P. M. Dengan, "Analisis perbandingan silhouette coefficient dan metode elbow pada pengelompokkan provinsi di indonesia berdasarkan indikator ipm dengan k-medoids 1,2,3," vol. 13, pp. 13–24, 2024, doi: 10.14710/j.gauss.13.1.13-24.

[6] E. A. Saputra and Y. Nataliani, "Analisis Pengelompokan Data Nilai Siswa untuk Menentukan Siswa Berprestasi Menggunakan Metode Clustering K-Means," *J. Inf. Syst. Informatics*, vol. 3, no. 3, pp. 424–439, 2021, doi: 10.51519/journalisi.v3i3.164.

[7] H. S. Kasman, "Perbandingan Efektifitas Algoritma K-Means Clustering -Topsis dan K-Medoids Clustering -Topsis dalam Menentukan Karyawan IT dengan Kinerja Terbaik A . Peneliti Terdahulu K-Means merupakan metode clustering non-hierarkis yang berusaha membagi data ke dalam," vol. 4, no. 2, 2024.

[8] B. A. Setiawan and Sulastri, "Perbandingan Clustering Optimalisasi Stok Barang Menggunakan Algoritma K – Means Dan Algoritma K – Medoids," pp. 978–979, 2021, [Online]. Available: https://unisbank.ac.id/ojs/index.php/sendi_u/article/view/8645

[9] P. Vania and B. Nurina Sari, "Perbandingan Metode Elbow dan Silhouette untuk Penentuan Jumlah Klaster yang Optimal pada Clustering Produksi Padi menggunakan Algoritma K-Means," *J. Ilm. Wahana Pendidik.*, vol. 9, no. 21, pp. 547–558, 2023, [Online]. Available: https://doi.org/10.5281/zenodo.10081332

[10] N. A. Maori and E. Evanita, "Metode Elbow dalam Optimasi Jumlah Cluster pada K-Means Clustering," *Simetris J. Tek. Mesin, Elektro dan Ilmu Komput.*, vol. 14, no. 2, pp. 277–288, 2023, doi: 10.24176/simet.v14i2.9630.

[11] F. Irhamni, F. Damayanti, B. Khusnul K, and M. A, "Optimalisasi Pengelompokan Kecamatan Berdasarkan Indikator Pendidikan Menggunakan Metode Clustering dan Davies Bouldin Index," *Semin. Nas. dan Teknol. UMJ*, no. 11, pp. 1–5, 2014.

[12] I. T. Umagapi, B. Umaternate, S. Komputer, P. Pasca Sarjana Universitas Handayani, B. Kepegawaian Daerah Kabupaten Pulau Morotai, and B. Riset dan Inovasi, "Uji Kinerja K-Means Clustering Menggunakan Davies-Bouldin Index Pada Pengelompokan Data Prestasi Siswa," *Semin. Nas. Sist. Inf. dan Teknol.*, vol. 7, no. 1, pp. 303–308, 2023.

[13] I. F. Fauzi, M. Gito Resmi, and T. I. Hermanto, "Penentuan Jumlah Cluster Optimal Menggunakan Davies Bouldin Index pada Algoritma K-Means untuk Menentukan Kelompok Penyakit," vol. 7, no. 2, pp. 2598–8069, 2023.

[14] A. Atira and B. Nurina Sari, "Penerapan Silhouette Coefficient, Elbow Method dan Gap Statistics untuk Penentuan Cluster Optimum dalam Pengelompokkan Provinsi di Indonesia Berdasarkan Indeks Kebahagiaan," *J. Ilm. Wahana Pendidik.*, vol. 9, no. 17, pp. 76–86, 2023, [Online]. Available: https://doi.org/10.5281/zenodo.8282638

[15] A. R. Aruadi, F. Andreas, and B. N. Sari, "Analisis Klaster K-Means pada Data Rata-Rata Konsumsi Kalori dan Protein Menurut Provinsi dengan Metode Davies Bouldin Index," *J. Pendidik. dan Konseling*, vol. 4, no. 4, pp. 5652–5661, 2022.

[16] D. Nasution, D. N. Sirait, I. Wardani, and Dwiyanto, "Optimasi Jumlah Cluster Metode K-Medoids Berdasarkan Nilai DBI Pada Pengelompokkan Data Luas Tanaman Dan Produksi Kelapa Sawit Di Sumatera Utara," *Kumpul. J. Ilmu Komput.*, vol. 9, no. 2, p. 381, 2022.