

## Penerapan Algoritma *K-Means* untuk Mengelompokkan Makanan Berdasarkan Nilai Nutrisi

Prayoga Alga Vredizon<sup>1</sup>, Hasbi Firmansyah<sup>2</sup>, Nadya Shafira Salsabila<sup>3</sup>, Wildani Eko Nugroho<sup>4</sup>

<sup>1,2</sup>Informatika, Universitas Pancasakti Tegal, Tegal, Indonesia

<sup>3</sup>Teknik Sipil, Universitas Pancasakti Tegal, Tegal, Indonesia

<sup>4</sup>D3 Komputer, Politeknik Harapan Bersama, Tegal, Indonesia

e-mail : yogavredizon23@gmail.com<sup>1</sup>, hasbifirmansyah@upstegal.ac.id<sup>2</sup>, nadyashafira@upstegal.ac.id<sup>3</sup>, wild4n1@gmail.com<sup>4</sup>

\*Penulis Korespondensi: E-mail: yogavredizon23@gmail.com

**Abstrak:** Penelitian ini bertujuan untuk mengelompokkan makanan yang memiliki nilai nutrisi yang serupa. Yang mana makanan dibagi ke dalam 3 *cluster* yaitu makanan yang mempunyai kadar nutrisi tinggi, sedang dan rendah. Hasil pengelompokan pada penelitian ini dapat digunakan untuk pemilihan dan konsumsi makanan dalam pemenuhan nutrisi dan juga dapat digunakan untuk mencegah timbulnya penyakit yang disebabkan oleh makanan. Seperti makanan pada *cluster* 0 bisa dipilih jika ingin menaikkan berat badan. Makanan *cluster* 1 dapat menjadi patokan jika dikonsumsi terlalu banyak dapat menyebabkan obesitas dan *cluster* 2 dapat dipilih jika ingin melakukan diet atau menurunkan berat badan. Hasil ini ditunjukkan dari hasil klusterisasi di mana *cluster* pertama diisi oleh makanan dengan kadar kalori dan protein yang cukup tinggi dan kadar lemak, karbohidrat yang rendah. *Cluster* kedua diisi oleh makanan dengan kadar kalori, protein dan lemak yang tinggi serta kadar karbohidrat yang rendah. *Cluster* ketiga diisi oleh makanan dengan kadar kalori, protein, lemak dan karbohidrat yang rendah. Penelitian ini menggunakan metode *clustering* dengan menerapkan algoritma *K-Means* karena efektif dalam melakukan klusterisasi terhadap tipe data numerik dan menguji menggunakan *Elbow Method* dan *Davies Bouldin Index*.

**Kata Kunci:** *Clustering; K-Means; Nutrisi*

**Abstract:** The purpose of this research is to group foods that have similar nutritional value. Food is divided into 3 clusters, namely those with high, medium and low nutritional levels. The grouping results in this research can be used for selecting and consuming food to fulfil nutritional needs and can also be used to prevent the emergence of diseases caused by food. For example, foods in cluster 0 can be chosen if you want to gain weight. Cluster 1 foods can be a benchmark if consumed too much can cause obesity and cluster 2 can be chosen if you want to go on a diet or lose weight. This result is shown from the clustering results where first cluster filled by food with high enough calories and protein and low levels of fat and carbohydrates. The second cluster is filled by foods with high levels of calories, protein and fat and low levels of carbohydrates. The third cluster is filled by foods with low levels of calories, protein, fat and carbohydrates. This research uses a clustering method by applying a *K-Means* algorithm because it is effective to clustering numeric data types and testing using *Elbow Method* and *Davies Bouldin Index*.

**Keywords:** *Clustering; K-Means; Nutrition*

### PENDAHULUAN

Makanan menjadi salah satu sumber utama demi keberlangsungan hidup manusia. Di Indonesia sendiri terdapat berbagai macam aneka makanan dari berbagai macam daerah. Sebagai contoh, di Medan mempunyai *bacang*, Maluku mempunyai *bagea*, Makassar mempunyai *palubasa*, Semarang mempunyai *lumpia*, dan masih banyak aneka macam lainnya. Makanan tersebut mempunyai karakteristik, tekstur dan rasa yang berbeda-beda.

Makanan sendiri diolah dari berbagai macam bahan, baik dari nabati maupun hewani. Maka dari itu, makanan mengandung banyak nutrisi di antaranya seperti kalori, protein, lemak dan karbohidrat. Nutrisi sangat diperlukan oleh manusia untuk melakukan kegiatan sehari-hari agar manusia dapat hidup sehat[1]. Manusia bisa mendapatkan nutrisi dari daging hewan ataupun tumbuhan. Daging hewan menyediakan banyak nutrisi seperti protein, lemak, karbohidrat, dan juga beberapa nutrisi lainnya seperti

vitamin dan mineral. Tidak hanya dari daging hewan saja, Nutrisi yang dibutuhkan tubuh juga dapat diperoleh dari telur. Telur mengandung 75% air, 12% protein, dan 12% lipid[2]. Nutrisi juga bisa didapatkan dari tumbuhan seperti sayur-sayuran, umbi-umbian, dan juga kacang-kacangan. Nutrisi yang dihasilkan oleh tumbuhan mendominasi protein sebesar 57%, lebih besar 43% daripada daging hewan. Yang mana sebesar 10% dari kerang, 6% dari ikan, 18% dari daging, dan hewan lainnya sebesar 9%[2].

Selain olahraga dan pola hidup sehat, nutrisi makanan juga memegang peranan penting bagi kesehatan manusia. Nutrisi dari makanan sangat diperlukan untuk memenuhi kebutuhan manusia. Jika makanan dikonsumsi dalam kadar nutrisi yang cukup maka makanan akan sangat bermanfaat dan membantu menguatkan metabolisme tubuh manusia[3]. Namun, nutrisi yang berlebihan juga dapat menjadi pemicu penyakit berbahaya seperti obesitas, penyakit jantung, stroke, hipertensi, dan kanker. Kondisi ini menjadi

penyebab utama kematian didunia, terhitung sebesar 70% kematian di seluruh dunia yang disebabkan oleh faktor makanan[4]. *The Principle Balance Nutrition* adalah cara yang sesuai untuk menanggulangi hal tersebut. Prinsip ini digunakan untuk memperbaiki kesehatan dan mengawasi berat badan pada tubuh dengan mengubah gaya hidup dengan mempertimbangkan kandungan nutrisi pada makanan sesuai dengan yang dibutuhkan oleh tubuh[5].

Berdasarkan latar belakang tersebut, peneliti tertarik melakukan penelitian untuk mengelompokkan makanan berdasarkan kandungan nutrisinya seperti kandungan karbohidrat, kalori, protein dan lemak. Pengelompokan ini bertujuan untuk memudahkan dalam pemilihan dan konsumsi makanan dalam pemenuhan nutrisi dan juga dapat mencegah terjadinya penyakit yang disebabkan oleh makanan. Penelitian ini juga bertujuan untuk mengubah gaya hidup dengan mempertimbangkan kandungan nutrisi pada makanan sesuai dengan yang dibutuhkan oleh tubuh.

Pengelompokan pada penelitian ini menggunakan salah satu algoritma pada metode *clustering* yang sesuai untuk mengolah data numerik yaitu Algoritma *K-Means*. Algoritma *K-Means* merupakan salah satu metode *clustering non-hirarki* yang paling membagi data ke dalam *cluster* tanpa memerlukan label ataupun kategori[6][7]. Masalah yang dihadapi pada algoritma ini adalah menentukan jumlah *cluster* yang optimal. Hal ini dapat menyebabkan kurang tepatnya akurasi yang dihasilkan oleh Algoritma *K-Means*. Oleh karena itu, beberapa metode pengujian dilakukan untuk menentukan jumlah *cluster* di antaranya menggunakan *Elbow Method* dan *Davies Bouldin Index*. Kedua metode tersebut digunakan karena dapat mengukur dan memvalidasi dalam menentukan jumlah *cluster* yang baik dan optimal[6].

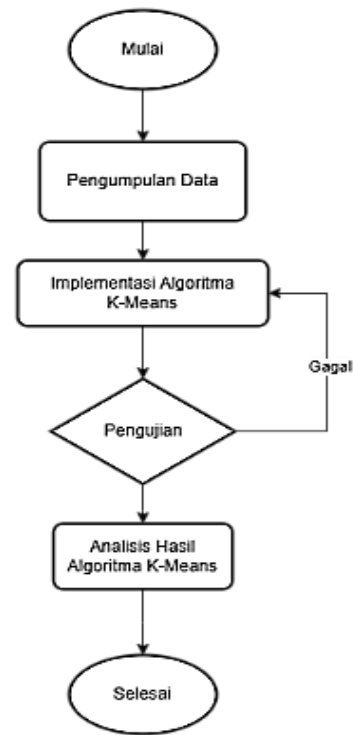
## METODE PENELITIAN

Sebelum melakukan penelitian, diperlukan gambaran bagaimana penelitian yang akan dilakukan dengan mengidentifikasi tahapan-tahapannya. Menentukan tahapan-tahapan sangat penting agar penelitian yang dilakukan dapat diaplikasikan secara rinci dan teratur[8].

Pada penelitian ini menerapkan beberapa tahapan untuk memperoleh hasil klasterisasi yang optimal. Di antaranya terdapat pengumpulan data, melakukan operasi Algoritma *K-Means*, melakukan pengujian dan menganalisis hasil yang diperoleh dari operasi Algoritma *K-Means*.

Pada penelitian dilakukan beberapa kali pengujian hingga menemukan hasil yang optimal. Hal ini juga bertujuan untuk meningkatkan akurasi tiap *cluster* dan me-minimalisir kesalahan pada penelitian ini.

Adapun diagram alur penelitian yang menjadi dasar dalam isi dan pembahasan seperti pada Gambar 1.



Gambar 1. Alur Metode Penelitian

## Pengumpulan Data

Pada penelitian ini *dataset* diperoleh dari pihak lain, data bersumber dari situs dataset Kaggle (<https://kaggle.com/datasets/anasfikrihanif/indonesian-food-and-drink-nutrition-dataset>) yang diunggah oleh Anas Fikri Hanis pada 1 November 2023 dengan 1346 daftar nutrisi per 100gram makanan di Indonesia. Gambar 2 merupakan cuplikan beberapa dataset nutrisi yang didapatkan dari situs Kaggle.

id	calories	proteins	fat	carbohydrate	name	image
1	280	9.200	28.400	0	Abon	<a href="https://img-cdn.medikontak.com/PhvY9X3jgn08vV">https://img-cdn.medikontak.com/PhvY9X3jgn08vV</a>
2	513	23.700	37	21.300	Abon haruwan	<a href="https://img-global.cpcdn.com/recipes/c8f530bd1bc">https://img-global.cpcdn.com/recipes/c8f530bd1bc</a>
3	0	0	0.200	0	Agar-agar	<a href="https://res.cloudinary.com/9k5z4ums3/image/upload">https://res.cloudinary.com/9k5z4ums3/image/upload</a>
4	45	1.100	0.400	10.800	Akar tonjeng segar	<a href="https://images.ikopedia.net/img/cache/200-square">https://images.ikopedia.net/img/cache/200-square</a>
5	37	4.400	0.500	3.800	Alatoge segar	<a href="https://imgagzi.com/assets/images/produk/produk_">https://imgagzi.com/assets/images/produk/produk_</a>
6	85	0.900	6.500	7.700	Alpukat segar	<a href="https://ikatakabar.com/assets/images/upload/news">https://ikatakabar.com/assets/images/upload/news</a>
7	96	3.700	0.600	19.100	Ampas kacang hijau	<a href="https://images.ikopedia.net/img/cache/215-square">https://images.ikopedia.net/img/cache/215-square</a>
8	414	26.600	18.300	41.300	Ampas Tahu	<a href="https://pdpres.dewire.id/upload/9eb1ba592ac72">https://pdpres.dewire.id/upload/9eb1ba592ac72</a>
9	75	4.100	2.100	10.700	Ampas tahu kukus	<a href="https://icdn.diadona.id/dadona/id/resize/640x320">https://icdn.diadona.id/dadona/id/resize/640x320</a>
10	67	5	2.100	8.100	Ampas tahu mentah	<a href="https://icdn-image.hipwee.com/wp-content/uploads">https://icdn-image.hipwee.com/wp-content/uploads</a>
11	184	18.800	14	0	Anak sapi daging gemuk segar	<a href="https://img.pngtree.com/png-clipart/20220124/koum">https://img.pngtree.com/png-clipart/20220124/koum</a>
12	174	19.600	10	0	Anak sapi daging kurus segar	<a href="https://asset.kompas.com/crops/BzqLdWVnY1u1">https://asset.kompas.com/crops/BzqLdWVnY1u1</a>
13	190	19.100	12	0	Anak sapi daging sedang segar	<a href="https://koran-jakarta.com/images/article/tips-memil">https://koran-jakarta.com/images/article/tips-memil</a>
14	99	4.600	1	18	Andalman segar	<a href="https://cf.shopee.co.id/file/3c33b2480e951a2634f">https://cf.shopee.co.id/file/3c33b2480e951a2634f</a>
15	25	1.600	0.200	5.300	Andewi	<a href="https://www.sabuharapan.com/uploads/cache/news">https://www.sabuharapan.com/uploads/cache/news</a>
16	30	0.500	0.200	6.800	Anggur hutan segar	<a href="https://cf.shopee.co.id/file/ed7f05c6de5f13ec39e9f">https://cf.shopee.co.id/file/ed7f05c6de5f13ec39e9f</a>
17	354	16.400	31.500	0	Anona	<a href="https://icdn.dindin.com/content-images/crommoun">https://icdn.dindin.com/content-images/crommoun</a>

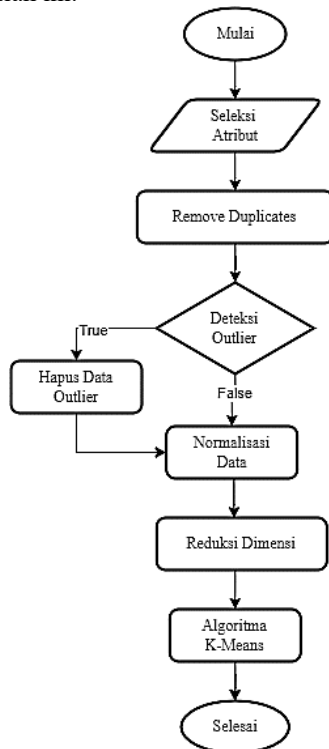
Gambar 2. Dataset Nutrisi Situs Kaggle

## Implementasi Algoritma K-Means

Sebelum implementasi Algoritma *K-Means*, diperlukan beberapa tahapan *preprocessing*. Hal ini penting untuk dilakukan karena dapat mengoptimalkan hasil yang diperoleh dari implementasi Algoritma *K-Means*.

Pada penelitian ini dilakukan beberapa tahapan *preprocessing*. Di antaranya adalah seleksi atribut, *Remove Duplicates*, Deteksi *Outlier*, Normalisasi Data dan Reduksi Dimensi. Semua tahapan *preprocessing* ini dilakukan karena menyesuaikan dengan *dataset* yang digunakan agar Algoritma *K-Means* dapat berjalan dengan baik dan menghasilkan *output* yang sesuai dan akurat.

Gambar 3. merupakan diagram alur yang akan dilakukan dalam penerapan Algoritma *K-Means* pada penelitian ini.



Gambar 3. Alur Algoritma *K-Means*

### Seleksi Atribut

Seleksi atribut dilakukan untuk memilih atribut yang akan digunakan dalam algoritma *K-Means* dan menghapus atribut yang tidak digunakan. Karena *K-Means* hanya bisa menerima atribut yang memiliki data dengan tipe *numeric*. Maka dari itu atribut yang bukan *numeric* harus diseleksi[5].

Tabel 1. Seleksi Atribut

Kolom	Keterangan
<i>Calories</i>	Kalori per 100gram/makanan (kal)
<i>Protein</i>	Protein per 100gram/makanan (gram)
<i>Fat</i>	Lemak per 100gram/makanan (gram)
<i>Carbohidrat</i>	Karbohidrat per 100gram/makanan (gram)

(Sumber: Tim Peneliti)

### Remove Duplicates

*Remove Duplicates* digunakan untuk menghapus data yang mempunyai nilai yang sama

dengan membandingkan semua data satu per satu. Dua data yang diindikasikan mempunyai nilai yang sama akan dihapus salah satu dan menyisakan satu data dari hasil perbandingan.

Pada penelitian ini dilakukan *remove duplicate* karena setiap makanan biasanya mempunyai nilai nutrisi yang sama. Maka dari itu, diputuskan untuk menghapus data yang duplikat agar operasi menjadi lebih optimal.

### Deteksi Outlier

Data *outlier* sangat mempengaruhi kinerja *K-Means*, sehingga deteksi *outlier* diperlukan untuk mencari data di mana data tersebut mempunyai perilaku yang berbeda dibandingkan data-data normal yang lain agar operasi algoritma *K-Means* menjadi lebih optimal[9].

Pendeteksi *outlier* pada penelitian ini menggunakan *Distance based Outlier Detection*. Dengan *Distance Based* jarak antara dua titik dipilih kemudian akan dicek. Jika titik *neighbour* dekat maka situasinya dianggap normal, tetapi jika titik *neighbour*nya terpisah jauh maka akan dianggap sebagai *Outliers*[10].

Setelah dilakukan pengujian, parameter yang tepat untuk digunakan pada penelitian ini menggunakan jarak 10 *neighbour*, 9 *outlier* dan menggunakan *Euclidean Distance* untuk mengukur jarak.

### Normalisasi Data

Normalisasi data yang digunakan menggunakan normalisasi *min-max*. Normalisasi *min-max* mengubah nilai data aslinya menjadi nilai yang berada pada rentang 0 sampai 1[9]. Tujuannya adalah membentuk data ke dalam nilai dengan rentang yang sama sehingga operasi menjadi lebih cepat[11].

Rumus Normalisasi *Min-max*[9] :

$$V_{norm} = \left( \frac{V_i - V_{min}}{V_{max} - V_{min}} \right) \quad (1)$$

Keterangan:

- v : atribut data
- $V_{min}, V_{max}$  : nilai absolut minimal dan maksimal dari v
- $V_i$  : nilai lama dari setiap entri dalam data

### Reduksi Dimensi

Metode yang digunakan untuk mereduksi dimensi suatu data dengan banyak *variable* adalah menggunakan *Principle Component Analysis (PCA)*. *PCA* mereduksi dimensi suatu data tanpa mengurangi informasi pada data tersebut secara signifikan[12]. *PCA* membentuk sekumpulan dimensi baru yang diberi peringkat berdasarkan perbedaan data. *PCA* menghasilkan komponen utama yang diperoleh dari dekomposisi nilai *eigen* dan vektor *eigen* matriks *kovarians*[13].

Pada penelitian ini, parameter yang digunakan untuk implementasi *PCA* menggunakan

*Fixed Number Dimensity* dengan 2 *Number of Component*. Di mana atribut yang sebelumnya sejumlah 4 kemudian direduksi menjadi 2 atribut.

**Algoritma K-Means**

K-Means merupakan metode *clustering non-hierarki* yang membagi data *n* ke dalam *k cluster*, sehingga hasil kesamaan antar *clusternya* tinggi jika *Within-Cluster Sum of Squarenya* minimal dan hasil kesamaan di dalam *clusternya* rendah jika *Between-Cluster Sum of Squarenya* maksimal[6].

K-Means merupakan *unsupervised learning algorithm*, artinya *input* data tidak mempunyai label atau tidak menyediakan kategori. Kelompok atau *cluster* pada K-Means direpresentasikan sebagai nilai K[7].

Pengelompokan data pada penelitian ini menggunakan algoritma K-Means dengan beberapa proses berikut[14] :

- 1) Menentukan nilai K
- 2) Memilih objek secara *random* pada *dataset* yang akan dijadikan sebagai titik pusat *cluster*. Titik pusat disebut juga sebagai *centroid*.
- 3) Tetapkan data ke *centroid* terdekat menggunakan rumus *Euclidian Distance*:

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

Keterangan:

n : jumlah dimensi  
 x, : atribut ke-k data x  
 y dan y

- 4) Alokasi data ke *cluster* terdekat dengan rumus :

$$\min \sum_{k=1}^k d(x,y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (3)$$

- 5) Menghitung pusat *cluster*

$$y_i = \frac{\sum_{i=1}^p c_i}{p} \quad (4)$$

Keterangan:

c<sub>i</sub> : data *cluster* ke-k  
 p : Jumlah data *cluster* ke-k

**Pengujian**

Pengujian dilakukan untuk melihat tingkat keberhasilan dan keakuratan dari algoritma K-Means. Jumlah *cluster* dapat ditentukan dengan beberapa cara, di antaranya adalah menggunakan *Elbow Method* dan *Davies Bouldin Index*[6].

**Elbow Method**

*Elbow Method* merupakan langkah dasar pada semua algoritma *unsupervision* dalam menentukan jumlah kluster untuk menentukan nilai optimal yang digunakan pada nilai K pada algoritma K-Means [7]. *Elbow Method* menjalankan nilai kluster dari 1 sampai *n*, setiap nilai pada *With Cluster Sum of Square* (WCSS) akan dihitung. WCSS merupakan jarak *sum of square* antara nilai rata-rata dengan nilai *centroid* dari setiap *cluster* yang diberikan. kemudian pilih titik yang paling membentuk siku dengan menggunakan *chartline*[6].

$$WCSS = \sum_{j=1}^k \sum_{i=1}^n ||x_i^{(j)} - c_j||^2 \quad (5)$$

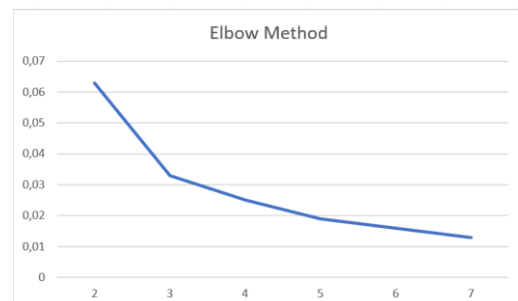
Keterangan:

k: nomor *cluster*,  
 n: nomor *object*,  
 x<sub>i</sub>: i<sup>th</sup> elemen pada *cluster*  
 c<sub>j</sub>: *centroid* dari *cluster* j

Tabel 2. Nilai K dan Avg. *Within-cluster Sum of Square*

K	WCSS
2	0.063
3	0.033
4	0.025
5	0.019
6	0.016
7	0.013

(Sumber: Tim Peneliti)



Gambar 4. Penentuan Jumlah *Cluster* untuk Algoritma K-Means

**Davies Bouldin Index**

*Davies Bouldin Index* (DBI) merupakan validasi *cluster* mencakup nilai separasi dan kohesi. Separasi adalah jarak antara *centroid* dari *cluster* dan kohesi adalah jumlah kemiripan data dengan *centroid cluster*[6]. Hasil klasterisasi yang baik ditentukan oleh jumlah dan kedekatan data yang di klasterisasi. Semakin kecil nilai DBI (non-negatif => 0), maka semakin baik hasil *cluster* yang diperoleh[8].

$$DBI = \frac{1}{k} \sum_{i=1}^k \max_{i \neq j} (R_{i,j}) \quad (6)$$

Keterangan:

k : Jumlah *cluster*  
 R<sub>i,j</sub> : rasio antar kluster  
 max : dicari rasio antar kluster yang terbesar

Tabel 3. Nilai *Davies Bouldin Index*

K	DBI
2	0.746
3	0.643
4	0.703
5	0.736
6	0.751
7	0.791

(Sumber: Tim Peneliti)

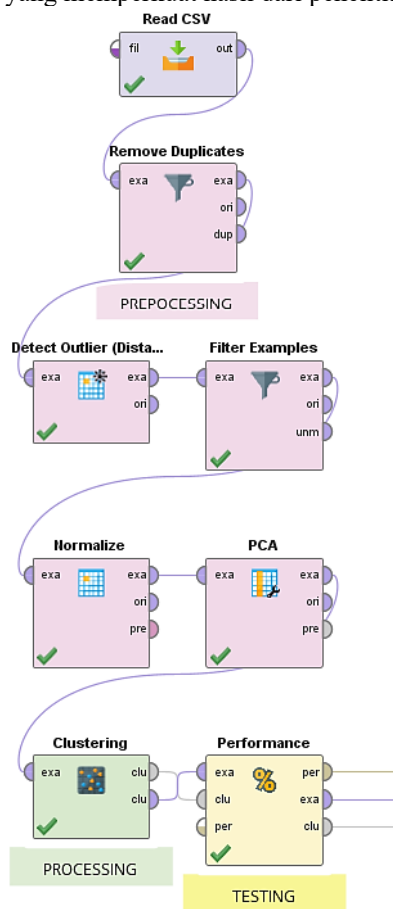
Berdasarkan *Elbow Method* dan *Davies Bouldin Index* yang telah diimplementasi didapati bahwa jumlah *cluster* yang terbaik terdapat pada angka 3.

**Analisis Hasil**

Menganalisis hasil yang sudah diimplementasi dari metode *clustering* dengan algoritma *K-Means*.

**HASIL DAN PEMBAHASAN**

Gambar 5. merupakan implementasi klusterisasi dengan Algoritma *K-Means* menggunakan program *Rapidminer Studio*. Berdasarkan penelitian dan pengujian, diperoleh hasil yang akan dipaparkan pada bagian ini. Selain itu, terdapat juga tabel dan visualisasi yang memperkuat hasil dari penelitian.



Gambar 5. Implementasi *K-Means* dengan *Rapidminer*

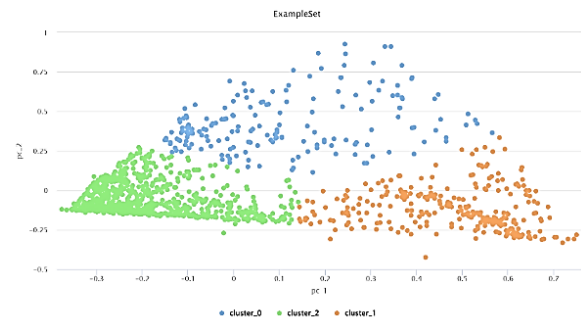
Implementasi algoritma *K-Means* dengan beberapa tahapan *preprocessing* seperti seleksi data, *remove duplicates*, normalisasi data, dan reduksi dimensi serta optimalisasi jumlah *cluster* dengan *Elbow Method* dan *Davies Bouldin Index* didapati bahwa jumlah *cluster* terbaik atau paling optimal ada pada angka 3.

Tabel 4. Nilai *WCSS*

Cluster	Avg. Within Centroid Distance
0	0.063
1	0.038
2	0.024

(Sumber: Tim Peneliti)

Dari data Tabel 4. menunjukkan bahwa nilai *Average Within Centroid Distance* mendekati angka 0 yang menandakan bahwa kedekatan antar data di dalam *cluster* sudah baik. Untuk lebih mudah melihat visualisasi dapat menggunakan *scatter plot*.



Gambar 6. *Scatter Plot* Hasil Algoritma *K-Means* dengan 3 *Cluster*

Setelah dilakukan operasi menggunakan algoritma *K-Means* didapati hasil klusterisasi seperti pada Tabel 5 dan 6.

Tabel 5. Jumlah Data Tiap *Cluster*

Cluster	Jumlah Data
0	282
1	182
2	802

(Sumber: Tim Peneliti)

Tabel 6. Beberapa Data Makanan Hasil Klusterisasi

Cluster	Daftar Makanan
0	Asam masak di pohon, Asan kandis kering, Bagea kelapa, Bagea kenari, Bakwan, Bantal, Beras Cerdas, Beras ganyong, Beras giling varian pelita mentah, Beras hitam mentah, Beras jagung kuning kering mentah, Beras jagung putih kering mentah, Beras Ketan Hitam, Beras Ketan Hitam tumbuk, Beras Ketan Putih, Ikan Asin bilis goreng, Ikan Asin gabus goreng, Ikan Asin japuh goreng, Ikan Asin pari goreng, Ikan Asin pepetek goreng, Ikan Asin sirinding goreng, Ikan Asin teri goreng, Ikan bandeng presto masakan,
1	Abon, Abon Haruan, Ampas, Tahu, Angsa, Arwan Sisir, Ayam,

Cluster	Daftar Makanan
	Ayam goreng <i>Kentucky</i> sayap, Ayam ampela goreng, Ayam goreng <i>church</i> texas sayap, Ayam goreng <i>church</i> texas dada, Ayam goreng kalasan paha, Ayam goreng <i>kentucky</i> dada, Ayam goreng <i>Kentucky</i> paha, Ayam goreng mbok berek dada, Ayam goreng paha, Ayam goreng Pasundan dada, Ayam goreng Pasundan paha, Ayam goreng <i>pioneer</i> dada, Ayam goreng Sukabumi dada, Ayam goreng Sukabumi paha, Ayam hati segar, Ayam Taliwang masakan, Ayam usus goreng, Emping beras, Emping komak, Wijen, Worst (sosis daging).
2	Agar-agar, Akar Tonjong segar, Aletoge segar, Alpukat segar, Ampas kacang hijau, Ampas tahu kukus, Ampas tahu mentah, Anak sapi daging gemuk segar, Anak sapi daging kurus segar, Anak sapi daging sedang segar, Andaliman segar, Andewi, Anggur hutan segar, Anyang sayur, Apel, Apel malang segar, Arbei, Ares sayur, Arrowroot segar, Asam kandis segar, Asam payak segar, Asinan Bogor sayuran, Ayam didih/darah segar, Babi ginjal segar, Babi hati segar, Bacang, Baje, Bakso Bakung segar, Barongko, Batang Tading, Batar daan, Batatas gembili segar, Batatas kelapa ubi bakar, Batatas kelapa ubi kukus, Batatas kelapa ubi segar, Batatas tali ubi rebus, Bawang Bombay, Bawang Merah Bawang Putih, Bayam kukus, Bayam Merah, Bayam merah segar, Bayam rebus, Bayam segar Bayam tumis + oncom, Bayam tumis bersantan, Coto mangkasara kuda masakan, Coto mangkasara sapi masakan, Cue Selar Kuning, Cuka, Daun kenikir segar, Daun kesum segar, Daun kol sawi segar, Daun Koro, Daun kubis segar, Daun Kumak, Daun Labu Siam, Daun Labu Waluh, Daun lamtoro segar, Daun leilem segar, Daun Leunca, Daun Lobak Daun Lompong Tales, Daun Mangkokan, Sawo duren segar, Sawo kecil segar, Sayur asem, Sayur bunga papaya, Sayur garu

Cluster	Daftar Makanan
	Sayur koku-koku, Sayur lebu, Sayur lilin-terubuk, Sayur ndusuk, Sayur sop

Tabel 7. Keterangan Tiap Cluster

Cluster	Keterangan
0	Cluster ini didapati bahwa data-data makanan memiliki kadar kalori yang tinggi yaitu 250 hingga 400 lebih kalori, Protein dan lemak dengan rata-rata di bawah 10mg dan Karbohidrat yang tinggi hingga di atas 40mg dari 100gram makanan.
1	Cluster ini didapati bahwa data-data makanan memiliki kadar kalori yang tinggi yaitu di atas 260kal, Protein dengan rata-rata di atas 20mg, lemak di atas 10mg dan Karbohidrat di bawah 5mg dari 100gram makanan
2	Cluster ini didapati bahwa data-data makanan memiliki kadar nutrisi yang rendah yaitu dengan kadar kalori di bawah 100kal, Protein dan lemak dengan rata-rata di bawah 10mg dan Karbohidrat yang cukup dengan rata-rata di atas 10mg dari 100gram makanan

(Sumber: Tim Peneliti)

Dengan hasil yang diperoleh dari implementasi Algoritma *K-Means*, terdapat 3 cluster yang masing-masing mempunyai hasil yang berbeda.

Pada cluster 0 diisi oleh makanan dengan kadar kalori dan karbohidrat yang tinggi dan protein dan lemak yang rendah. Dengan ini makanan pada cluster 0 dapat dipilih sebagai pengganti makanan pokok karena karbohidrat dan kalori berperan sebagai energi utama. Selain itu, juga dapat digunakan untuk menambah berat badan jika dikonsumsi dalam jumlah yang cukup, jika dikonsumsi berlebihan dalam makanan pada cluster 0 dapat mengakibatkan kegemukan[15].

Pada cluster 1 diisi oleh makanan dengan kadar kalori yang tinggi, protein yang cukup tinggi, dan kadar lemak dan karbohidrat yang rendah. Dengan ini makanan pada cluster 1 dapat diperhitungkan jika ingin mengonsumsinya. Walaupun mempunyai kadar protein yang cukup tinggi, namun makanan pada cluster 1 mempunyai kadar kalori yang tinggi. Hal tersebut baik untuk dikonsumsi beberapa kali satu minggu untuk mendapatkan protein yang cukup. Namun, jika dikonsumsi secara berlebihan dan tidak disertai dengan aktivitas yang cukup dapat mengakibatkan obesitas[16].

Pada *cluster 2* diisi oleh makanan dengan kadar nutrisi yang rendah. Makanan pada *cluster* ini baik dikonsumsi untuk melakukan diet atau mengurangi berat badan karena memiliki kadar kalori, karbohidrat dan lemak yang rendah[17]. Namun, jika tidak diseimbangi dengan konsumsi nutrisi yang cukup dapat membuat tubuh lemas dan kekurangan nutrisi[15].

## KESIMPULAN DAN SARAN

Berdasarkan implementasi *K-Means* untuk mengelompokkan makanan berdasarkan nilai nutrisi dari tahapan seleksi hingga pengujian, didapati bahwa jumlah kluster yang optimal berada pada angka 3. Jumlah ini ditentukan menggunakan metode optimalisasi *cluster* dengan *Elbow Method* dan *Davies Bouldin Index Cluster 0* atau pertama didapati data-data makanan dengan kadar kalori, karbohidrat yang tinggi, dan kadar protein, lemak yang rendah. *Cluster 1* atau kedua didapati data-data makanan dengan kadar kalori, protein dan lemak yang tinggi dan kadar karbohidrat yang rendah. *Cluster 2* atau ketiga didapati data-data makanan dengan kadar kalori, protein, lemak dan karbohidrat yang rendah. Ada beberapa tujuan dari penelitian ini, contohnya pada *cluster 0* dapat digunakan untuk memilih makanan sebagai makanan pokok, *cluster 1* dapat diperhatikan karena mempunyai kadar kalori yang tinggi, *cluster 2* dapat dipilih untuk melakukan diet.

## DAFTAR PUSTAKA

- [1] "Noodle Grouping Based On Nutritional Similarity With Hierarchical Cluster Analysis Method", Doi: 10.37577/Sainteks.V%Vi%I.543.
- [2] S. Langyan, P. Yadava, F. N. Khan, Z. A. Dar, R. Singh, Dan A. Kumar, "Sustaining Protein Nutrition Through Plant-Based Foods," *Frontiers In Nutrition*, Vol. 8. Frontiers Media S.A., 18 Januari 2022. Doi: 10.3389/Fnut.2021.772573.
- [3] R. Anjariansyah Dan A. Triayudi, "Clustering Kebutuhan Makanan Untuk Meminimasi Standar Deviasi Angka Kebutuhan Gizi Menggunakan Algoritma K-Means Dan K-Medoids," *Jurnal Media Informatika Budidarma*, Vol. 6, No. 1, Hlm. 597, Jan 2022, Doi: 10.30865/Mib.V6i1.3522.
- [4] S. S. Gropper, "The Role Of Nutrition In Chronic Disease," *Nutrients*, Vol. 15, No. 3. Mdpi, 1 Februari 2023. Doi: 10.3390/Nu15030664.
- [5] S. S. Nagari Dan L. Inayati, "Implementation Of Clustering Using K-Means Method To Determine Nutritional Status," *Jurnal Biometrika Dan Kependudukan*, Vol. 9, No. 1, Hlm. 62, Jun 2020, Doi: 10.20473/Jbk.V9i1.2020.62-68.
- [6] R. Gustriansyah, N. Suhandi, Dan F. Antony, "Clustering Optimization In Rfm Analysis Based On K-Means," *Indonesian Journal Of Electrical Engineering And Computer Science*, Vol. 18, No. 1, Hlm. 470–477, 2019, Doi: 10.11591/Ijeecs.V18.I1.Pp470-477.
- [7] D. I. Durojaye Dan G. Obunadike, "Analysis And Visualization Of Market Segmentation In Banking Sector Using Kmeans Machine Learning Algorithm," *Fudma Journal Of Sciences*, Vol. 6, No. 1, Hlm. 387–393, Apr 2022, Doi: 10.33003/Fjs-2022-0601-910.
- [8] I. W. Septiani, Abd. C. Fauzan, Dan M. M. Huda, "Implementasi Algoritma K-Medoids Dengan Evaluasi Davies-Bouldin-Index Untuk Klasterisasi Harapan Hidup Pasca Operasi Pada Pasien Penderita Kanker Paru-Paru," *Jurnal Sistem Komputer Dan Informatika (Json)*, Vol. 3, No. 4, Hlm. 556, Jul 2022, Doi: 10.30865/Json.V3i4.4055.
- [9] A. Ambarwari, Q. Jafar Adrian, Dan Y. Herdiyeni, "Analysis Of The Effect Of Data Scaling On The Performance Of The Machine Learning Algorithm For Plant Identification," *Jurnal Resti (Rekayasa Sistem Dan Teknologi Informasi)*, Vol. 4, No. 1, Hlm. 117–122, Feb 2020, Doi: 10.29207/Resti.V4i1.1517.
- [10] R. Trivedi Dan N. Srivastava, "Outlier Detection Using K-Means Clustering Algorithm With Different Distance Metrics," *International Journal Of Science And Research*, Doi: 10.21275/Sr22324115201.
- [11] Ahmad Harmain, P. Paiman, H. Kurniawan, K. Kusriani, Dan Dina Maulina, "Normalisasi Data Untuk Efisiensi K-Means Pada Pengelompokan Wilayah Berpotensi Kebakaran Hutan Dan Lahan Berdasarkan Sebaran Titik Panas," *Teknimedia: Teknologi Informasi Dan Multimedia*, Vol. 2, No. 2, Hlm. 83–89, Jan 2022, Doi: 10.46764/Teknimedia.V2i2.49.
- [12] A. Toraismaya, L. R. Sasongko, Dan F. S. Rondonuwu, "Principal Component Dan K-Means Cluster Analysis Untuk Data Spektrum Black Tea Grades Guna Penilaian Kualitas Alternatif," *Journal Of Fundamental Mathematics And Applications (Jfma)*, Vol. 3, No. 2, Hlm. 148–157, Nov 2020, Doi: 10.14710/Jfma.V3i2.8663.
- [13] D. Hedyati Dan I. M. Suartana. 2020. "Penerapan Principal Component Analysis (Pca) Untuk Reduksi Dimensi Pada Proses Clustering Data Produksi Pertanian Di Kabupaten Bojonegoro."
- [14] R. E. Pawening, "Algoritma K-Means Untuk Mengukur Kepuasan Mahasiswa

- Menggunakan E-Learning,” *Journal Of Technology And Informatics (Joti)*, Vol. 3, No. 1, Hlm. 27–33, Okt 2021, Doi: 10.37802/Joti.V3i1.201.
- [15] W. Aulia Dan D. Maharani. 2020. “Karbohidrat Dalam Tubuh: Manfaat Dan Dampak Defisiensi Karbohidrat,” 1302.
- [16] S. Kandinasti, “Obesitas: Pentingkah Memperhatikan Konsumsi Makanan Di Akhir Pekan? Obesity: Is It Important To Concern About Food Intake In The Weekend?,” *Amerta Nutr*, Hlm. 1–12, 2018, Doi: 10.2473/Amnt.V2i4.2018.307-316.
- [17] L. Dwi Asih Dan M. Widyastiti, “Meminimumkan Jumlah Kalori Di Dalam Tubuh Dengan Memperhitungkan Asupan Makanan Dan Aktivitas Menggunakan Linear Programming”, *Ekologia*, Vol. 16 No. 1, April 2016: 38-44