

## CLUSTERING MENU MAKANAN BERDASARKAN KEBUTUHAN KALORI HARIAN MENGGUNAKAN ALGORITME K-MEANS

**Trio Nugroho, Ultach Enri, Iqbal Maulana**

Program Studi Informatika, Universitas Singaperbangsa Karawang  
 Jl. HS. Ronggo Waluyo, Telukjambe Timur, Karawang, Indonesia  
 2010631170123@student.unsika.ac.id

### ABSTRAK

Masyarakat Indonesia masih sangat jauh dari kata pola hidup sehat dengan memilih makanan fastfood atau junkfood untuk dikonsumsi sehari-hari. Hal ini menyebabkan adanya kekurangan ataupun kelebihan kalori harian. Penelitian ini bertujuan untuk menerapkan metode algoritme K-Means dalam menentukan menu makanan yang sesuai dengan kebutuhan kalori harian. Dalam menentukan kebutuhan kalori harian dapat menggunakan Harris-Benedict. Dataset yang digunakan sebanyak 359 data yang digunakan dalam penelitian ini. Data ini diolah menggunakan metode klasterisasi dengan menggunakan algoritme k-means melalui proses Knowledge Discovery in Databases (KDD). Tahap awal yang dilakukan melakukan seleksi data dengan menghapus kolom "id", selanjutnya melakukan pengecekan nilai null. Transformation data melakukan normalisasi data dengan menggunakan min-max normalization dengan rentang nilai 0 sampai 1. Hasil data mining dengan K-Means menunjukkan bahwa cluster yang digunakan berjumlah 2 dengan menggunakan metode Elbow. Cluster 0 merupakan kelompok dataset makanan dengan tingkat protein dan kalori yang tinggi sedangkan cluster 1 merupakan kelompok dataset makanan dengan tingkat protein dan kalori yang rendah. Hasil evaluasi menggunakan silhouette score menunjukkan cluster yang terbaik adalah 2 dengan nilai rata-rata 0.61 yang merupakan kategori kriteria struktur baik. Model clustering menggunakan k-means kemudian diimplementasikan ke dalam aplikasi website sederhana menggunakan Flask sebagai web server dan TailwindCSS sebagai front-end. Aplikasi ini memungkinkan pengguna untuk menentukan menu makanan berdasarkan kalori harian. Penelitian ini menunjukkan bahwa algoritme k-means berhasil untuk menentukan menu makanan berdasarkan kalori harian.

**Kata kunci :** Pola Hidup Sehat, Harris-Benedict, Clustering, K-Means, Flask

### 1. PENDAHULUAN

Masyarakat di Indonesia masih sangat jauh dari kata pola hidup sehat. Menurut Legatum Institute pada 2023 aspek kesehatan Indonesia berada di peringkat 87 dari 167 negara di dunia. Salah satu cara yang dapat dilakukan untuk menerapkan pola hidup sehat adalah dengan cara memakan makanan bergizi. Memilih makanan tidak asal sembarang memilihnya harus sesuai dengan kebutuhan kalori setiap individu. Apabila kekurangan kalori yang dibutuhkan akan menyebabkan penurunan berat badan dan juga sebaliknya apabila kelebihan kalori yang dibutuhkan maka akan menyebabkan kelebihan berat badan atau obesitas[1].

Untuk mengetahui berapa kalori yang dibutuhkan dalam sehari dapat menggunakan persamaan Harris-Benedict dengan hasil Basal Metabolic Rate yang mempertimbangkan beberapa faktor dari setiap individu. Setelah itu menentukan kebutuhan kalori dalam satu hari sehingga akan mendapatkan Total Daily Energy Expenditure (TDEE)[2].

Status gizi merupakan salah satu poin yang penting dalam mencapai kesehatan yang optimal[3].

Status gizi remaja di Indonesia berdasarkan data penelitian Riskesdas yang dilakukan pada tahun 2018 untuk lebih jelasnya dapat dilihat pada Tabel 1.

Tabel 1. Prevalensi status gizi

Usia	Kategori Postur Badan				
	Sangat Kurus	Kurus	Normal	Gendut	Obesitas
13 – 15	1,9%	6,8%	75,3%	11,2%	4,8%
16 – 18	1,4%	6,7%	78,3%	9,5%	4%
>18	-	9,3%	55,3%	13,6%	21,7%

Rekomendasi makanan bergizi dapat dijadikan patokan untuk penyesuaian kalori yang dibutuhkan oleh setiap individu terpenuhi dengan memperhatikan komposisi makanan tersebut. Dalam proses menentukan makanan bergizi dapat menggunakan data[2].

Data tersebut mempunyai informasi berupa lauk, sayuran, makanan pokok, dan buah. Informasi data yang tersedia akan dimanfaatkan untuk membuat fitur untuk menentukan menu makanan bergizi. Salah satu cara yang dapat digunakan untuk mendapatkan menu makanan bergizi dengan bisa dengan menggunakan teknik clustering. Clustering adalah metode untuk mengelompokkan data berdasarkan kesamaan dalam satu kelompok. Penggunaan pengelompokan sangat relevan dalam sistem rekomendasi ini, karena dapat membantu dalam menentukan dengan lebih mudah menu makanan yang cocok dengan kebutuhan kalori individu[4].

Dalam teknik *clustering* terdapat algoritme *K-Means*, metode atau teknik yang digunakan untuk mengelompokkan data non-hierarkis (partisi), yang mana memungkinkan menjadi dua kelompok atau lebih. Teknik tersebut membagi data menjadi satu kelompok, yang mana data yang memiliki karakteristik mirip dikelompokkan bersama, sedangkan data yang memiliki karakteristik yang berbeda dikelompokkan secara terpisah[5].

Pada penelitian ini akan melakukan penerapan algoritme *K-Means* dalam menentukan menu makanan berdasarkan kebutuhan kalori harian serta penerapan *deployment* hasil model. Penelitian ini diharapkan berhasil menerapkan metode *K-Means* dalam menentukan menu makanan berdasarkan kebutuhan kalori dan *deployment* terhadap hasil model.

## 2. TINJAUAN PUSTAKA

### 2.1. Kebutuhan Kalori

Energi yang diperoleh dari konsumsi makanan dan minuman merupakan sumber peningkatan kalori dalam tubuh. Dengan kata lain, energi atau kalori tersebut diperlukan oleh tubuh agar dapat menjalankan aktivitas dengan optimal[6]. Setiap individu mempunyai kebutuhan kalori yang berbeda-beda mengingat adanya perbedaan umur, jenis kelamin, tinggi badan, berat badan, dan tingkat aktivitas yang berbeda setiap individu. Dalam menentukan

kebutuhan kalori setiap individu dapat menggunakan kalkulator kalori seperti Basal Metabolic Rate (BMR) ataupun Harris-Benedict[7].

### 2.2. Harris-Benedict

Jumlah kalori yang dibutuhkan oleh tubuh selama satu hari dalam tubuh dapat ditentukan melalui perhitungan Harris-Benedict. Perhitungan Harris-Benedict dapat dipengaruhi oleh usia, tinggi badan, berat badan, dan jenis kelamin, dan aktivitas fisik[2]. Berikut persamaan untuk BMR

$$BMR(L) = 66 + (13,7 \times BB) + (5 \times TB) - (6,8 \times U) \tag{1}$$

$$BMR(P) = 665 + (9,6 \times BB) + (1,8 \times TB) - (4,7 \times U) \tag{2}$$

Keterangan:

BMR(L) : Basal Metabolic Rate (Laki-Laki)

BMR(P) : Basal Metabolic Rate (Perempuan)

BB : Berat Badan

TB : Tinggi Badan

U : Usia

Untuk menggunakan rumus Harris-Benedict harus mengetahui perhitungan faktor aktivitas. Faktor aktivitas fisik memiliki perhitungan yang berbeda-beda dapat dilihat pada Tabel 2.

Tabel 2. Kategori aktivitas fisik

No	Aktivitas	Laki-laki	Perempuan	Keterangan
1	Sangat Ringan	1,3	1,3	100% Waktu digunakan hanya untuk duduk atau berdiri.
2	Ringan	1,56	1,55	75% Waktu digunakan hanya untuk duduk atau berdiri dan 25% Waktu untuk berdiri atau bergerak.
3	Sedang	1,76	1,7	60% Waktu digunakan hanya untuk duduk atau berdiri dan 40% Waktu untuk aktivitas tertentu.
4	Berat	2,1	2	40% Waktu digunakan hanya untuk duduk atau berdiri dan 60% Waktu untuk aktivitas tertentu

Rumus *Harris-Benedict* mempunyai persamaan yang dapat dilihat pada persamaan.

$$TDEE = BMR \times AF \tag{3}$$

Keterangan:

BMR = Basal Metabolic Rate

AF = Aktivitas Fisik

TDEE = Total Daily Energy Expenditure (Kebutuhan kalori dalam sehari)

### 2.3. Data Mining

Data mining merupakan tahapan untuk memperoleh informasi yang bermanfaat dari basis data yang memiliki kapasitas yang besar dan harus dilakukan pengolahan agar menghasilkan informasi baru bagi penggunaannya dan bisa membantu untuk melakukan pengambilan keputusan. Dalam penelitian yang dilakukan oleh Wu et al. (2021) mengungkapkan proses *data mining* mempunyai 6 langkah yaitu sebagai berikut yaitu pemilihan data, pengolahan dan

integrasi data, pembersihan dan transformasi data, data mining, evaluasi, dan penilaian hasil[8].

### 2.4. Clustering

Clustering adalah teknik statistik multidimensi yang bertujuan untuk mengelompokkan individu-individu yang mirip ke dalam kelas-kelas yang seragam berdasarkan nilai-nilai yang diamati dalam sejumlah variabel. Kelas-kelas yang dihasilkan dapat diorganisir sesuai dengan berbagai struktur[9]

### 2.5. K-Means

*K-Means* adalah metode atau teknik pengelompokan data yang membagi data-data tersebut ke dalam beberapa kelompok yang berbeda. Secara iteratif, *K-Means* berhasil dalam meminimalkan rata-rata jarak setiap data ke pusat *clusternya* atau centroid. Penggunaan algoritme *K-Means* dalam proses pengelompokan atau *clustering* tergantung pada data yang tersedia dan hasil yang ingin dicapai[10].

**2.6. Metode Elbow**

Metode *Elbow* merupakan metode yang dapat mengoptimalkan pemilihan *centroid* atau titik *cluster* dalam proses *data mining*. Tujuan dilakukannya metode ini untuk memilih nilai *cluster* (k) yang kecil dengan masih memiliki nilai *witness* yang rendah. Untuk mengetahui titik yang optimal dalam metode ini akan membentuk siku dalam grafik[11]. Adapun tahapan dalam menggunakan algoritme *K-Means* sebagai berikut:

1. Menentukan berapa banyak jumlah *cluster* yang akan digunakan atau nilai k.
2. Menentukan pusat *cluster* atau *centroid* secara *random*.
3. Mengukur jarak antara setiap titik data dan setiap pusat. Hal tersebut dapat menggunakan persamaan *Euclidean Distance*.

$$d(a,b) = \sqrt{\sum_{k=1}^n (x_{ak} - x_{bk})^2} \tag{4}$$

Keterangan:

d(a,b) : Distance (jarak) antara objek a dan b

n : Jumlah atribut

$x_{ak}$  : Nilai pusat dari objek a pada dimensi k

$x_{bk}$  : Nilai pusat dari objek b pada dimensi k

**2.7. Min-Max Normalization**

Metode *Min-Max Normalization* merupakan metode normalisasi yang melakukan transformasi linear dengan menggunakan nilai yang terkecil atau nilai minimum dan nilai terbesar atau nilai maksimum. Metode ini berguna untuk menghasilkan nilai-nilai yang seimbang dengan rentang nilai 0 sampai 1[12].

**2.8. Silhouette Coefficient**

*Silhouette Coefficient* merupakan sebuah metode evaluasi yang digunakan untuk memvalidasi *cluster* yang terbentuk dengan menggabungkan dua buah metode yaitu antara metode kohesi dan metode separasi. Metode ini bisa dibilang bahwa metode yang paling banyak digunakan untuk melakukan validasi *cluster* [13]. Adapun kriteria dalam pengukuran *silhouette coefficient* yang dapat dilihat pada Tabel 3.

Tabel 3. Kriteria struktur

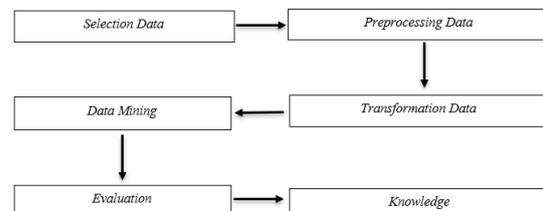
Nilai <i>Silhouette Coefficient</i>	Kriteria
0.71 – 1.00	Struktur kuat
0.51 – 0.70	Struktur baik
0.26 – 0.50	Struktur lemah
≤ 0.25	Struktur buruk

**2.9. Flask**

*Flask* adalah sebuah framework yang digunakan untuk mengembangkan sistem aplikasi berbasis web. Kerangka kerja ini ditulis dalam bahasa pemrograman *Python*. Ketika menggunakan *Flask*, kemudahan datang dari fakta bahwa ia menggunakan *Python*, sehingga mempermudah pengembangan aplikasi dalam satu bahasa. Ini memungkinkan integrasi model jaringan saraf tiruan ke dalam kerangka kerja *Flask* dengan lebih mudah [14]

**3. METODE PENELITIAN**

Metodologi yang dipilih pada penelitian ini adalah menggunakan metodologi *Knowledge Discovery in Database* (KDD) yang terdiri dari enam tahapan yaitu *Selection Data* (seleksi data), *Preprocessing Data* (pemrosesan data), *Transformation Data* (transformasi data), *Data Mining*, *Evaluation Data* (evaluasi), dan *Knowledge* (pengetahuan).



Gambar 1. Tahapan Metode Penelitian

**3.1. Selection Data**

Pada tahap pertama yaitu *selection data*, dilakukan pengumpulan data merupakan data publik yang diambil dari sumber website [www.panganku.org](http://www.panganku.org) yang dimiliki oleh Kementerian Kesehatan Republik Indonesia (Kemenkes) pada tahun 2017 yang mana merupakan data terakhir yang diperbaharui dan juga perkembangan dari tahun 2009.

**3.2. Preprocessing Data**

*Preprocessing data* atau pemrosesan data. Pada tahap ini dilakukan pembersihan data atau yang disebut dengan data *cleaning* yang digunakan untuk menghilangkan duplikasi data agar tidak terjadi *missing value* atau data yang tidak lengkap dengan nilai pada atribut *null* (data kosong). Pengecekan data dilakukan karena data tersebut dapat menyebabkan analisis data menjadi tidak akurat dan dapat membuat informasi menjadi melebar atau tidak relevan.

**3.3. Transformation Data**

*Transformation data* yang dilakukan perubahan data yang kompleks menjadi data yang mudah diolah seperti mengurutkan data, normalisasi data, atau dengan simbolisasi data. Semisal ada data yang masih dalam format nominal maka harus diubah menjadi format kategorikal numerik. Hal ini karena algoritme *K-Means* mengelompokan data berdasarkan jarak, maka dari itu data yang digunakan lebih cocok data dengan format numerik atau data yang berisi dengan angka.

**3.4. Data Mining**

*Data mining* yang merupakan tahap utama dalam pengelolaan data dilakukan mining data dengan teknik klasterisasi menggunakan metode *K-means* serta optimalisasi penentuan cluster menggunakan metode *elbow*.

**3.5. Evaluation Data**

*Evaluation data* yaitu melakukan proses pengujian hasil dari tahap data mining yang digunakan. Pada penelitian ini digunakan sebuah metode evaluasi untuk menguji hasil dengan metode *Silhouette Coefficient*.

**3.6. Knowledge**

*Knowledge* yaitu tahap terbentuknya cluster sebagai informasi atau pengetahuan dari hasil dalam penelitian ini. Hasil tersebut nantinya akan dideploy dalam bentuk *website*, sehingga pengguna dapat menggunakan dan mengimplementasikan hasil yang didapatkan.

**4. HASIL DAN PEMBAHASAN**

Penelitian yang dilakukan menggunakan metodologi *Knowledge Discovery in Databases* (KDD) dalam melakukan analisis algoritma *clustering*, yaitu algoritme *K-Means*. Pada penelitian ini dibantu menggunakan bahasa pemrograman Python dan beberapa *Librariesnya* seperti Pandas, Sklearn. *Libraries* ini membantu dan memudahkan dalam pembuatan dataset, pengolahan dataset, permodelan, dan menampilkan grafik. Jadi sebagian besar tabel dan grafik yang dihasilkan pada bab ini dihasilkan dari *Libraries* tersebut.

**4.1. Selection Data**

Pada tahap *selection data*, melakukan pemilihan dataset. Dataset yang akan digunakan sebanyak 359 entri data yang pada awalnya terdapat 5 atribut seperti pada Tabel 4 di bawah.

Tabel 4. Dataset sebelum diseleksi

<b>Id</b>	<b>Nama makanan</b>	<b>Kalori</b>	<b>Protein</b>	<b>jenis</b>
1	Nasi	180	3.0	Pokok
2	Beef Yakiniku	132	9.8	Lauk
3	Bayam Rebus	23	1.2	Sayuran
4	Alpukat	85	0.9	Buah

Setelah melakukan seleksi data, hanya 4 atribut yang digunakan dengan menghilangkan atribut “id” dikarenakan tidak relevan dengan penelitian ini. Data yang telah diseleksi dapat dilihat seperti Tabel 5 di bawah.

Tabel 5. Dataset sesudah diseleksi

<b>Nama makanan</b>	<b>Kalori</b>	<b>Protein</b>	<b>jenis</b>
Nasi	180	3.0	Pokok
Beef Yakiniku	132	9.8	Lauk
Bayam Rebus	23	1.2	Sayuran
Alpukat	85	0.9	Buah

**4.2. Preprocessing Data**

Pada tahap *preprocessing* data, melakukan pengecekan nilai null di setiap kolom atau *attributes*. Nilai null dapat dilihat jika kurangnya jumlah entri data dari jumlah maksimal entri data yang mana jumlah maksimal entri data tersebut adalah 359.

Jumlah nilai *null* atau nilai kosong di setiap kolom dapat dilihat pada Tabel 6.

Tabel 6. Jumlah nilai *null*

<b>Atribut</b>	<b>Null_Counts</b>
Nama Makanan	0
Kalori	0
Protein	0
Jenis	0

Terlihat bahwa tidak ada nilai yang kosong, hal tersebut dapat dilihat pada kolom “*null\_counts*” hasilnya adalah 0.

**4.3. Transformation Data**

Pada tahap *transformation data*, melakukan pengubahan dataset dengan menggunakan teknik *min-max normalization*. Metode ini melakukan transformasi linear terhadap data asli, yang mana pada proses normalisasi data dilakukan untuk mendapatkan nilai dengan rentang antara 0 dan 1. Untuk mendapatkan nilai hasil normalisasi dapat menggunakan persamaan berikut:

$$x' = \frac{x - \min(x)}{\max(x) - \min(x)} \tag{5}$$

Keterangan:

$x'$  : Nilai yang sudah dinormalisasi

$x$  : Nilai asli

$\min(x)$  : Nilai minimum dataset

$\max(x)$  : Nilai maksimum dataset

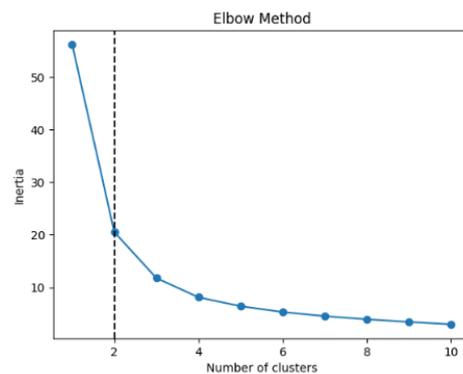
Hasil sampel dari *min-max normalization* ini dapat dilihat pada Tabel 7 berikut.

Tabel 7. Hasil *min-max normalization*

<b>Nama makanan</b>	<b>Kalori</b>	<b>Protein</b>	<b>jenis</b>
Nasi	0.364606	0.120332	Pokok
Nasi Tim	0.236674	0.095436	Pokok
Nasi Merah	0.298507	0.112033	Pokok
Ketupat Ketan	0.432836	0.161826	Pokok

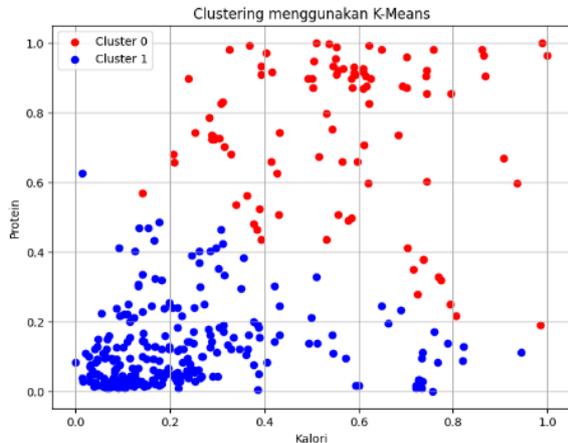
**4.4. Data Mining**

Pada proses *data mining* akan menggunakan teknik *clustering* dengan menggunakan algoritme *K-Means*. Tahap pertama dalam algoritme *K-Means* adalah menentukan jumlah *cluster* menggunakan metode *Elbow*.



Gambar 2. Metode *Elbow*

Pada gambar di atas menunjukkan bahwa *cluster 2* merupakan *cluster* yang paling optimal dengan membentuk sudut siku dan adanya penurunan di setiap nilai *cluster* dan nilai yang stabil atau turun secara perlahan-lahan berada di *cluster 2*.



Gambar 3. Visualisasi *Cluster K-Means*

Gambar di atas menunjukkan hasil visualisasi *cluster* yang berjumlah 2. Hasil analisis, *cluster 0* merupakan kelompok dataset makanan dengan tingkat protein dan kalori tinggi dengan jumlah dataset sebanyak 95 dataset. Sedangkan *cluster 1* merupakan kelompok dataset makanan dengan tingkat protein dan kalori rendah dengan jumlah dataset sebanyak 264 dataset. Adapun rincian karakteristik setiap *cluster* yang dapat dilihat pada Tabel 8 dan Tabel 9 di bawah.

Tabel 8. Rincian karakteristik *cluster 0*

<b>Cluster 0</b>	<b>Kalori</b>	<b>Protein</b>
Min	0.141365	0.190871
Max	1	1
Mean	0.557116	0.743918

*Cluster 0* merupakan kelompok *cluster* dengan tingkat kalori dan protein yang tinggi. *Cluster 0* terdiri dari 95 dataset. Pada *cluster 0* makanan dengan kalori terendah = 0.141365, kalori tertinggi = 1, dan rata-ratanya = 0.557116. Sedangkan protein terendah = 0.190871, protein tertinggi = 1, dan rata-ratanya = 0.743918.

Tabel 9. Rincian karakteristik *cluster 1*

<b>Cluster 1</b>	<b>Kalori</b>	<b>Protein</b>
Min	0	0
Max	0.944563	0.626556
Mean	0.219531	0.114579

*Cluster 1* merupakan kelompok *cluster* dengan tingkat kalori dan protein yang rendah. *Cluster 1* terdiri dari 264 dataset. Pada *cluster 1* makanan dengan kalori terendah = 0, kalori tertinggi = 0.944563, dan rata-ratanya = 0.219531. Sedangkan protein terendah = 0, protein tertinggi = 0.626556, dan rata-ratanya = 0.114579.

Dengan demikian, langkah selanjutnya adalah mengaplikasikan algoritme *K-Means* untuk melakukan pengelompokan dataset tersebut berdasarkan jumlah *cluster* yang telah ditentukan. Sampel hasil clustering menggunakan algoritme *K-Means*. Berikut hasil sampel setiap *cluster* dapat dilihat pada Tabel 10 dan Tabel 11.

Tabel 10. Sampel hasil *cluster 0*

<b>Nama makanan</b>	<b>Kalori</b>	<b>Protein</b>	<b>Jenis</b>	<b>cluter</b>
Mi Kering	0.793177	0.248963	Pokok	0
Bihun Goreng Instan	0.793177	0.248963	Pokok	0
Misoa	0.716417	0.348547	Pokok	0
Jagung Tepung Putih	0.737739	0.377593	Pokok	0

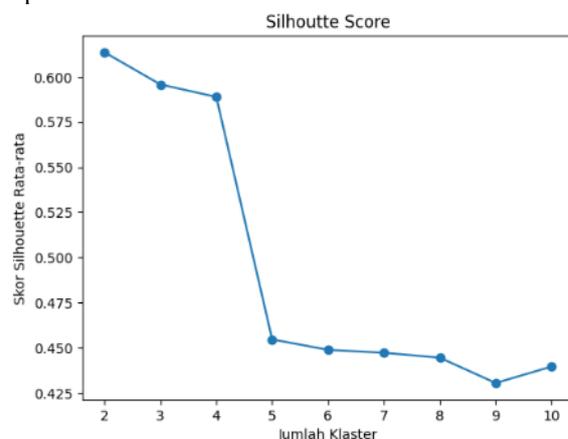
Tabel 11. Sampel hasil *cluster 1*

<b>Nama makanan</b>	<b>Kalori</b>	<b>Protein</b>	<b>Jenis</b>	<b>cluter</b>
Nasi	0.364605	0.120331	Pokok	1
Nasi Tim	0.236673	0.095435	Pokok	1
Nasi Merah	0.298507	0.112033	Pokok	1
Ketupat Ketan	0.432835	0.161825	Pokok	1

Pada Tabel 10 dan Tabel 11 merupakan sampel hasil dari setiap *cluster 0* dan *cluster 1*. Sampel yang digunakan pada *cluster 0* terdiri 4 dataset dari 95 dataset. Sedangkan sampe yang digunakan pada *cluster 1* terdiri 4 dataset dari 264 dataset.

#### 4.5. Evaluation

Pada tahap evaluasi berguna untuk mengukur kualitas hasil dari *clustering*. Metode dalam evaluasi hasil *cluster* yaitu dengan menggunakan metode *Silhouette Coefficient*. Semakin tinggi nilai *Silhouette Coefficient*, semakin baik kualitas *clustering* yang diperoleh.



Gambar 4. Metode *Silhouette Score*

Gambar di atas memperlihatkan nilai rata-rata *silhouette score* setiap *cluster*. Terlihat di *cluster 2*

merupakan nilai tertinggi dengan nilai *silhouette score* 0. 61 yang merupakan kategori kriteria struktur baik.

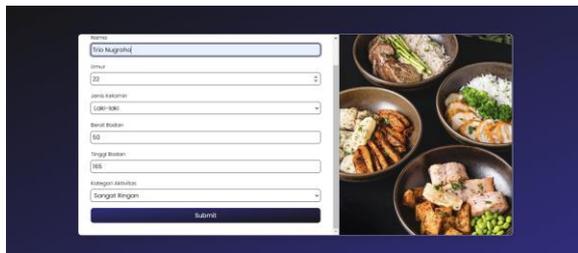
4.6. Knowledge

Setelah melakukan seluruh tahapan, selanjutnya adalah masuk ke tahap terakhir yaitu *knowledge*. Hal ini berguna untuk informasi atau pengetahuan dari hasil penelitian ini yang mana akan dideploy dalam bentuk *website*. Langkah pertama yang dilakukan adalah dengan memasukan input dari faktor-faktor yang mempengaruhi kebutuhan kalori harian. Data yang akan diinputkan sebagai contoh berjenis kelamin laki-laki, berumur 22 tahun, berat badan 50 kilogram, dan tinggi badan 165 centimeter. Maka data-data tersebut dapat diinputkan ke dalam persamaan *Harris-Benedict*.

$$BMR(L) = 66 + (13,7 \times 50) + (5 \times 165) - (6,8 \times 22) = 1426.4 \tag{5}$$

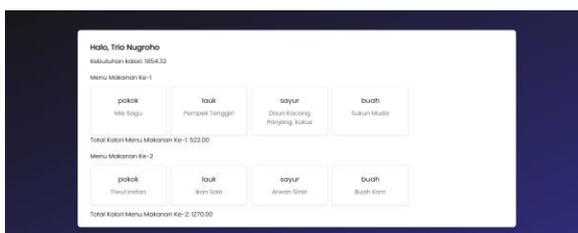
$$TDEE = 1426.4 \times 1.3 \tag{6}$$

Setelah memasukan input data, maka kebutuhan kalori harian dengan menggunakan persamaan *Harris-Benedict* mendapatkan 1854.32 kalori. Dengan jumlah kalori harian yang sudah didapatkan maka akan disesuaikan dengan menu makanan berdasarkan *cluster* dan jumlah kalorinya. Selanjutnya, model yang sudah dibuat akan dideploy menjadi halaman *website* sederhana dengan menggunakan *Flask* sebagai *web server* dan *TailwindCSS* sebagai *front-end*. Halaman pertama akan memunculkan *input form* yang digunakan untuk menghitung jumlah kalori yang dibutuhkan dalam satu hari.



Gambar 5. Halaman Input

Setelah melakukan *submit* akan muncul halaman kedua yang merupakan hasil dari *model* yang sudah dibuat menggunakan Algoritme *K-Means* dan akan memunculkan menu-menu makanan dalam satu hari. Untuk lebih jelasnya dapat dilihat pada Gambar 6.



Gambar 6. Halaman Hasil

Gambar di atas merupakan hasil dari permodelan yang sudah *dideploy*. Menu makanan ke-1 merupakan hasil dari *cluster* 1 dan menu makanan ke-2 merupakan hasil dari *cluster* 0.

5. KESIMPULAN DAN SARAN

Penelitian ini berhasil menentukan menu makanan berdasarkan kebutuhan kalori harian, serta penelitian ini menggunakan proses metodologi KDD (*Knowledge Discovery in Database*) untuk menghasilkan model dengan penerapan algoritme *K-Means*. Langkah awal, melakukan seleksi data dengan menghapus kolom “id”. Pada tahap *preprocessing data* tidak adanya nilai yang kosong atau nilai *null*. Setelah itu, merupakan normalisasi dengan *min-max normalization*. *Data mining* menggunakan algoritme *K-Means* dengan menentukan jumlah *cluster* menggunakan metode *Elbow*, metode ini menghasilkan jumlah *cluster* 2.

Hasil analisis, menunjukkan *cluster* 0 merupakan bahwa kelompok dataset makanan dengan tingkat protein dan kalori tinggi dan *cluster* 1 merupakan bahwa kelompok dataset makanan dengan tingkat protein dan kalori rendah. Model yang dihasilkan menggunakan algoritme *K-Means* berhasil *dideployment* ke dalam bentuk *website* dengan *Flask* sebagai *web server* dan *TailwindCSS* sebagai *front-end*. *Website* tersebut menunjukkan menu makanan dalam satu hari berdasarkan kebutuhan kalori harian. Meskipun algoritme *K-Means* berhasil dalam menentukan menu makanan, disarankan untuk menggunakan dataset yang lebih besar dan algoritme yang berbeda. Selain itu, melakukan evaluasi yang lebih inovatif, sehingga dapat menghasilkan model dengan performa yang lebih optimal.

DAFTAR PUSTAKA

[1] B. Sembiring and H. Kristian Siburian, “Ahp Penerapan Algoritma Genetika Pada Proses Menurunkan Berat Badan Menggunakan Aplikasi Diet Sehat Berbasis Mobile,” *Jurnal Informasi dan Teknologi Ilmiah*, vol. 8, no. 2, 2021.

[2] S. I. et al Indra, “Rekomendasi Menu Makanan Bergizi Dengan Decision Support System Menggunakan Algoritma Genetika,” *Jurnal Ilmiah Komputasi*, vol. 22, no. 1, Mar. 2023, doi: 10.32409/jikstik.22.1.3316.

[3] F. S. Maedy, T. A. E. Permatasari, and S. Sugiatmi, “Hubungan Status Gizi dan Stres terhadap Siklus Menstruasi Remaja Putri di Indonesia,” *Muhammadiyah Journal of Nutrition and Food Science (MJNF)*, vol. 3, no. 1, p. 1, Jul. 2022, doi: 10.24853/mjnf.3.1.1-10.

[4] H. Mutiasari, T. W. Purboyo, and R. A. Nugrahaeni, “SISTEM REKOMENDASI FILM MENGGUNAKAN METODE K-MEANS CLUSTERING (MOVIE RECOMMENDATION SYSTEM USING K-MEANS CLUSTERING METHOD).”

- [5] I. Nuryani and D. Darwis, "ANALISIS CLUSTERING PADA PENGGUNA BRAND HP MENGGUNAKAN METODE K-MEANS." [6] D. Megah Sari, S. Sulfayanti, and N. Arifin, "Rekomendasi Makanan Pendamping Asi Berdasarkan Kebutuhan Kalori Menggunakan Algoritma Naive Bayes," *Jurnal Ilmiah Sistem Informasi dan Teknik Informatika (JISTI)*, vol. 5, no. 2, pp. 131–141, Oct. 2022, doi: 10.57093/jisti.v5i2.141.
- [7] N. Cholis Anggoro and M. Akbar, "Chatbot Pemilihan Makanan dan Minuman Berdasarkan Kalori menggunakan Natural Language Processing," 2023.
- [8] W. T. Wu *et al.*, "Data mining in clinical big data: the frequently used databases, steps, and methodological models," *Military Medical Research*, vol. 8, no. 1. BioMed Central Ltd, Dec. 01, 2021, doi: 10.1186/s40779-021-00338-z.
- [9] R. Anjariansyah and 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, p. 597, Jan. 2022, doi: 10.30865/mib.v6i1.3522.
- [10] T. G. Manik, W. I. Rahayu, R. Nuraini, and S. Fathonah, "PERBANDINGAN METODE FUZZY C-CMEANS DAN K-MEANS CLUSTERING PADA DATA PENGGUNAAN OBAT DI R.S NATIONAL HOSPITAL SURABAYA," 2023.
- [11] N. T. Hartanti, "Metode Elbow dan K-Means Guna Mengukur Kesiapan Siswa SMK Dalam Ujian Nasional," *Jurnal Nasional Teknologi dan Sistem Informasi*, vol. 6, no. 2, pp. 82–89, Aug. 2020, doi: 10.25077/teknosi.v6i2.2020.82-89.
- [12] N. Puspitasari, G. Lempas, H. Hamdani, H. Haviuddin, and A. Septiarini, "Perbandingan Algoritma K-Means dan Algoritma K-Medoids Pada Kasus Covid-19 di Indonesia," *Building of Informatics, Technology and Science (BITS)*, vol. 4, no. 4, Mar. 2023, doi: 10.47065/bits.v4i4.2994.
- [13] D. I. Yunistya, R. Goejantoro, F. Deny, and T. Amijaya, "The Application Of K-Harmonic Means Method In District/City Grouping (Case Study: Poverty in Kalimantan Island in 2020) Penerapan Metode K-Harmonic Means Dalam Pengelompokan Kabupaten/Kota (Studi Kasus: Kemiskinan di Pulau Kalimantan Tahun 2020)," vol. 19, no. 1, pp. 51–64, 2022, doi: 10.20956/j.v19i1.21116.
- [14] R. N. M. Kholilul, E. Prakarsa Mandyartha, and A. M. Rizki, "Seminar Nasional Informatika Bela Negara (SANTIKA) Rancang Bangun Sistem Deteksi Huruf Rusia Berbasis Web Flask," *Seminar Nasional Informatika Bela Negara (SANTIKA)*, vol. Vol. 2, 2021.