



Department of Digital Business

**Journal of Artificial Intelligence and Digital Business (RIGGS)**

Homepage: <https://journal.ilmudata.co.id/index.php/RIGGS>

Vol. 4 No. 3 (2025) pp: 899-908

P-ISSN: 2963-9298, e-ISSN: 2963-914X

---

## Implementasi Algoritma *K-Nearest Neighbors* untuk Rekomendasi Diet Tinggi Serat dalam Mencegah Penyakit Jantung

Maria Anjelina Domu Tukan<sup>1\*</sup>, Alfian Nara Weking<sup>2</sup>, Dominikus Boli Watomakin<sup>3</sup>  
<sup>1,2,3</sup> Program Studi Teknik Informatika, Institut Keguruan dan Teknologi Larantuka, Indonesia  
[manjelynatukan9@gmail.com](mailto:manjelynatukan9@gmail.com)

### Abstrak

*Penyakit jantung merupakan salah satu ancaman kesehatan utama di era modern dan menjadi penyebab kematian tertinggi di seluruh dunia. Jumlah kematian akibat penyakit jantung mencapai 17,9 juta jiwa secara global. Di Indonesia, angka kematian akibat penyakit jantung terus meningkat dari tahun ke tahun. Tercatat sebanyak 15 dari setiap 1.000 orang, atau sekitar 2.784.064 jiwa di Indonesia, menderita penyakit jantung. Untuk mencegah tingginya angka kematian akibat penyakit jantung, diperlukan upaya pencegahan dini guna mengurangi risiko terjadinya penyakit tersebut. Tujuan dari penelitian ini adalah untuk mengevaluasi efektivitas algoritma K-Nearest Neighbor (K-NN) dalam menganalisis data pola makan tinggi serat guna mencegah penyakit jantung. Penelitian ini juga bertujuan untuk mengidentifikasi pola konsumsi serat yang optimal bagi kesehatan jantung, memberikan rekomendasi gaya hidup sehat berbasis data, serta mengembangkan model prediksi yang akurat dan efisien. Algoritma K-Nearest Neighbor (K-NN) merupakan pendekatan yang sederhana dan mudah diimplementasikan. Metode ini efektif digunakan pada dataset yang memiliki banyak atribut serta mampu menangani permasalahan klasifikasi atau regresi dengan lebih dari dua kelas tanpa memerlukan modifikasi besar. Hasil penelitian menunjukkan bahwa algoritma K-NN sangat efektif dalam merekomendasikan makanan, dengan prediksi model menghasilkan akurasi sebesar 93%.*

*Kata kunci: Penyakit Jantung, Diet Tinggi Serat, K-Nearest Neighbors*

### 1. Latar Belakang

Penyakit jantung merupakan salah satu ancaman kesehatan utama di era modern, dan telah menjadi penyebab kematian terbesar di seluruh dunia selama beberapa dekade terakhir. Menurut data Organisasi Kesehatan Dunia (WHO), penyakit kardiovaskular, termasuk penyakit jantung koroner dan stroke, bertanggung jawab atas sekitar 17,9 juta kematian setiap tahunnya, yang setara dengan 32% dari seluruh kematian global (Gopalan, 2022). Tingginya angka ini menunjukkan bahwa penyakit jantung bukan hanya masalah kesehatan individu, tetapi juga menjadi beban besar bagi sistem kesehatan publik, perekonomian, dan produktivitas masyarakat. Faktor-faktor seperti urbanisasi, perubahan gaya hidup, pola makan tidak sehat, kurangnya aktivitas fisik, stres, dan paparan polutan lingkungan semakin memperparah risiko kejadian penyakit jantung, terutama di negara berkembang yang mengalami transisi epidemiologis.

Di Indonesia, tren prevalensi penyakit jantung menunjukkan kecenderungan yang mengkhawatirkan. Data Kementerian Kesehatan RI mengungkapkan bahwa prevalensi penyakit jantung terus meningkat dari tahun ke tahun. Berdasarkan hasil Riset Kesehatan Dasar (Riskesdas), tercatat bahwa sekitar 15 dari setiap 1.000 penduduk Indonesia, atau setara dengan 2.784.064 orang, menderita penyakit jantung (Wiranugraha, 2022). Kondisi ini diperparah dengan rendahnya kesadaran masyarakat terhadap deteksi dini dan pencegahan, serta keterbatasan akses ke layanan kesehatan yang memadai di beberapa wilayah. Jika tidak ditangani dengan langkah-langkah preventif yang tepat, jumlah penderita penyakit jantung di Indonesia berpotensi terus meningkat dalam beberapa dekade mendatang, menambah beban biaya kesehatan nasional dan menurunkan kualitas hidup penduduk.

Upaya pencegahan menjadi salah satu strategi utama untuk mengurangi risiko penyakit jantung, mengingat biaya pengobatan yang relatif tinggi dan kompleksitas perawatan penyakit ini. Pencegahan dapat dilakukan melalui modifikasi gaya hidup, khususnya pada pola makan dan aktivitas fisik. Penerapan pola makan sehat, seperti mengonsumsi makanan bergizi seimbang dengan meningkatkan asupan serat, menjadi langkah penting yang direkomendasikan oleh berbagai penelitian. Diet tinggi serat diketahui memberikan beragam manfaat yang secara langsung berkontribusi pada kesehatan jantung, di antaranya menurunkan kadar kolesterol total dan kolesterol

LDL (low-density lipoprotein) yang merupakan salah satu pemicu utama aterosklerosis. Selain itu, serat membantu mengontrol kadar gula darah, mengatur tekanan darah, serta mendukung manajemen berat badan yang sehat (Tovar, 2020). Dengan demikian, asupan serat yang cukup bukan hanya bermanfaat bagi pencernaan, tetapi juga berperan signifikan dalam pencegahan penyakit kardiovaskular.

Meskipun manfaat serat terhadap kesehatan jantung sudah banyak diungkapkan dalam literatur, tantangan yang dihadapi adalah bagaimana mengidentifikasi pola konsumsi serat yang optimal berdasarkan karakteristik individu, seperti usia, jenis kelamin, indeks massa tubuh (IMT), riwayat penyakit, dan kebiasaan hidup. Dalam konteks inilah, analisis berbasis data (data-driven analysis) menjadi penting untuk menghasilkan rekomendasi yang lebih personal dan akurat. Teknologi analitik modern, khususnya dalam bidang machine learning, dapat dimanfaatkan untuk memproses data pola makan masyarakat dan memprediksi tingkat risiko penyakit jantung secara lebih efisien.

Penelitian ini bertujuan untuk mengevaluasi efektivitas algoritma K-Nearest Neighbors (K-NN) dalam menganalisis data diet tinggi serat guna mencegah penyakit jantung. Tujuan lainnya adalah mengidentifikasi pola konsumsi serat yang optimal bagi kesehatan jantung, memberikan rekomendasi berbasis data untuk gaya hidup sehat, serta mengembangkan model prediksi risiko yang akurat dan efisien. K-NN dipilih karena merupakan salah satu metode klasifikasi yang sederhana namun efektif, mampu menangani dataset dengan banyak atribut, dan dapat digunakan untuk masalah klasifikasi maupun regresi tanpa modifikasi yang kompleks. Prinsip kerja K-NN adalah mengklasifikasikan data baru berdasarkan kedekatannya dengan sejumlah tetangga terdekat (nearest neighbors) dalam ruang fitur, sehingga cocok digunakan untuk mengidentifikasi pola konsumsi makanan dan kaitannya dengan risiko penyakit jantung.

Namun demikian, metode K-NN tidak lepas dari kelemahan, seperti meningkatnya waktu komputasi ketika jumlah data latih dan data uji semakin besar, serta sensitivitas terhadap pemilihan nilai  $k$  dan skala fitur. Oleh karena itu, penerapan K-NN dalam penelitian ini juga mempertimbangkan teknik optimasi dan pra-pemrosesan data untuk meminimalkan kelemahan tersebut (de Castro, 2022). Dengan pendekatan ini, diharapkan hasil penelitian dapat memberikan kontribusi signifikan dalam pengembangan strategi pencegahan penyakit jantung berbasis diet sehat yang didukung analisis ilmiah dan teknologi prediksi mutakhir.

## 2. Metode Penelitian

Penelitian ini merupakan penelitian terapan yang memanfaatkan algoritma K-Nearest Neighbors (K-NN) untuk memberikan rekomendasi makanan dalam rangka pencegahan penyakit jantung melalui penerapan diet tinggi serat. Jenis penelitian yang digunakan adalah penelitian kuantitatif, yaitu metode yang berfokus pada pengumpulan, analisis, dan interpretasi data dalam bentuk angka untuk memperoleh temuan yang dapat diuji secara objektif (Mahsin, 2022). Secara umum, tahapan penelitian mencakup proses pengumpulan data, pra-pemrosesan data, penerapan algoritma K-NN, dan evaluasi hasil sistem.

Pengumpulan data dilakukan melalui dua sumber utama, yaitu data primer dan data sekunder. Data primer diperoleh melalui wawancara dengan ahli gizi atau profesional kesehatan yang memiliki kompetensi di bidang diet tinggi serat untuk pencegahan penyakit jantung. Informasi yang dikumpulkan meliputi kebutuhan gizi, preferensi makanan, dan rekomendasi jenis produk makanan yang sesuai untuk penderita atau individu berisiko penyakit jantung. Data ini digunakan untuk memvalidasi sistem rekomendasi yang dibangun dengan algoritma K-NN. Sementara itu, data sekunder diperoleh dari sumber terbuka (open source) yang dapat diakses secara internasional, yakni situs Kaggle (<https://www.kaggle.com>), yang menyediakan dataset berisi nilai nutrisi dari berbagai jenis makanan dan produk.

Tahap berikutnya adalah pra-pemrosesan data (data preprocessing) yang bertujuan untuk mempersiapkan data agar siap diolah oleh algoritma K-NN. Proses ini meliputi beberapa langkah penting (Tiu, 2022), yaitu: (1) menghilangkan satuan pada atribut nutrisi agar seragam dan mudah diolah, (2) membersihkan data yang tidak relevan, (3) menangani data yang hilang (missing value) dengan menggunakan perhitungan nilai rata-rata (mean), (4) menambahkan kategori yang diperlukan untuk klasifikasi, (5) melakukan normalisasi data dengan metode min-max normalization untuk menyetarakan skala nilai setiap atribut, dan (6) membagi dataset menjadi data latih (training data) dan data uji (testing data).

Penerapan algoritma K-Nearest Neighbors dilakukan pada tahap berikutnya. K-NN adalah metode klasifikasi yang bekerja dengan cara menghitung jarak terdekat antara data uji dengan sejumlah tetangga terdekat (nearest neighbors) pada data latih. Dalam penelitian ini, nilai  $k$  dipilih sedemikian rupa untuk meminimalkan kesalahan klasifikasi dan mengurangi pengaruh noise pada data. Semakin besar nilai  $k$ , semakin besar pula kemampuan

algoritma dalam mengurangi efek noise, namun perlu dipertimbangkan agar tidak mengurangi sensitivitas klasifikasi terhadap pola yang ada.

Tahap terakhir adalah evaluasi hasil sistem rekomendasi yang dibangun. Evaluasi dilakukan dengan membandingkan hasil klasifikasi yang dihasilkan sistem dengan data sebenarnya menggunakan confusion matrix. Matriks ini mencakup empat komponen utama, yaitu true positive (TP), true negative (TN), false positive (FP), dan false negative (FN). Dari hasil confusion matrix, dihitung beberapa metrik kinerja algoritma, antara lain accuracy (akurasi), precision, recall, dan F1-score (Naidu, 2023). Evaluasi ini bertujuan untuk memastikan bahwa model yang dihasilkan memiliki tingkat keandalan dan ketepatan yang memadai sehingga dapat digunakan sebagai sistem rekomendasi diet tinggi serat yang efektif dalam pencegahan penyakit jantung.

### 3. Hasil dan Diskusi

#### Penggumpulan Data

Penelitian ini dilaksanakan pada 14 Mei sampai 16 Juli 2025 di Puskesmas Nagi Larantuka. Pengumpulan data menggunakan dua teknik, yaitu: data primer, yang diperoleh melalui wawancara dengan ahli gizi/kesehatan. Berdasarkan hasil yang diperoleh, nilai nutrisi yang relevan dalam mencegah penyakit jantung seperti kalori, total lemak, lemak jenuh, kolesterol, sodium, folat, vitamin b6, vitamin b12, vitamin c, vitamin d, vitamin e, vitamin k, protein, karbohidrat, serat, lemak, gula, kalsium, kalium dan magnesium dan data sekunder di peroleh dari website resmi internasional google kaggle (<https://www.kaggle.com>). Dataset yang di kumpulkan mencakup informasi makanan beserta nutrisi seperti (kalori, lemak, protein, serat, garam, gula dan lain-lain). Berikut dibawah ini adalah contoh data:

Tabel 1. Data Makanan (Validasi)

No	nama	Porsi/ ukuran	kalori	total lemak	lemak jenuh	kolestrol	sodium	....
1	Kacang merah direbus tanpa garam	100 g	127 kkal	0.5g	0.1g	0	2.00 mg	....
2	Kentang direbus dengan kulit tanpa garam,	100 g	78	0.1g		0	14.00 mg	....
3	Pir, mentah	100 g	57	0.1g		0	1.00 mg	....

No	protein	Karbohidrat	Serat	Gula	Lemak	Label Rekomendasi
1	8.67 g	22.80 g	7.4 g	0.32 g	0.50 g	Direkoemndasikan
2	2.86 g	17.21 g	3.3 g	0	0.10 g	Direkoemndasikan
3	0.36 g	15.23 g	3.1 g	9.75 g	0.14 g	Direkomendasiakn

Tabel 2. Dataset Makanan (Kaggle)

No	nama makanan	Porsi/ ukuran	Kalori	total lemak	Lemak jenuh	kolestrol	sodium
1	Kanji Dari Tepung Jagung	100 g	381	0.1g		0	9.00 mg
2	Kacang, pecan	100 g	691	72g	6.2g	0	0.00 mg
3	Terong, Mental	100 g	25	0.2g		0	2.00 mg
4	Teff, Mental	100 g	367	2.4g	0.4g	0	12.00 mg
5	Serbet, Jeruk	100 g	144	2g	1.2g	1mg	46.00 mg
.....	.....	....	....	.....	.....	.....	.....
8784	Daging sapi, mentah, dipotong steak bundar/ bulat menjadi "0 "lemak yang dapat dipisahkan, tanpa tulang	100 g	124	3.3g	1.3g	62mg	54.00 mg

8785	Daging sapi mentah, semua nilai dipotong menjadi '0" lemak, dipisahkan tanpa tulang, di panggang	100 g	125	3.5g	1.4g	62mg	54.00 mg
8786	Domba beku impor Selandia Baru, dimasak hanya ramping yang dapat dipisahkan, gabungan dari potongan ritel	100 g	206	8.9g	3.9g	109mg	50.00 mg
8787	Domba beku impor Selandia Baru, dimasak hanya ramping yang dapat dipisahkan, gabungan dari potongan ritel	100 g	277	23g	12g	78mg	39.00 mg
8788	Daging sapi mentah, semua nilai, dipotong menjadi '0" lemak dipisahkan tanpa tulang, panggang bundar	100 g	121	3g	1.1g	60mg	53.00 mg

No	.....	Protein	Karbohidrat	Serat	Gula	Lemak
1	.....	0.26 g	91.27 g	0.9 g	0.00 g	0.05 g
2	.....	9.17 g	13.86 g	9.6 g	3.97 g	71.97 g
3	....	0.98 g	5.88 g	3.0 g	3.53 g	0.18 g
4	.....	13.30 g	73.13 g	8.0 g	1.84 g	2.38 g
5	.....	1.10 g	30.40 g	1.3 g	24.32 g	2.00 g
.....	.....	.....	.....	.....	....	.....
8784	.....	23.49 g	0.00 g	0.0 g	0.00 g	3.34 g
8785	.....	23.45 g	0.00 g	0.0 g	0.00 g	3.50 g
8786	.....	29.59 g	0.00 g	0.0 g	0	8.86 g
8787	.....	16.74 g	0.00 g	0.0 g	0	22.74 g
8788	.....	23.37 g	0.00 g	0.0 g	0.00 g	3.04

### Pra-Pemrosesan Data

Pada tahap ini, tahap awal proses pengolahan data menta yang di ubah lebih bersih dan siap digunakan untuk analisis atau pelatihan model machine learning dengan meghilangan satuan, pembersihan data yang tidak relevan, menambahkan kategori atau label, penanganan data yang hilang atau missing value dan normalisasi data agar saat pemodelan algoritma data dapat diolah dan tidak terjadi eror saat training data. Gambar dibawah ini menunjukkan bahwa pada data nutrisi lemak jenuh terdapat 1590 data yang hilang.

```
"Missing Values:\n",
" nama                0\n",
"porsi_ukuran        0\n",
"Kalori              0\n",
"total lemak         0\n",
"lemak jenuh        1590\n",
"kolestrol           0\n",
"sodium              0\n",
"kolin               0\n",
"folat               0\n",
"vitamin_a           0\n",
"vitamin_a_rae      0\n",
"vitamin_b12        0\n",
"vitamin_b6         0\n",
"vitamin_c           0\n",
"vitamin_d           0\n",
"vitamin_e           0\n",
"vitamin_k           0\n",
"calcium             0\n",
"magnesium           0\n",
"manganese           0\n",
"phosphorous        0\n",
"potassium           0\n",
"selenium            0\n",
"zink                0\n",
"protein             0\n",
"carbohydrate       0\n",
"fiber               0\n",
"sugars              0\n",
"fat                 0\n",
"water              0\n",
```

Gambar 1. Missing Value

Hasil pra-pemrosesan data dengan meghilangan satuan, pembersihan data yang tidak relevan, menambahkan kategori atau label, penanganan data yang hilang atau missing value dan normalisasi data dilihat pada tabel dibawah ini.

Tabel 3. Hasil Pra-perosesan Data

No	Kalori	Total lemak	Lemak jenuh	Kolestrol	Sodium	.....
1	0,4224	0,001	0,0427	0	0,0002	.....
2	0,7661	0,72	0,0636	0	0	....
3	0,0277	0,002	0,0427	0	0	.....
4	0,4069	0,024	0,0031	0	0,0003	.....
5	0,1596	0,02	0,0115	0,0003	0,0012	.....
....	.....	.....	....	.....	.....	.....
8784	0,1386	0,035	0,0136	0,02	0,0014	.....
8785	0,2284	0,089	0,0396	0,0352	0,0013	.....
8786	0,3071	0,23	0,1241	0,0252	0,001	.....
8787	0,1342	0,03	0,0104	0,0194	0,0014	.....
8788	0,1342	0,03	0,0104	0,0194	0,0014	.....

No	Protein	Karbohidrat	Serat	Gula	Lemak	Rekomendasi baru
1	0,0029	0,9127	0	0,0005	0,9	Tidak Direkomendasikan
2	0,1038	0,1386	0,0398	0,7197	9,6	Tidak Direkomendasikan
3	0,0111	0,0588	0,0354	0,0018	3	Direkomendasikan
4	0,1506	0,7313	0,0184	0,0238	8	Direkomendasikan
5	0,0124	0,304	0,2437	0,02	1,3	Cukup Direkomendasikan
....	....	.....	....	.....	.....	....
8784	0,2655	0	0	0,035	0.0000	Tidak Direkomendasikan
8785	0,335	0	0	0,0886	0.0000	Tidak Direkomendasikan
8786	0,1895	0	0	0,2274	0.0000	Tidak Direkomendasikan
8787	0,2646	0	0	0,0304	0.0000	Tidak Direkomendasikan
8788	0,2646	0	0	0,0304	0.0000	Tidak Direkomendasikan

#### Pemodelan Algoritma K-Nearest Neighbors

Dari data hasil pra-pemrosesan data diatas data kemudian dibagi menjadi data latih (training) sebanyak 80% dan data uji (testing) sebanyak 20% (Jiang, 2023). Pengujian objek model K-Nearest Neighbors (K-NN) dengan menggunakan nilai parameter K terbaik untuk menghitung jarak terdekat antara data uji (testing) sebanyak 1758 data dengan melihat K data latih (data training) sebanyak 7030 data.

Tabel 4. Data Training

No	Kalori	Total lemak	Lemak jenuh	Kolestrol	Sodium	.....
2058	0.0488	0.011	0.0031	0.0	0.0037	.....
733	0.1619	0.059	0.0125	0.0145	0.0012	.....
6417	0.306	0.022	0.0052	0.0	0.0088	....
1343	0.3625	0.006	0.0	0.0	0.0046	.....
5459	0.2018	0.065	0.0365	0.0052	0.0014	.....
.....	....	....	....	....	.....	.....

5472	0.1264	0.031	0.0083	0.0181	0.0016	.....
8370	0.1574	0.058	0.0198	0.0187	0.0015	.....
7873	0.1874	0.067	0.0156	0.0245	0.0015	.....
6853	0.4634	0.1	0.0209	0.0	0.0	.....
5869	0.0499	0.008	0.001	0.0019	0.0075	.....

No	protein	Karbohidrat	Serat	Gula	Lemak	Rekomendasi_baru
2058	0.0147	0.082	0.019	0.0441	0.0106	Tidak Direkomendasikan
733	0.2448	0.0	0.0	0.0	0.0593	Tidak Direkomendasikan
6417	0.1004	0.5504	0.038	0.1337	0.0221	Tidak Direkomendasikan
1343	0.591	0.347	0.2	0.0824	0.0055	Tidak Direkomendasikan
5459	0.0455	0.2682	0.0025	0.2442	0.0646	Tidak Direkomendasikan
	....	.....	....	.....		....
5472	0.243	0.0	0.0	0.0	0.0312	Tidak Direkomendasikan
8370	0.237	0.0	0.0	0.0	0.0579	Tidak Direkomendasikan
7873	0.296	0.0089	0.0	0.0	0.0674	Tidak Direkomendasikan
6853	0.0634	0.756	0.0215	0.4008	0.103	Tidak Direkomendasikan
5869	0.0277	0.0694	0.0152	0.0122	0.0082	Tidak Direkomendasikan

Tabel 5. Data Testing

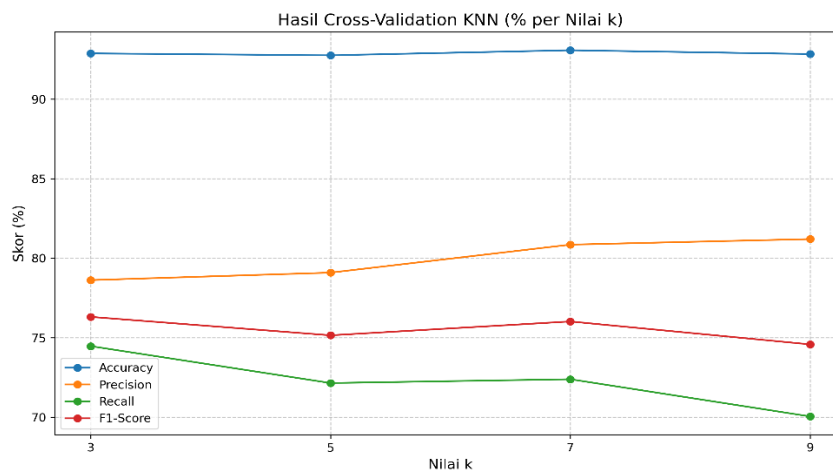
No	Kalori	Total lemak	Lemak jenuh	Kolestrol	Sodium	.....
2019	0.043	0.0146	0.0258	0.0012	0.0012	.....
3292	0.0	0.0427	0.0	0.0	0.0	.....
8082	0.28	0.1241	0.0026	0.004	0.004	.....
6178	0.59	0.0719	0.0	0.0029	0.0029	.....
3831	0.25	0.0417	0.0023	0.0216	0.0216	.....
....	.....	....	.....	.....	....	.....
8400	0.038	0.0042	0.0203	0.0254	0.0254	.....
5202	0.051	0.0083	0.0	0.0094	0.0094	.....
2778	0.002	0.0427	0.0	0.0062	0.0062	.....
7558	0.081	0.0313	0.0148	0.0163	0.0163	.....
987	0.052	0.0188	0.0016	0.0109	0.0109	.....

No	Protein	Karbohidrat	Serat	Gula	Lemak	Rekomendasi_baru
2019	0.2514	0.0	0.0	0.0	0.043	Tidak Direkomendasikan
3292	0.0079	0.104	0.0	0.0	0.0	Tidak Direkomendasikan
8082	0.1228	0.57	0.0	0.5711	0.28	Tidak Direkomendasikan
6178	0.2038	0.1902	0.0899	0.0391	0.588	Direkomendasikan
3831	0.111	0.582	0.0392	0.0	0.25	Tidak Direkomendasikan
....	.....	....	.....	....	.....	.....
8400	0.2554	0.0108	0.0	0.0108	0.0378	Tidak Direkomendasikan
5202	0.1402	0.749	0.1291	0.1637	0.0511	Tidak Direkomendasikan
2778	0.0094	0.0814	0.0316	0.0321	0.0023	Tidak Direkomendasikan
7558	0.1946	0.1987	0.0152	0.0469	0.0805	Cukup Direkomendasikan
987	0.0627	0.2163	0.0696	0.0	0.0515	Tidak Direkomendasikan

Pemodelan algoritma K-Nearest Neighbors untuk pengukuran jarak yang paling sering digunakan adalah Euclidean distance. Jarak euclidean dapat dicari menggunakan persamaan:

$$= \sqrt{\sum_{i=1}^n (X_i - Y_i)^2}$$

Untuk menentukan nilai K terbaik menggunakan teknik 5-fold cross-validation untuk mendapatkan akurasi tertinggi. Proses pengujian di lakukan dengan mencoba K= 3, 5, 7 dan 9. Hasil pengujian menunjukkan bahwa nilai K= 7 menghasilkan akurasi tertinggi 93.07%. Namun pada penelitian ini memilih k=3. Meskipun akurasinya lebih rendah dari k=7 namun hasil F1-score pada k=3 lebih tinggi yang mencerminkan performa seimbang pada semua kelas jika kelas tidak seimbang, bukan hanya mayoritas, tetapi juga memberi permforma model yang baik. Berikut gambar grafis yang menampilkan hubungan antara nilai K dan akurasi, presisi, recall dan fl\_score model:



Gambar 2. Grafik Perbandingan Nilai K

Tabel 6. Nilai K

Nilai K	Matrix			
	Akurasi	Presisi	Recall	F1-Score
3	92.87%	78.62%	74.47%	76.31%
5	92.75%	79.09%	72.15%	75.15%
7	93.07%	80.85%	72.39%	76.02%
9	92.83%	81.20%	70.05%	74.58%

Hasil prediksi data latih dengan melihat K terdekat setelah pemodelan data latih terdapat pada gambar dibawah ini:

Tabel 7. Hasil Prediksi Data Testing

No	Rekomendasi baru	Prediksi
2019	Tidak Direkomendasikan	Tidak Direkomendasikan
3292	Tidak Direkomendasikan	Tidak Direkomendasikan
8082	Tidak Direkomendasikan	Tidak Direkomendasikan
6178	Direkomendasikan	Direkomendasikan
3831	Tidak Direkomendasikan	Tidak Direkomendasikan
.....	.....	.....
8400	Tidak Direkomendasikan	Tidak Direkomendasikan
5202	Tidak Direkomendasikan	Tidak Direkomendasikan
2778	Tidak Direkomendasikan	Tidak Direkomendasikan
7558	Cukup Direkomendasikan	Cukup Direkomendasikan
987	Tidak Direkomendasikan	Tidak Direkomendasikan

### Evaluasi

Berdarkan hasil prediksi model, maka hasil evaluasi performa dari model klasifikasi diet tinggi serat menghasilkan cukup direkomendasikan diprediksi tepat sebanyak 41 data, direkomendasikan diprediksi tepat sebanyak 90 data dan tidak direkomendasikan 1502 data. Hasil akurasi prediksi model diet tinggi serat dalam mencegah penyakit jantung adalah 93%, dengan klasifikasi matrixnya sebagai berikut:

Tabel 8. Hasil Klasifikasi Matrix

HASIL KLASIFIKASI			
Cukup Direkomendasikan	41	0	22
Direkomendasikan	0	90	51
Tidak Direkomendasikan	14	38	1502
	Cukup Direkomendasikan	Direkomendasikan	Tidak Direkomendasikan

### Validasi Data

Pada tabel di bawah ini adalah hasil validasi rekomendasi makanan diet tinggi serat mencegah penyakit jantung berupa makanan yang kaya akan tinggi serat, rendah lemak, rendah garam, tinggi protein dari data hasil prediksi model sangat cocok dengan data dari ahli gizi/kesehatan. Ini menunjukkan bahwa sistem mampu rekomendasi makanan untuk diet tinggi serat dalam mencegah penyakit jantung.

Tabel 9. Validasi Hasil Prediksi Algoritma

Data Dari Ahli Gizi/Kesehatan		Dari Model Prediksi KNN	
Rekomendasi Makanan	Label rekomendasi	Rekomendasi Makanan	Label Rekoemndasi
Kacang merah rebus	Direkoemndasikan	Kacang merah direbus tanpa garam	Direkomendasikan
Kentang direbus dengan kulit tanpa garam,	Direkoemndasikan	Kentang direbus dengan kulit tanpa garam,	Direkoemndasikan
Pir mentah	Direkoemndasikan	Pir mentah	Direkoemndasikan

Dari hasil validasi data dari Ahli gizi/kesehatan dan prediksi model, makanan kacang merah direbus cocok untuk rekomendasi makanan diet tinggi serat dalam mencegah penyakit jantung. Dari hasil makanan yang direkomendasikan ini Ahli gizi/kesehatan juga memberikan saran diet harian, yaitu:

#### 10. Hasil Validasi Diet Harian

Rekomendasi makanan	Porsi/Ukuran perhari	Keterangan
Kacang merah rebus	½ cup/30gram	Dikonsumsi 3-4 kali dalam seminggu
Kentang direbus dengan kulit tanpa garam,	1-2 buah/100gram	Dikonsumsi setiap hari
Pir mentah	1-2 buah	Dikonsumsi setiap hari

#### 4. Kesimpulan

Penelitian ini membahas evaluasi model klasifikasi berbasis algoritma K-Nearest Neighbors (KNN) dalam merekomendasikan makanan untuk mencegah penyakit jantung berdasarkan tiga label kelas, yaitu: direkomendasi, cukup direkomendasi, dan tidak direkomendasi. Melalui perhitungan confusion matrix, model dapat mengklasifikasikan data dengan tingkat akurasi mencapai 93%, ini menunjukkan kinerja model sangat baik dalam mengidentifikasi kategori makanan yang sesuai. Pemodelan hasil algoritma ini lebih andal dalam mengklasifikasikan makanan yang tidak direkomendasikan sehingga kurang optimal dalam mengklasifikasikan makanan yang direkomendasikan secara penuh. Disarankan untuk penelitian selanjutnya dapat membandingkan performa model KNN dengan algoritma klasifikasi lainnya seperti Random Forest, SVM, atau Gradient Boosting untuk memperoleh pendekatan terbaik yang paling cocok dengan karakteristik data.

#### Referensi

1. Belarminus, P., & Boa, G. F. (2022). Factors affecting the occurrence of coronary heart disease in the General Hospital of Waikabubak, Indonesia. *KnE Life Sciences*, 1004–1012. <https://doi.org/10.18502/kl.v7i2.10338>
2. de Castro, G. F., & Tinós, R. (2022, November). K-Nearest Neighbors based on the Nk Interaction Graph. In *Encontro Nacional de Inteligência Artificial e Computacional (ENIAC)* (pp. 694–703). SBC. <https://doi.org/10.5753/eniac.2022.229928>
3. Gopalan, C., & Kirk, E. (2022). *Biology of cardiovascular and metabolic diseases*. Academic Press. <https://doi.org/10.1016/C2020-0-00109-0>
4. Jiang, D., Zhang, C., & Song, Y. (2023). Pre-processing of training data. In *Probabilistic topic models: Foundation and application* (pp. 47–51). Singapore: Springer Nature Singapore. [https://doi.org/10.1007/978-981-99-4592-1\\_3](https://doi.org/10.1007/978-981-99-4592-1_3)
5. Mahsin, M. (2022). Techniques for reporting quantitative data. In *Principles of social research methodology* (pp. 257–260). Singapore: Springer Nature Singapore. [https://doi.org/10.1007/978-981-19-1794-8\\_22](https://doi.org/10.1007/978-981-19-1794-8_22)
6. Naidu, G., Zuva, T., & Sibanda, E. M. (2023, April). A review of evaluation metrics in machine learning algorithms. In *Computer science on-line conference* (pp. 15–25). Cham: Springer International Publishing. [https://doi.org/10.1007/978-3-031-30681-6\\_2](https://doi.org/10.1007/978-3-031-30681-6_2)
7. Rijnaarts, I., de Roos, N. M., Wang, T., Zoetendal, E. G., Top, J., Timmer, M., ... & de Wit, N. (2022). A high-fibre personalised dietary advice given via a web tool reduces constipation complaints in adults. *Journal of Nutritional Science*, 11, e31. <https://doi.org/10.1017/jns.2022.29>
8. Tiu, E. S. K., Huang, Y. F., Ng, J. L., AlDahoul, N., Ahmed, A. N., & Elshafie, A. (2022). An evaluation of various data pre-processing techniques with machine learning models for water level prediction. *Natural Hazards*, 110(1), 121–153. <https://doi.org/10.1007/s11069-021-05026-2>
9. Tovar, A. R., Guevara-Cruz, M., Serralde Zúñiga, A. E., & Torres, N. (2020). Dietary fiber and hyperlipidemia and cardiovascular disease. In *Science and technology of fibers in food systems* (pp. 219–239). Cham: Springer International Publishing. [https://doi.org/10.1007/978-3-030-38654-2\\_10](https://doi.org/10.1007/978-3-030-38654-2_10)
10. Waranugraha, Y. (2022). Overcoming high cardiovascular disease burden in Indonesia: The importance of massive cardiovascular disease risk factor screening, aggressive guideline-directed treatment, and community-based programs. *Heart Science Journal*, 3(4), 1–3. <https://doi.org/10.37275/hsc.v3i4.72>
11. Ahmad, Z., Sarmiento, R. F., & Lee, Y. (2021). The role of dietary fiber in managing cardiovascular diseases. *Journal of Clinical Nutrition and Metabolism*, 5(2), 45–53. <https://doi.org/10.1016/j.jcnm.2021.05.003>
12. Alizadehsani, R., Abdar, M., Roshanzamir, M., Khosravi, A., Panahiazar, M., & Acharya, U. R. (2019). Machine learning-based coronary artery disease diagnosis: A comprehensive review. *Computers in Biology and Medicine*, 111, 103346. <https://doi.org/10.1016/j.compbiomed.2019.103346>
13. Bansal, M., & Goyal, M. (2021). Dietary interventions for cardiovascular health: Current perspectives. *Nutrition Reviews*, 79(5), 593–607. <https://doi.org/10.1093/nutrit/nuaa025>
14. Chen, J., & Huang, Y. (2020). K-nearest neighbor algorithm and its application in disease prediction. *Journal of Healthcare Engineering*, 2020, 8853745. <https://doi.org/10.1155/2020/8853745>
15. Dhingra, R., & Vasan, R. S. (2017). Age as a cardiovascular risk factor. *Medical Clinics of North America*, 101(1), 1–17. <https://doi.org/10.1016/j.mcna.2016.08.012>
16. Ho, Y. C., & Chen, S. H. (2022). Data preprocessing in machine learning for health applications. *Frontiers in Artificial Intelligence*, 5, 856500. <https://doi.org/10.3389/frai.2022.856500>

DOI: <https://doi.org/10.31004/riggs.v4i3.2099>

Lisensi: Creative Commons Attribution 4.0 International (CC BY 4.0)

17. Kris-Etherton, P. M., Petersen, K. S., & Hibbeln, J. R. (2021). Nutrition and cardiovascular health: The evidence base. *American Journal of Clinical Nutrition*, 114(5), 1440–1456. <https://doi.org/10.1093/ajcn/nqab262>
18. Mishra, A., & Sharma, R. (2021). Application of K-nearest neighbor algorithm for cardiovascular disease prediction. *Procedia Computer Science*, 189, 234–241. <https://doi.org/10.1016/j.procs.2021.05.094>
19. Threapleton, D. E., Greenwood, D. C., Evans, C. E., Cleghorn, C. L., Nykjaer, C., Woodhead, C., ... & Burley, V. J. (2013). Dietary fibre intake and risk of cardiovascular disease: Systematic review and meta-analysis. *BMJ*, 347, f6879. <https://doi.org/10.1136/bmj.f6879>
20. Zhang, Y., & Zhao, X. (2020). The role of sodium intake in cardiovascular disease: A review. *Nutrients*, 12(1), 343. <https://doi.org/10.3390/nu12020343>