

ANALISIS PREDIKTIF KANDUNGAN MAKRONUTRIEN PADA MAKANAN TRADISIONAL PAGAR ALAM MENGGUNAKAN ALGORITMA K-NEAREST NEIGHBOR REGRESSION

Lizky Suprianto¹, Yadi², Sigit Candra Setya^{3*}

^{1,2,3}Program Studi Teknik Informatika, Institut Teknologi Pagar Alam, Pagar Alam, Indonesia

*Penulis korespondensi: sigitcs1969@gmail.com

ABSTRAK

Makanan tradisional Pagar Alam merupakan bagian dari warisan kuliner daerah yang masih dikonsumsi luas oleh masyarakat, namun informasi kandungan gizinya, khususnya zat gizi makro, masih sangat terbatas karena pengujian laboratorium memerlukan biaya, waktu, dan sumber daya yang besar. Penelitian ini bertujuan membangun model prediksi kandungan karbohidrat, protein, dan lemak pada enam jenis makanan tradisional Pagar Alam (Gunjing, Kelicuk, Lemang, Lempuk, Wajik, dan Tekwan) menggunakan algoritma *K-Nearest Neighbor (KNN) Regression* sebagai alternatif estimasi gizi yang lebih cepat dan efisien dibandingkan analisis laboratorium konvensional. Data komposisi bahan diperoleh melalui observasi lapangan dan wawancara langsung dengan pelaku usaha kuliner tradisional, kemudian nilai makronutrien dihitung menggunakan metode agregasi proporsional berdasarkan Tabel Komposisi Pangan Indonesia (TKPI). Fitur model dibentuk dari enam variabel agregat per resep, yaitu jumlah bahan, total berat bahan, dan jumlah bahan pada setiap metode pengolahan (goreng, kukus, rebus, sangrai), yang dinormalisasi menggunakan Min-Max Scaler sebelum dihitung jarak Euclidean-nya. Karena keterbatasan jumlah sampel, validasi model dilakukan menggunakan skema *Leave-One-Out Cross-Validation (LOO-CV)* dengan pengujian tiga nilai K (1, 2, dan 3). Hasil evaluasi menunjukkan bahwa K = 2 memberikan performa terbaik dengan RMSE 18,29 gram, MAE 11,16 gram, dan MSE 334,43 gram², lebih baik dibandingkan K = 1 dan K = 3. Analisis lebih lanjut menunjukkan bahwa *error* prediksi terbesar dipengaruhi oleh keterbatasan variasi data dan sifat fitur yang bersifat agregat resep, bukan komposisi gizi langsung per bahan. Analisis kontribusi terhadap Angka Kecukupan Gizi (AKG) menunjukkan bahwa keenam makanan tradisional tersebut memberikan kontribusi karbohidrat per porsi antara 4,65% hingga 44,20% AKG, dengan Lemang sebagai penyumbang tertinggi. Penelitian ini menyimpulkan bahwa pendekatan KNN Regression berbasis fitur resep memiliki potensi sebagai metode estimasi gizi awal (*proof of concept*), namun memerlukan penambahan jumlah dan variasi sampel pada penelitian selanjutnya agar akurasi prediksi dapat ditingkatkan.

Kata kunci: Makanan Tradisional, Makronutrien, *K-Nearest Neighbor*, *Machine Learning*, Pagar Alam.

1 PENDAHULUAN

Perkembangan teknologi informasi telah memperluas penerapan pendekatan komputasional ke berbagai sektor, termasuk pangan dan gizi. *Machine learning* memungkinkan sistem komputer mempelajari pola dari data komposisi nutrisi sehingga proses estimasi kandungan gizi dapat dilakukan lebih efisien dibandingkan metode konvensional berbasis pengujian laboratorium yang menuntut waktu, biaya, dan sumber daya besar (Davies *et al.*, 2021). Berbagai kajian melaporkan bahwa pendekatan ini mampu memodelkan data nutrisi secara kuantitatif dan menghasilkan estimasi gizi yang akurat berbasis data (Pangestu *et al.*, 2025), bahkan model regresi berbasis *machine learning* dilaporkan mencapai akurasi tinggi dalam memprediksi kepadatan nutrisi makanan secara umum (Yang, 2025). Meskipun demikian, sebagian besar penelitian tersebut masih berfokus pada makanan secara umum dan

belum banyak mengkaji makanan tradisional dengan karakteristik komposisi bahan lokal dan metode pengolahan khas daerah.

Salah satu algoritma *machine learning* yang banyak diterapkan pada data gizi adalah *K-Nearest Neighbor* (KNN), yang bekerja berdasarkan prinsip kemiripan antardata sehingga efektif memodelkan hubungan antarvariabel gizi, baik untuk klasifikasi maupun estimasi nilai secara kuantitatif (Ozturk Kiyak *et al.*, 2023). KNN tidak memerlukan asumsi distribusi data tertentu sehingga cocok diterapkan pada data gizi yang memiliki variasi tinggi dan pola yang kompleks (Pratama *et al.*, 2024). Jajanan tradisional yang dijumpai di pasar tradisional merupakan bagian dari pola konsumsi masyarakat sehari-hari lintas kelompok usia, memiliki nilai budaya dan ekonomi, serta berkontribusi terhadap asupan energi harian. Namun, konsumsinya umumnya tidak disertai informasi gizi yang memadai sehingga pemilihan makanan lebih banyak didasarkan pada rasa dan kebiasaan dibandingkan pertimbangan keseimbangan gizi (Setiawan *et al.*, 2024). Sebagian jajanan tradisional, seperti yang berbahan dasar beras ketan dan gula, memiliki kepadatan energi yang relatif tinggi (Mahmud & Hermana, 2020). Data Survei Kesehatan Indonesia menunjukkan prevalensi obesitas dewasa yang terus meningkat dan telah melampaui dua puluh persen, yang berkaitan dengan pola konsumsi tinggi energi dan ketidakseimbangan asupan zat gizi makro (Kementerian Kesehatan Republik Indonesia, 2025). Meskipun penelitian ini tidak dimaksudkan untuk menetapkan hubungan kausal antara konsumsi jajanan tradisional dan kejadian obesitas, informasi kandungan makronutrien tetap relevan sebagai dasar pemahaman risiko gizi terhadap pola konsumsi masyarakat.

Penentuan kandungan gizi makanan tradisional secara konvensional dilakukan melalui pengujian laboratorium terakreditasi menggunakan metode seperti Kjeldahl untuk protein, ekstraksi Soxhlet untuk lemak, dan gravimetri untuk karbohidrat. Metode ini menghasilkan data dengan akurasi dan reliabilitas tinggi, tetapi memerlukan biaya operasional yang besar serta waktu analisis yang dapat mencapai beberapa hari hingga minggu, sehingga kurang efisien untuk kebutuhan estimasi yang bersifat cepat. Sebagai alternatif, penelitian ini mengusulkan pendekatan estimasi berbasis *machine learning* yang tidak dimaksudkan untuk menggantikan pengujian laboratorium, melainkan sebagai sarana estimasi awal yang lebih cepat, murah, dan mudah diakses, dengan memanfaatkan data komposisi bahan, takaran, dan metode pengolahan makanan tradisional khas Pagar Alam sebagai dasar pembelajaran model.

Berdasarkan permasalahan tersebut, penelitian ini bertujuan untuk membangun model prediksi kandungan karbohidrat, protein, dan lemak pada enam jenis makanan tradisional Pagar Alam menggunakan algoritma KNN Regression, mengevaluasi performa model pada beberapa nilai parameter K menggunakan skema *Leave-One-Out Cross-Validation* (LOO-CV), serta menganalisis kontribusi kandungan makronutrien hasil prediksi terhadap Angka Kecukupan Gizi (AKG) yang dianjurkan bagi masyarakat Indonesia. Hasil penelitian ini diharapkan dapat memberikan estimasi awal kandungan gizi makanan tradisional Pagar Alam secara kuantitatif sebagai alat bantu pendukung informasi, tanpa menggantikan peran analisis laboratorium sebagai standar acuan.

2 METODE

Langkah-langkah Penelitian Secara umum, tahapan penelitian ini meliputi pengumpulan data primer dan sekunder, prapemrosesan data (termasuk agregasi komposisi dan pembentukan matriks fitur), normalisasi data dengan *Min-Max Scaler*, pemodelan menggunakan algoritma *K-Nearest Neighbor (KNN) Regression* dengan skema validasi *Leave-One-Out Cross-Validation* (LOO-CV), serta evaluasi performa model dan analisis kontribusi makronutrien terhadap Angka Kecukupan Gizi (AKG).

Objek penelitian ini adalah komposisi bahan, takaran, dan metode pengolahan enam jenis makanan tradisional khas Pagar Alam, yaitu Gunjing, Kelicuk, Lemang, Lempuk, Wajik,

dan Tekwan. Data primer diperoleh melalui observasi lapangan dan wawancara langsung dengan pelaku usaha kuliner tradisional di Kota Pagar Alam, mencakup jenis bahan, berat bahan dalam satu kali proses produksi, serta jumlah porsi atau produk yang dihasilkan. Data sekunder berupa nilai makronutrien setiap bahan diperoleh dari Tabel Komposisi Pangan Indonesia (TKPI) (Mahmud & Hermana, 2020). Nilai kandungan karbohidrat, protein, dan lemak pada setiap makanan tidak diperoleh melalui pengujian laboratorium, melainkan dihitung menggunakan metode agregasi proporsional, yaitu dengan mengalikan berat setiap bahan terhadap nilai makronutrien rujukan TKPI per 100 gram, kemudian menjumlahkan kontribusi seluruh bahan dalam satu resep. Hasil perhitungan ini selanjutnya digunakan sebagai nilai aktual (*ground truth*) dalam pelatihan dan evaluasi model prediksi.

2.1 Variabel Penelitian dan Pembentukan Fitur

Karena setiap sampel data merepresentasikan satu jenis makanan dan bukan satu bahan baku, variabel input (fitur) yang dibentuk bersifat agregat pada level resep. Berdasarkan hasil agregasi data komposisi, diperoleh enam fitur numerik, yaitu Jumlah Bahan, Total Gram, dan jumlah bahan yang diproses dengan masing-masing metode pengolahan (Metode_Goreng, Metode_Kukus, Metode_Rebus, dan Metode_Sangrai). Variabel metode pengolahan direpresentasikan langsung sebagai fitur numerik berupa banyaknya bahan yang diproses dengan metode tersebut, sehingga tidak diperlukan proses *One-Hot Encoding* tambahan pada tahap ini. Variabel output (target) yang diprediksi adalah tiga nilai makronutrien hasil perhitungan TKPI, yaitu kandungan karbohidrat, protein, dan lemak per porsi, sehingga model yang dibangun bersifat regresi multi-output. Matriks fitur (X) berukuran 6 baris $\times$ 6 kolom dan matriks target (y) berukuran 6 baris $\times$ 3 kolom, sesuai jumlah makanan, jumlah fitur, dan jumlah target yang ditetapkan, sebagaimana disajikan pada Tabel 1.

Tabel 1. Matriks Fitur (X) Sebelum Normalisasi

Makanan	Jml. Bahan	Total Gram	Mtd. Goreng	Mtd. Kukus	Mtd. Rebus	Mtd. Sangrai
Gunjing	5	45,56	5	0	0	0
Kelicuk	4	206,67	0	4	0	0
Lemang	4	201,20	0	4	0	0
Lempuk	2	62,50	0	0	2	0
Wajik	3	115,00	0	1	2	0
Tekwan	5	54,70	0	0	5	0

2.2 Pra-Pemrosesan Data: Normalisasi Min-Max

Fitur numerik pada penelitian ini memiliki rentang skala yang berbeda jauh, misalnya Total Gram yang mencapai nilai 206,67 sedangkan fitur Metode_Goreng hanya bernilai 0 sampai 5. Apabila tidak dinormalisasi, fitur berskala besar akan mendominasi perhitungan jarak Euclidean pada algoritma KNN (Ozturk Kiyak *et al.*, 2023). Oleh karena itu, seluruh fitur dinormalisasi menggunakan *Min-Max Scaler* ke rentang [0, 1] dengan persamaan:

$$X' = \frac{X - X_{min}}{X_{max} - X_{min}}$$

dengan X' adalah nilai fitur setelah normalisasi, X adalah nilai fitur asli, Xmin adalah nilai minimum, dan Xmax adalah nilai maksimum fitur tersebut pada data latih. Normalisasi tidak dilakukan satu kali terhadap seluruh data, melainkan diulang pada setiap iterasi *Leave-One-Out Cross-Validation* (LOO-CV). *Scaler* hanya di-fit menggunakan data latih pada iterasi

tersebut, kemudian diterapkan (*transform*) pada data uji, guna mencegah kebocoran data (*data leakage*), yaitu kondisi ketika informasi data uji secara tidak langsung memengaruhi proses pelatihan model.

2.3 Skema Validasi: Leave-One-Out Cross-Validation

Dataset penelitian ini hanya terdiri dari 6 sampel makanan, sehingga pembagian data secara konvensional, misalnya 80:20, hanya akan menyisakan 1–2 sampel sebagai data uji, jumlah yang tidak representatif secara statistik. Untuk mengatasi keterbatasan tersebut, validasi model dilakukan menggunakan *Leave-One-Out Cross-Validation* (LOO-CV), yaitu skema validasi yang pada setiap iterasi menjadikan satu sampel sebagai data uji dan lima sampel sisanya sebagai data latih. Proses ini diulang sebanyak jumlah sampel data sehingga terdapat enam iterasi. Pada setiap iterasi, model *KNeighborsRegressor* dilatih menggunakan data latih yang telah dinormalisasi dengan metrik jarak Euclidean, kemudian digunakan untuk memprediksi nilai makronutrien pada satu sampel data uji yang tidak dilibatkan dalam pelatihan.

2.4 Algoritma K-Nearest Neighbor Regression

Algoritma *K-Nearest Neighbor* (KNN) Regression memprediksi nilai target suatu data uji berdasarkan rata-rata nilai target dari K data latih yang memiliki jarak terdekat dalam ruang fitur (Ozturk Kiyak *et al.*, 2023). Jarak antar data dihitung menggunakan metrik jarak Euclidean, yang dirumuskan sebagai:

$$d(p, q) = \sqrt{\sum_{i=1}^n (p_i - q_i)^2}$$

dengan $d(p, q)$ adalah jarak Euclidean antara data uji p dan data latih q , n adalah jumlah fitur ($n = 6$), serta p_i dan q_i adalah nilai fitur ke- i pada data uji dan data latih yang telah dinormalisasi. Setelah jarak terhadap seluruh data latih dihitung, dipilih K data dengan jarak terkecil, kemudian nilai prediksi dihitung sebagai rata-rata aritmetik nilai target K tetangga tersebut:

$$\hat{y} = \frac{1}{K} \sum_{k=1}^K y_k$$

dengan $\hat{y}$ adalah nilai prediksi dan y_k adalah nilai target tetangga ke- k . Algoritma ini dipilih karena tidak memerlukan asumsi distribusi data tertentu, mampu menangkap hubungan nonlinier antarvariabel, serta relatif sederhana untuk diimplementasikan pada dataset berukuran terbatas (Pratama *et al.*, 2024).

2.5 Pengujian Nilai K dan Metrik Evaluasi

Pengujian nilai K dilakukan terhadap tiga nilai, yaitu $K = 1$, $K = 2$, dan $K = 3$. Pembatasan nilai K maksimum sebesar 3 dilakukan karena setiap iterasi LOO-CV hanya menyediakan 5 data latih, sehingga nilai K yang terlalu besar berisiko menghasilkan estimasi yang terlalu rata-rata (*over-smoothing*) dan kehilangan karakteristik individual tiap makanan. Performa setiap nilai K dievaluasi menggunakan tiga metrik, yaitu *Mean Absolute Error* (MAE), *Mean Squared Error* (MSE), dan *Root Mean Squared Error* (RMSE), yang dihitung berdasarkan selisih antara nilai aktual hasil perhitungan TKPI dan nilai prediksi model pada seluruh sampel hasil LOO-CV, mencakup ketiga target makronutrien secara bersamaan. Ketiga metrik dirumuskan sebagai:

$$MAE = \frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}_i|$$

$$MSE = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2$$

$$RMSE = \sqrt{MSE}$$

Dengan N adalah jumlah total observasi ($6 \text{ sampel} \times 3 \text{ target} = 18$), y_i adalah nilai aktual, dan $\hat{y}_i$ adalah nilai prediksi. Nilai K dengan RMSE terendah dipilih sebagai konfigurasi terbaik karena RMSE memberikan penalti lebih besar terhadap kesalahan prediksi yang nilainya jauh dari nilai aktual, sehingga lebih representatif sebagai kriteria seleksi model dibandingkan MAE.

2.6 Analisis Kontribusi Makronutrien terhadap Angka Kecukupan Gizi (AKG)

Untuk menilai relevansi gizi dari hasil prediksi, dilakukan analisis kontribusi kandungan karbohidrat, protein, dan lemak setiap makanan terhadap Angka Kecukupan Gizi (AKG) harian yang dianjurkan. Nilai AKG yang digunakan mengacu pada Peraturan Menteri Kesehatan Republik Indonesia Nomor 28 Tahun 2019 tentang Angka Kecukupan Gizi yang Dianjurkan untuk Masyarakat Indonesia, yaitu karbohidrat 300 gram, protein 60 gram, dan lemak 67 gram per hari untuk orang dewasa dengan kebutuhan energi 2.150 kkal. Persentase kontribusi AKG dihitung dengan persamaan:

$$\%AKG = \left(\frac{\text{Kandungan Makronutrien per Porsi}}{\text{Nilai AKG yang Dianjurkan}} \right) \times 100\%$$

Selain perbandingan terhadap AKG dalam satuan gram, dilakukan pula analisis proporsi energi dari masing-masing makronutrien terhadap *Acceptable Macronutrient Distribution Range* (AMDR), yaitu rentang proporsi energi dari karbohidrat, protein, dan lemak yang dianjurkan untuk menjaga kecukupan gizi dan menurunkan risiko penyakit kronis, masing-masing sebesar 45–65% energi dari karbohidrat, 10–35% energi dari protein, dan 20–35% energi dari lemak (National Academies of Sciences, Engineering, and Medicine, 2024; Lee *et al.*, 2015). Kontribusi energi setiap makronutrien dihitung menggunakan faktor konversi Atwater, yaitu 4 kkal/gram untuk karbohidrat dan protein, serta 9 kkal/gram untuk lemak, kemudian dibandingkan terhadap total energi per porsi makanan.

3 HASIL DAN PEMBAHASAN

3.1 Gambaran Umum Hasil Penelitian

Penelitian ini menghasilkan model prediksi kandungan makronutrien (karbohidrat, protein, dan lemak) pada enam jenis makanan tradisional Pagar Alam menggunakan algoritma *K-Nearest Neighbor* (KNN) Regression, sehingga estimasi nilai gizi suatu makanan dapat diperoleh tanpa harus melalui pengujian laboratorium secara langsung. Berdasarkan hasil observasi lapangan dan perhitungan menggunakan Tabel Komposisi Pangan Indonesia (TKPI), kandungan karbohidrat keenam makanan tradisional yang diteliti berada pada rentang 13,95–132,60 gram, protein 1,25–9,10 gram, dan lemak 1,50–18,15 gram per porsi. Variasi nilai yang cukup lebar ini mencerminkan keragaman komposisi bahan dan metode pengolahan antar jenis makanan tradisional yang diteliti.

Tabel 2. Total Kandungan Makronutrien Hasil Observasi

Nama Makanan	Karbohidrat (g)	Protein (g)	Lemak (g)
Gunjing	13,95	1,82	6,55
Kelicuk	102,40	8,40	8,60
Lemang	132,60	9,10	18,15
Lempuk	26,49	1,25	1,50
Wajik	65,24	4,50	4,40
Tekwan	22,02	5,62	2,72

Tabel 2 menunjukkan bahwa Lemang memiliki kandungan karbohidrat dan lemak tertinggi di antara keenam makanan, sejalan dengan dominasi beras ketan putih, kelapa parut, dan gula merah dalam komposisinya, sedangkan Tekwan menunjukkan komposisi makronutrien yang relatif berbeda karena merupakan satu-satunya makanan berbahan dasar ikan dalam penelitian ini, sehingga kandungan proteinnya secara proporsional lebih tinggi dibandingkan makanan berbahan dasar beras ketan.

3.2 Implementasi *K-Nearest Neighbor Regression*: Pembentukan Fitur dan *Encoding*

Implementasi algoritma KNN Regression diawali dengan pembentukan matriks fitur (X) dan target (y) dari data hasil agregasi per makanan menggunakan pustaka *scikit-learn*. Pada tahap perancangan awal, metode pengolahan (kategorikal) umumnya memerlukan transformasi *One-Hot Encoding* (OHE) agar dapat diproses secara numerik oleh algoritma berbasis jarak. Akan tetapi, berdasarkan struktur data hasil agregasi pada penelitian ini, variabel metode pengolahan tidak direpresentasikan sebagai satu kategori tunggal per makanan, melainkan sebagai jumlah bahan yang diproses dengan masing-masing metode (Metode_Goreng, Metode_Kukus, Metode_Rebus, Metode_Sangrai) dalam satu resep, sebagaimana telah ditunjukkan pada Tabel 1. Representasi ini secara fungsional setara dengan hasil *One-Hot Encoding* yang diperluas menjadi nilai frekuensi (bukan hanya 0/1), karena satu resep dapat menggunakan lebih dari satu metode pengolahan dengan proporsi bahan yang berbeda — misalnya Wajik yang diproses dengan kombinasi metode kukus (1 bahan) dan rebus (2 bahan). Dengan demikian, proses *encoding* tambahan secara eksplisit tidak diperlukan lagi pada tahap pembentukan fitur, karena variabel kategorikal metode pengolahan telah dikonversi menjadi bentuk numerik sejak proses agregasi data per resep dilakukan. Pendekatan ini sekaligus mempertahankan informasi proporsi penggunaan setiap metode pengolahan yang akan hilang apabila digunakan OHE konvensional yang hanya mencatat ada/tidaknya suatu kategori.

3.3 Contoh Perhitungan Normalisasi Min-Max pada Iterasi LOO-CV Tekwan

Sebagai ilustrasi rinci proses perhitungan, diambil satu kasus representatif, yaitu iterasi LOO-CV ketika Tekwan menjadi data uji dan kelima makanan lainnya (Gunjing, Kelicuk, Lemang, Lempuk, dan Wajik) menjadi data latih. Nilai minimum dan maksimum setiap fitur dihitung hanya dari data latih, sebagaimana disajikan pada Tabel 3.

Tabel 3. Nilai Minimum, Maksimum Data Latih, dan Nilai Asli Data Uji (Tekwan)

Statistik	Jml. Bahan	Total Gram	Mtd. Goreng	Mtd. Kukus	Mtd. Rebus
Xmin (data latih)	2	45,56	0	0	0
Xmax (data latih)	5	206,67	5	4	2
Tekwan (data uji)	5	54,70	0	0	5

Berdasarkan Tabel 3, terlihat bahwa nilai fitur Metode_Rebus pada Tekwan (5) melampaui nilai maksimum pada data latih (2), karena Tekwan diproses dengan metode rebus pada seluruh lima bahannya sedangkan tidak ada makanan pada data latih dengan jumlah bahan rebus sebanyak itu. Kondisi ini menyebabkan hasil normalisasi pada fitur tersebut bernilai lebih dari 1, yang merupakan konsekuensi wajar dari pendekatan *Min-Max Scaler* yang di-fit ulang pada setiap iterasi LOO-CV untuk mencegah kebocoran data. Sebagai contoh perhitungan, normalisasi fitur Total Gram pada Tekwan dihitung sebagai berikut:

$$X'(Total\ Gram) = (54,70 - 45,56) / (206,67 - 45,56) = 9,14 / 161,11 = 0,0567$$

Perhitungan yang sama diterapkan pada seluruh fitur dan seluruh data latih, menghasilkan matriks fitur ternormalisasi sebagaimana disajikan pada Tabel 4.

Tabel 4. Hasil Normalisasi Min-Max Data Latih dan Data Uji (Iterasi Tekwan)

Makanan	Jml. Bahan Norm.	Total Gram Norm.	Mtd. Goreng Norm.	Mtd. Kukus Norm.	Mtd. Rebus Norm.
Gunjing	1,0000	0,0000	1,0000	0,0000	0,0000
Kelicuk	0,6667	1,0000	0,0000	1,0000	0,0000
Lemang	0,6667	0,9660	0,0000	1,0000	0,0000
Lempuk	0,0000	0,1051	0,0000	0,0000	1,0000
Wajik	0,3333	0,4310	0,0000	0,2500	1,0000
Tekwan (uji)	1,0000	0,0567	0,0000	0,0000	2,5000

3.4 Contoh Perhitungan Jarak Euclidean dan Prediksi KNN

Berdasarkan data ternormalisasi pada Tabel 4, jarak Euclidean antara Tekwan (data uji) dan setiap data latih dihitung menggunakan keenam fitur (termasuk Metode_Sangrai yang seluruhnya bernilai 0 sehingga tidak memengaruhi hasil). Sebagai contoh, jarak antara Tekwan dan Wajik dihitung sebagai berikut:

$$d(\text{Tekwan}, \text{Wajik}) = \sqrt{(1,0000 - 0,3333)^2 + (0,0567 - 0,4310)^2 + (0 - 0)^2 + (0 - 0,2500)^2 + (2,5000 - 1,0000)^2 + (0 - 0)^2}$$

$$d(\text{Tekwan}, \text{Wajik}) = \sqrt{0,4444 + 0,1401 + 0 + 0,0625 + 2,2500 + 0} = \sqrt{2,8970} = 1,7021$$

Perhitungan yang sama dilakukan terhadap seluruh data latih, menghasilkan jarak Euclidean sebagaimana disajikan pada Tabel 5.

Tabel 5. Jarak Euclidean Tekwan terhadap Seluruh Data Latih

Data Latih	Jarak Euclidean	Peringkat
Wajik	1,7021	1 (terdekat)
Lempuk	1,8034	2 (terdekat)
Gunjing	2,6932	3
Lemang	2,8615	4
Kelicuk	2,8724	5

Berdasarkan Tabel 5, dengan $K = 2$, dua tetangga terdekat Tekwan adalah Wajik (jarak 1,7021) dan Lempuk (jarak 1,8034). Nilai prediksi makronutrien Tekwan dihitung sebagai rata-rata nilai aktual kedua tetangga tersebut:

$$\text{Karbohidrat: } \hat{y} = \frac{65,24+26,49}{2} = 45,86 \text{ gram}$$

$$\text{Protein: } \hat{y} = \frac{4,50+1,25}{2} = 2,875 \approx 2,88 \text{ gram}$$

$$\text{Lemak: } \hat{y} = \frac{4,40+1,50}{2} = 2,95 \text{ gram}$$

3.5 Contoh Perhitungan *Error* pada Kasus Tekwan

Nilai prediksi tersebut selanjutnya dibandingkan dengan nilai aktual Tekwan (Karbohidrat 22,02 g; Protein 5,62 g; Lemak 2,72 g) untuk memperoleh *error* absolut pada kasus ini:

$$\text{Error Karbohidrat: } |22,02 - 45,86| = 23,84 \text{ gram}$$

$$\text{Error Protein: } |5,62 - 2,88| = 2,74 \text{ gram}$$

$$\text{Error Lemak: } |2,72 - 2,95| = 0,23 \text{ gram}$$

Hasil tersebut menunjukkan bahwa *error* prediksi pada makronutrien lemak relatif kecil (0,23 gram), sedangkan *error* pada karbohidrat jauh lebih besar (23,84 gram). Hal ini terjadi karena posisi Tekwan dalam ruang fitur ternormalisasi lebih dekat dengan makanan berbahan dasar beras ketan (Wajik dan Lempuk) dibandingkan dengan karakteristik komposisi gizi sebenarnya, mengingat fitur yang digunakan bersifat agregat resep (jumlah bahan, total berat, dan metode pengolahan) dan tidak secara langsung merepresentasikan jenis kandungan gizi tiap bahan baku.

3.6 Hasil Prediksi Seluruh Makanan dengan $K = 2$

Prosedur perhitungan yang sama, sebagaimana diilustrasikan pada kasus Tekwan, diterapkan pada seluruh enam iterasi LOO-CV menggunakan $K = 2$, sehingga diperoleh hasil prediksi makronutrien untuk setiap makanan sebagaimana disajikan pada Tabel 6.

Tabel 6. Perbandingan Nilai Aktual dan Prediksi KNN ($K=2$, LOO-CV) — gram per porsi

Makanan	Karbo Aktual	Karbo Prediksi	Protein Aktual	Protein Prediksi	Lemak Aktual	Lemak Prediksi
Gunjing	13,95	43,63	1,82	5,06	6,55	3,56
Kelicuk	102,40	98,92	8,40	6,80	8,60	11,27
Lemang	132,60	83,82	9,10	6,45	18,15	6,50
Lempuk	26,49	43,63	1,25	5,06	1,50	3,56
Wajik	65,24	24,25	4,50	3,44	4,40	2,11
Tekwan	22,02	45,86	5,62	2,88	2,72	2,95

Tabel 6 menunjukkan bahwa Gunjing dan Lempuk memperoleh nilai prediksi yang identik pada ketiga makronutrien (Karbohidrat 43,63 g; Protein 5,06 g; Lemak 3,56 g). Hal ini terjadi karena, ketika masing-masing dikeluarkan sebagai data uji pada iterasinya sendiri, kedua makanan tersebut sama-sama memiliki dua tetangga terdekat yang sama dalam ruang fitur ternormalisasi, yaitu Wajik dan Tekwan, sehingga nilai prediksi $K = 2$ merupakan rata-rata target dari dua tetangga yang identik. Kondisi ini menjadi indikasi awal mengenai keterbatasan variasi data yang dibahas lebih lanjut pada Subbab Faktor Penyebab *Error*.

3.7 Evaluasi Performa Model: MAE, MSE, dan RMSE pada Tiga Nilai K

Evaluasi performa model dilakukan dengan menghitung MAE, MSE, dan RMSE pada seluruh hasil LOO-CV mencakup ketiga target makronutrien secara bersamaan ($N = 18$ observasi), untuk tiga nilai K yang diuji. Sebagai ilustrasi, perhitungan MAE keseluruhan pada $K = 2$ menggunakan 18 nilai *error* absolut yang diperoleh dari 6 makanan $\times$ 3 makronutrien. Sebagai contoh, tiga di antaranya adalah *error* pada Tekwan sebagaimana dihitung pada Subbab Contoh Perhitungan *Error* (23,84; 2,74; dan 0,23 gram), yang kemudian dijumlahkan dengan 15 *error* absolut lainnya dan dibagi 18:

$$MAE(K=2) = (23,84 + 2,74 + 0,23 + \dots + |error\ ke-18|) / 18 = 11,1629\ gram$$

Prosedur perhitungan yang identik diterapkan pada $K = 1$ dan $K = 3$, serta untuk metrik MSE dan RMSE, menghasilkan rekapitulasi sebagaimana disajikan pada Tabel 7.

Tabel 7. Hasil Pengujian Nilai K Menggunakan LOO-CV

Nilai K	MAE (gram)	MSE (gram ²)	RMSE (gram)
K = 1	15,1590	531,3468	23,0510
K = 2	11,1629	334,4308	18,2874
K = 3	12,9018	480,9555	21,9307

Berdasarkan Tabel 7, ketiga nilai K menghasilkan performa yang berbeda. Pada $K = 1$, model hanya mengandalkan satu tetangga terdekat untuk setiap prediksi sehingga sangat sensitif terhadap variasi data dan menghasilkan *error* tertinggi (RMSE = 23,0510 gram). Pada $K = 2$, *error* menurun signifikan menjadi RMSE = 18,2874 gram, MAE = 11,1629 gram, dan MSE = 334,4308 gram², yang merupakan nilai *error* terendah di antara ketiga nilai K. Pada $K = 3$, *error* kembali meningkat (RMSE = 21,9307 gram) karena rata-rata diambil dari tiga tetangga terdekat sementara data latih pada setiap iterasi hanya berjumlah 5 sampel, sehingga prediksi menjadi terlalu rata (*over-smoothing*) dan kurang mampu mengikuti variasi nilai aktual antarmakanan. Berdasarkan kriteria RMSE terendah, $K = 2$ ditetapkan sebagai konfigurasi model terbaik dan digunakan untuk menghasilkan prediksi akhir pada Tabel 6.

3.8 Selisih *Error* per Jenis Makronutrien

Selain evaluasi gabungan, dilakukan pula perhitungan MAE secara terpisah untuk setiap jenis makronutrien guna mengetahui jenis makronutrien mana yang diprediksi lebih akurat oleh model pada $K = 2$, sebagaimana disajikan pada Tabel 8.

Tabel 8. MAE per Makronutrien ($K=2$, LOO-CV)

Makronutrien	MAE (gram)	Rentang Nilai Aktual (gram)
Karbohidrat	27,3193	13,95 – 132,60
Protein	2,5184	1,25 – 9,10
Lemak	3,6509	1,50 – 18,15

Error absolut terbesar terjadi pada prediksi karbohidrat (MAE = 27,3193 gram), diikuti lemak (MAE = 3,6509 gram), dan protein dengan *error* terkecil (MAE = 2,5184 gram). Besarnya MAE karbohidrat secara nominal sejalan dengan rentang nilai aktualnya yang juga paling lebar (13,95 hingga 132,60 gram) dibandingkan protein dan lemak, sehingga selisih absolut pada karbohidrat secara alami lebih besar. Untuk memperoleh gambaran yang lebih adil mengenai performa relatif model pada masing-masing makronutrien, MAE setiap makronutrien dibandingkan terhadap rata-rata nilai aktualnya, menghasilkan *error* relatif sebesar 45,19% untuk karbohidrat, 49,24% untuk protein, dan 52,26% untuk lemak. Hasil ini mengindikasikan

bahwa secara proporsional ketiga makronutrien memiliki tingkat kesalahan yang relatif sebanding, meskipun secara nominal tampak berbeda jauh.

3.9 Faktor Penyebab *Error* Prediksi

Error prediksi yang muncul pada model KNN Regression ini, terutama pada makanan Lemang dan Wajik yang memiliki selisih karbohidrat cukup besar (48,78 gram dan 40,99 gram), bersumber dari karakteristik dasar algoritma KNN yang melakukan prediksi berdasarkan rata-rata nilai target dari tetangga terdekat dalam ruang fitur. Pada kasus Lemang, nilai aktual karbohidrat (132,60 gram) merupakan nilai tertinggi di antara seluruh sampel, namun karena Lemang dikeluarkan dari data latih pada iterasinya sendiri sesuai skema LOO-CV, prediksi yang dihasilkan hanya merupakan rata-rata dari tetangga terdekat yang nilai aktualnya lebih rendah, sehingga terjadi *under-estimation*. Pola sebaliknya terjadi pada makanan dengan nilai aktual rendah seperti Tekwan, sebagaimana ditunjukkan pada contoh perhitungan sebelumnya, di mana prediksi karbohidrat (45,86 gram) justru lebih tinggi dari nilai aktualnya (22,02 gram) karena tetangga terdekatnya (Wajik dan Lempuk) memiliki nilai aktual yang lebih tinggi (*over-estimation*).

3.10 Pengaruh Karakteristik Fitur terhadap Hasil KNN

Fitur yang digunakan dalam penelitian ini, yaitu jumlah bahan, total gram, dan jumlah bahan per metode pengolahan, bersifat agregat pada level resep dan tidak secara langsung merepresentasikan komposisi gizi dari masing-masing bahan baku. Akibatnya, jarak Euclidean antarmakanan dalam ruang fitur ternormalisasi lebih dipengaruhi oleh kemiripan pola resep (jumlah bahan, total berat, dan metode pengolahan yang dominan) dibandingkan kemiripan kandungan gizi yang sesungguhnya. Hal ini terlihat pada kasus Gunjing dan Lempuk yang menghasilkan prediksi identik karena keduanya memiliki dua tetangga terdekat yang sama, yaitu Wajik dan Tekwan, dalam ruang fitur, meskipun nilai aktual makronutrien kedua makanan tersebut berbeda cukup jauh. Fitur metode pengolahan juga cenderung mengelompokkan makanan berdasarkan teknik memasak yang sama, misalnya seluruh makanan yang dikukus atau direbus, namun pengelompokan ini tidak selalu konsisten dengan kemiripan nilai gizi aktualnya, sehingga turut menjadi sumber *error* pada beberapa kasus prediksi.

3.11 Keterbatasan Jumlah dan Variasi Data

Hasil evaluasi model pada penelitian ini sangat dipengaruhi oleh keterbatasan jumlah dan variasi data. Dataset yang digunakan hanya terdiri dari 6 jenis makanan tradisional Pagar Alam, dengan total 23 baris data resep dan 13 bahan unik. Jumlah sampel yang kecil ini menyebabkan setiap iterasi LOO-CV hanya memiliki 5 data latih, sehingga ruang fitur yang dipelajari model menjadi sangat terbatas dan rentan terhadap kasus dua sampel yang kebetulan memiliki tetangga terdekat yang sama, seperti yang terjadi pada Gunjing dan Lempuk. Selain itu, variasi metode pengolahan dalam dataset juga tidak merata, karena hanya tiga metode yang benar-benar muncul dalam data, yaitu goreng, kukus, dan rebus, sementara kategori Metode_Sangrai tidak memiliki sampel sama sekali sehingga fitur tersebut tidak memberikan kontribusi pembeda dalam perhitungan jarak. Keterbatasan ini menjadi salah satu penyebab utama tingginya *error* pada beberapa makanan dengan karakteristik resep yang ekstrem, seperti Lemang dengan total gram bahan yang jauh lebih besar dibandingkan makanan lain.

3.12 Analisis Kelayakan Model

Berdasarkan hasil evaluasi, model KNN Regression dengan $K = 2$ menghasilkan RMSE sebesar 18,2874 gram dan MAE sebesar 11,1629 gram pada skala keseluruhan target, yang menunjukkan performa lebih baik dibandingkan $K = 1$ dan $K = 3$. Namun, secara absolut, *error* tersebut masih tergolong besar apabila dibandingkan dengan rentang nilai aktual pada beberapa makanan, terutama yang memiliki kandungan makronutrien rendah seperti Gunjing dan Tekwan. Ditinjau dari rata-rata *error* relatif terhadap nilai aktual, ketiga makronutrien menunjukkan tingkat kesalahan yang sebanding, yaitu sekitar 45% untuk karbohidrat, 49%

untuk protein, dan 52% untuk lemak, yang mengindikasikan bahwa model belum cukup andal untuk digunakan sebagai alat prediksi makronutrien secara mandiri pada tahap penelitian ini. Dengan mempertimbangkan jumlah sampel yang sangat kecil (6 makanan), model ini lebih tepat dipandang sebagai studi awal (*proof of concept*) yang menunjukkan bahwa pendekatan KNN berbasis fitur resep memiliki potensi untuk dikembangkan, namun memerlukan penambahan jumlah sampel makanan tradisional dan variasi metode pengolahan agar akurasi prediksi dapat ditingkatkan pada penelitian selanjutnya. Temuan ini konsisten dengan kajian sebelumnya yang menekankan bahwa kinerja algoritma KNN sangat dipengaruhi oleh kecukupan jumlah data latih serta representativitas fitur yang digunakan dalam ruang jarak (Ozturk Kiyak *et al.*, 2023; Gilal *et al.*, 2024).

3.13 Analisis Kontribusi Makronutrien terhadap Angka Kecukupan Gizi (AKG)

Analisis kontribusi makronutrien terhadap Angka Kecukupan Gizi (AKG) dilakukan untuk mengetahui seberapa besar sumbangan kandungan karbohidrat, protein, dan lemak dari masing-masing makanan tradisional Pagar Alam terhadap kebutuhan gizi harian yang dianjurkan. Nilai AKG yang digunakan merujuk pada Peraturan Menteri Kesehatan Republik Indonesia Nomor 28 Tahun 2019 tentang Angka Kecukupan Gizi yang Dianjurkan untuk Masyarakat Indonesia, yaitu karbohidrat 300 gram, protein 60 gram, dan lemak 67 gram per hari untuk orang dewasa dengan kebutuhan energi 2.150 kkal. Persentase kontribusi dihitung dengan persamaan $\%AKG = (\text{kandungan makronutrien per porsi} \div \text{nilai AKG yang dianjurkan}) \times 100\%$.

3.14 Contoh Perhitungan Kontribusi AKG (Kasus Tekwan)

Sebagai contoh perhitungan lengkap, kandungan makronutrien aktual Tekwan (Karbohidrat 22,02 g; Protein 5,62 g; Lemak 2,72 g per porsi) dibandingkan terhadap nilai AKG yang dianjurkan sebagai berikut:

$$\%AKG \text{ Karbohidrat} = (22,02 / 300) \times 100\% = 7,34\%$$

$$\%AKG \text{ Protein} = (5,62 / 60) \times 100\% = 9,37\%$$

$$\%AKG \text{ Lemak} = (2,72 / 67) \times 100\% = 4,06\%$$

Hasil tersebut menunjukkan bahwa satu porsi Tekwan memenuhi sekitar 7,34% kebutuhan karbohidrat harian, 9,37% kebutuhan protein harian, dan 4,06% kebutuhan lemak harian orang dewasa. Perhitungan yang identik diterapkan pada kelima makanan lainnya, menghasilkan rekapitulasi kontribusi AKG sebagaimana disajikan pada Tabel 9.

Tabel 9. Kontribusi Kandungan Makronutrien Keenam Makanan Tradisional Pagar Alam terhadap AKG (%)

Makanan	Karbo (g)	%AKG Karbo	Protein (g)	%AKG Protein	Lemak (g)	%AKG Lemak
Gunjing	13,95	4,65%	1,82	3,03%	6,55	9,78%
Kelicuk	102,40	34,13%	8,40	14,02%	8,60	12,84%
Lemang	132,60	44,20%	9,10	15,17%	18,15	27,09%
Lempuk	26,49	8,83%	1,25	2,08%	1,50	2,24%
Wajik	65,24	21,75%	4,50	7,50%	4,40	6,57%
Tekwan	22,02	7,34%	5,62	9,37%	2,72	4,06%

Berdasarkan Tabel 9, keenam makanan tradisional Pagar Alam memiliki kontribusi makronutrien yang bervariasi terhadap AKG harian. Lemang mencatatkan kontribusi karbohidrat tertinggi, yaitu 44,20% AKG, yang mencerminkan dominasi beras ketan putih dan gula merah pada komposisinya, diikuti Kelicuk (34,13%) dan Wajik (21,75%). Gunjing,

Lempuk, dan Tekwan menunjukkan kontribusi karbohidrat yang relatif lebih rendah, masing-masing 4,65%, 8,83%, dan 7,34% (berdasarkan nilai aktual TKPI 22,40 g untuk Tekwan pada perhitungan komposit), konsisten dengan komposisi bahan dan ukuran porsi masing-masing makanan.

Dari sisi protein, Lemang memberikan kontribusi tertinggi sebesar 15,17% AKG, diikuti Kelicuk (14,02%) dan Tekwan (9,37–12,12%, bergantung pada metode perhitungan porsi). Meskipun Tekwan merupakan satu-satunya makanan berbahan dasar ikan dalam dataset ini, kontribusi proteinnya tidak jauh berbeda dari makanan berbahan beras ketan seperti Kelicuk dan Lemang, yang disebabkan oleh ukuran porsi Tekwan yang relatif lebih kecil dibandingkan total berat resep yang besar pada Kelicuk dan Lemang. Lempuk mencatatkan kontribusi protein paling rendah (2,08%), sesuai komposisinya yang hanya terdiri dari dua bahan, yaitu durian dan gula pasir.

Untuk lemak, Lemang kembali menempati posisi tertinggi dengan kontribusi 27,09% AKG, bersumber dari kandungan kelapa parut dalam jumlah besar pada resepnya. Kelicuk menyumbang 12,84% AKG lemak, sedangkan Gunjing memberikan kontribusi 9,78% meskipun porsinya jauh lebih kecil, menunjukkan kepadatan lemak per porsi Gunjing yang relatif tinggi akibat penggunaan santan dan kelapa parut. Lempuk dan Tekwan menunjukkan kontribusi lemak paling rendah, masing-masing 2,24% dan 4,06%, sesuai komposisi bahan yang tidak banyak mengandung sumber lemak.

3.15 Distribusi Energi Makronutrien terhadap AMDR

Selain kontribusi terhadap AKG dalam satuan gram, dilakukan pula analisis proporsi energi makronutrien terhadap *Acceptable Macronutrient Distribution Range* (AMDR), yang menganjurkan proporsi energi sebesar 45–65% dari karbohidrat, 10–35% dari protein, dan 20–35% dari lemak bagi orang dewasa (National Academies of Sciences, Engineering, and Medicine, 2024; Lee *et al.*, 2015). Sebagai contoh, distribusi energi pada satu porsi Tekwan dihitung menggunakan faktor konversi Atwater, yaitu 4 kkal/gram untuk karbohidrat dan protein serta 9 kkal/gram untuk lemak:

$$\text{Energi Tekwan} = (22,02 \times 4) + (5,62 \times 4) + (2,72 \times 9) = 88,08 + 22,48 + 24,48 = 135,04 \text{ kkal}$$

$$\% \text{ Energi Karbohidrat} = (88,08 / 135,04) \times 100\% = 65,23\%$$

$$\% \text{ Energi Protein} = (22,48 / 135,04) \times 100\% = 16,65\%$$

$$\% \text{ Energi Lemak} = (24,48 / 135,04) \times 100\% = 18,13\%$$

Hasil perhitungan menunjukkan bahwa proporsi energi karbohidrat pada Tekwan (65,23%) berada tepat di batas atas rentang AMDR (45–65%), sedangkan proporsi energi protein (16,65%) berada dalam rentang anjuran (10–35%) dan proporsi energi lemak (18,13%) berada sedikit di bawah batas bawah rentang AMDR (20–35%). Pola ini konsisten dengan karakteristik Tekwan sebagai makanan berbahan dasar tepung tapioka dan ikan, yang energinya didominasi oleh karbohidrat dari tepung meskipun memiliki kandungan protein yang relatif tinggi dibandingkan makanan lain dalam penelitian ini. Secara umum, makanan tradisional Pagar Alam yang berbahan dasar beras ketan, seperti Lemang dan Kelicuk, diperkirakan memiliki proporsi energi karbohidrat yang melampaui batas atas AMDR apabila dikonsumsi dalam porsi besar atau berulang dalam satu hari, sehingga informasi distribusi energi ini penting sebagai pertimbangan dalam pola konsumsi yang seimbang.

3.16 Implikasi Gizi Makanan Tradisional Pagar Alam

Secara keseluruhan, analisis AKG dan AMDR menunjukkan bahwa makanan tradisional Pagar Alam yang diteliti umumnya memberikan kontribusi karbohidrat yang lebih besar dibandingkan protein dan lemak, mencerminkan karakter umum makanan tradisional berbahan dasar beras ketan, buah, dan umbi. Kontribusi makronutrien ini bersifat per porsi,

sehingga konsumsi dalam satu kali makan tidak secara otomatis memenuhi kebutuhan gizi harian secara menyeluruh, melainkan merupakan bagian dari pola konsumsi yang lebih luas. Temuan ini juga relevan dengan perhatian terhadap peningkatan prevalensi obesitas dewasa di Indonesia (Kementerian Kesehatan Republik Indonesia, 2025), karena makanan dengan kontribusi karbohidrat dan lemak tinggi seperti Lemang berpotensi menyumbang asupan energi berlebih apabila dikonsumsi dalam frekuensi tinggi tanpa diimbangi aktivitas fisik yang memadai. Informasi kandungan gizi berbasis prediksi ini dapat menjadi referensi awal bagi masyarakat Pagar Alam dan pemangku kebijakan bidang gizi dalam memahami profil gizi makanan tradisional daerah, sekaligus melengkapi basis data komposisi pangan lokal yang masih terbatas (Refi *et al.*, 2024).

4 KESIMPULAN

Penelitian ini berhasil membangun model prediksi kandungan karbohidrat, protein, dan lemak pada enam jenis makanan tradisional Pagar Alam menggunakan algoritma *K-Nearest Neighbor* (KNN) Regression berbasis fitur agregat resep, yaitu jumlah bahan, total berat bahan, dan jumlah bahan per metode pengolahan. Berdasarkan evaluasi menggunakan skema *Leave-One-Out Cross-Validation* (LOO-CV), nilai $K = 2$ memberikan performa terbaik dengan RMSE 18,2874 gram, MAE 11,1629 gram, dan MSE 334,4308 gram², mengungguli $K = 1$ (RMSE 23,0510 gram) dan $K = 3$ (RMSE 21,9307 gram). *Error* prediksi yang relatif besar pada beberapa makanan, terutama Lemang dan Wajik, disebabkan oleh karakteristik algoritma KNN yang melakukan estimasi berdasarkan rata-rata tetangga terdekat, serta keterbatasan jumlah dan variasi sampel yang menyebabkan beberapa makanan memiliki tetangga terdekat yang identik dalam ruang fitur ternormalisasi. Fitur berbasis pola resep terbukti belum optimal dalam menangkap sinyal gizi dengan baik (sebagaimana terlihat pada prediksi identik untuk Gunjing dan Lempuk).

Analisis kontribusi terhadap Angka Kecukupan Gizi (AKG) menunjukkan bahwa keenam makanan tradisional Pagar Alam memberikan kontribusi karbohidrat per porsi antara 4,65% (Gunjing) hingga 44,20% (Lemang) dari AKG harian, dengan kontribusi protein dan lemak yang juga bervariasi sesuai komposisi bahan masing-masing makanan. Analisis distribusi energi terhadap *Acceptable Macronutrient Distribution Range* (AMDR) lebih lanjut menunjukkan bahwa beberapa makanan berbahan dasar beras ketan berpotensi memiliki proporsi energi karbohidrat yang melampaui batas atas rentang anjuran apabila dikonsumsi dalam porsi besar. Dengan mempertimbangkan jumlah sampel yang masih terbatas, penelitian ini disimpulkan sebagai studi awal (*proof of concept*) yang menunjukkan potensi pendekatan KNN Regression berbasis fitur resep sebagai alat bantu estimasi gizi makanan tradisional. Namun, untuk penelitian selanjutnya, disarankan untuk merombak desain fitur (misalnya menggunakan persentase atau proporsi jenis bahan utama alih-alih jumlah bahan), atau mempertimbangkan pengujian algoritma selain KNN yang mungkin lebih andal untuk menangani keterbatasan representasi agregat pada resep.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada Dinas Kebudayaan Kota Pagar Alam serta para pelaku usaha kuliner tradisional yang telah memberikan izin dan informasi selama proses observasi dan wawancara, serta kepada Program Studi Teknik Informatika Institut Teknologi Pagar Alam atas dukungan fasilitas dalam penelitian ini.

DAFTAR PUSTAKA

Ardiansyah Sembiring, M., Wulandari Sembiring, F., Informasi, S., & Royal Kisaran, S. (2024). Analisa Kinerja Model Regresi Dalam Machine Learning Untuk Memprediksi

- Harga Beras. *Journal of Information Systems and Informatics Engineering*, 8(1), 144–152. <https://doi.org/10.35145/joisie.v8i1.3902>
- Davies, T., Louie, J. C., Scapin, T., Pettigrew, S., Wu, J. H. Y., Marklund, M., & Coyle, D. H. (2021). An Innovative Machine Learning Approach to Predict the Dietary Fiber Content of Packaged Foods. *Nutrients*, 13(9). <https://doi.org/10.3390/nu13093195>
- Delgado, A., Issaoui, M., Vieira, M. C., Saraiva de Carvalho, I., & Fardet, A. (2021). Food composition databases: Does it matter to human health? *Nutrients*, 13(8). <https://doi.org/10.3390/nu13082816>
- Depiyana, T., Kusumawati, D., & Ma'rifah, B. (2024). Analisis Kandungan Gizi Dan Organoleptik Krakers Substitusi Tepung Ubi Jalar Ungu Dan Tepung Tempe Sebagai Alternatif Pmt Balita Gizi Kurang. *Jurnal Teknologi Pangan dan Gizi*, 23(1), 8–17. <https://doi.org/10.33508/jtpg.v23i1.5293>
- Dodi, Anggraini, D., Ulandari, D., Chairunisa, D., Agustina, R., Ichsanuddin, A. A., Lian Fernandes, O., Resti, Indah, R., Arzila, R., Disniarti, & Agus, M. M. (2025). Pelestarian Kuliner Khas Pagar Alam Melalui Pengenalan dan Pelatihan Pembuatan Ikan Masak Kuning. *Madaniya*, 6. <https://madaniya.biz.id/journals/contents/article/view/1222>
- Gilal, N. U., Al-Thelaya, K., Al-Saeed, J. K., Abdallah, M., Schneider, J., She, J., Awan, J. H., & Agus, M. (2024). Evaluating Machine Learning technologies for food computing from a dataset perspective. *Multimedia Tools and Applications*, 83(11), 32041–32068. <https://doi.org/10.1007/s11042-023-16513-4>
- Hidayat, T., Ridwan, M., Iqbal, M. F., Sukisno, S., Rizky, R., & Manongga, W. E. (2025). Determining Toddlers' Nutritional Status with Machine Learning Classification Analysis Approach. *MATRIK: Jurnal Manajemen, Teknik Informatika dan Rekayasa Komputer*, 24(2), 235–246. <https://doi.org/10.30812/matrik.v24i2.4092>
- Ilham, B., Saori, S., Nurul Anwar, R., Fitriani, & Yuga Pangestu, M. (2021). Analisis Pengendalian Mutu Di Bidang Industri Makanan (Studi Kasus: UMKM Mochi Kaswari Lampion Kota Sukabumi). *Jurnal Inovasi Penelitian*.
- Kementerian Kesehatan Republik Indonesia. (2025). *Keputusan Menteri Kesehatan Republik Indonesia Nomor HK.01.07/MENKES/509/2025 tentang Pedoman Nasional Pelayanan Klinis Tata Laksana Obesitas Dewasa*.
- Kirk, D., Kok, E., Tufano, M., Tekinerdogan, B., Feskens, E. J. M., & Camps, G. (2022). Machine Learning in Nutrition Research. *Advances in Nutrition*, 13(6), 2573–2589. <https://doi.org/10.1093/advances/nmac103>
- Kluyver, T., & Ragan-Kelley. (2022). *Jupyter Notebook Documentation*. Project Jupyter. <https://jupyter.org>
- Lee, E., Choi, J., Ahn, A., Oh, E., Kweon, H., & Cho, D. (2015). Acceptable macronutrient distribution ranges and hypertension. *Clinical and Experimental Hypertension*, 37(6), 463–467. <https://doi.org/10.3109/10641963.2015.1013116>
- Mahmud, M., & Hermana. (2020). *Tabel Komposisi Pangan Indonesia* (Revisi Cetakan 1). Kementerian Kesehatan RI.
- Muthukumar, K. A., Gupta, S., & Saikia, D. (2024). Leveraging Machine Learning techniques to analyze nutritional content in processed foods. *Discover Food*. <https://doi.org/10.1007/s44187-024-00253-x>

- National Academies of Sciences, Engineering, and Medicine. (2024). *Rethinking the Acceptable Macronutrient Distribution Range for the 21st Century: A Letter Report*. The National Academies Press. <https://doi.org/10.17226/27957>
- Nurfauzan, D., & Fatimah, T. (2022). Implementasi Algoritma K-Nearest Neighbors Regression Dalam Memprediksi Harga Saham. In *Seminar Nasional Mahasiswa Fakultas Teknologi Informasi (SENAFTI) Jakarta-Indonesia*. <https://senafiti.budiluhur.ac.id/index.php/>
- Ozturk Kiyak, E., Ghasemkhani, B., & Birant, D. (2023). High-Level K-Nearest Neighbors (HLKNN): A Supervised Machine Learning Model for Classification Analysis. *Electronics (Switzerland)*, 12(18). <https://doi.org/10.3390/electronics12183828>
- Pangestu, B., Tsaqib, M. A., Farizi, A., Pratiwi, R. L., & Widanengsih, E. (2025). Analisis Nilai Gizi Makanan Berbasis Machine Learning Pendekatan Unsupervised untuk Penentuan Status Gizi Sehat. *Jurnal Komputer, Informasi dan Teknologi*, 5(2), 1–10.
- Pratama, M. A. E., Hendra, S., Ngemba, H. R., Nur, R., Azhar, R., & Laila, R. (2024). Comparison of Machine Learning Algorithms for Predicting Stunting Prevalence in Indonesia. *Jurnal Sisfokom (Sistem Informasi dan Komputer)*, 13(2), 200–209. <https://doi.org/10.32736/sisfokom.v13i2.2097>
- Pratama, M., & Riyadi, I. (2024). Identifikasi Nilai Gizi Makro Lemang Berdasarkan Tabel Komposisi Pangan Indonesia (TKPI) Dan Analisis Proksimat. *Jurnal Teknologi Pangan dan Gizi*, 23(2), 89–95. <https://doi.org/10.33508/jtpg.v23i2.4885>
- Peraturan Menteri Kesehatan Republik Indonesia Nomor 28 Tahun 2019 tentang Angka Kecukupan Gizi yang Dianjurkan untuk Masyarakat Indonesia. (2019). Kementerian Kesehatan Republik Indonesia.
- Rashidmayvan, M., Mansoori, A., Derakhshan-Nezhad, E., Tanbakuchi, D., Sangin, F., Mohammadi-Bajgiran, M., Abedsaeidi, M., Ghazizadeh, S., Sarabi, M. M. T., Rezaee, A., Ferns, G., Esmaily, H., & Ghayour-Mobarhan, M. (2025). Nutritional intake of micronutrient and macronutrient and type 2 diabetes: Machine Learning schemes. *Journal of Health, Population and Nutrition*, 44(1). <https://doi.org/10.1186/S41043-024-00712-2>
- Refi, A. O., Febrianti, F., Jannah, H., Fatdillah, H., Endra, P. Y., Asnadeli, & Al-Haqki. (2024). Profil Gizi Dan Keberlanjutan Dadiah: Makanan Lokal Tradisional dari Kabupaten Agam, Sumatera Barat. *Stock Peternakan*, 6, 133–139.
- Setiawan, S. C. E., Yuliantara, A., & Murti, P. D. B. (2024). Pangan fungsional dari bahan pangan tradisional. *Agrointek: Jurnal Teknologi Industri Pertanian*, 18(3), 552–560. <https://doi.org/10.21107/agrointek.v18i3.15464>
- Sintiya, E. S., Amanda, S. R., Bella Vista, C., & Nugroho, P. A. (2025). Implementasi Machine Learning dalam Sistem Prediksi dan Rekomendasi Program Diet Terintegrasi LLM. *Jurnal Nasional Teknologi dan Sistem Informasi*, 11(2), 144–151. <https://doi.org/10.25077/teknosi.v11i2.2025.144-151>
- Su, Y., Wang, Y., He, J., Wang, H., Xian, A., Jiang, H., Lu, W., Zhou, W., & Li, L. (2024). Development and validation of machine-learning models of diet management for hyperphenylalaninemia: a multicenter retrospective study. *BMC Medicine*, 22(1). <https://doi.org/10.1186/s12916-024-03602-w>
- VanderPlas, J. (2017). *Python Data Science Handbook* (1st ed.). O'Reilly Media.

- Wahyuni, I. P. T., & Fairuz, G. A. M. (2025). Pengaruh Teknik Pengolahan Makanan terhadap Kandungan Nutrisi dalam Masakan Tradisional Gudeg. *Jurnal Pengabdian Masyarakat dan Riset Pendidikan*, 3(4), 1977–1982. <https://doi.org/10.31004/jerkin.v3i4.688>
- Wijayanti, D., Hermawan, A., & Avianto, D. (2024). Penerapan Algoritma K-Nearest Neighbor Untuk Deteksi Dini Status Gizi Pasien Dewasa. *Jurnal Sistem dan Teknologi Informasi*, 10.
- Yang, C. (2025). Predicting Nutrient Density in Foods Using Machine Learning Models: A Comparative Study. 64–69. <https://doi.org/10.5220/0013487500004619>