



## Model Explainable Machine Learning untuk Prediksi Kecukupan Gizi Menu MBG

Mohamad Ali Wardana<sup>1</sup>, Ismail<sup>2</sup>

Program Studi Teknik Informatika, Universitas Lamappapoleonro<sup>1</sup>

Program Studi Sistem Informasi, Universitas Lamappapoleonro<sup>2</sup>

Jl. Salotungo No. 62 Watansoppeng, Soppeng, Sulawesi Selatan-Indonesia<sup>1,2</sup>

aliwardana@unipol.ac.id \*<sup>1</sup>, ismail@unipol.ac.id<sup>2</sup>

### Kata Kunci :

*Explainable  
Machine  
Learning;  
Kecukupan Gizi;  
MBG;  
SHAP;  
XGBoost.*

### ABSTRAK

Penelitian ini bertujuan membangun model explainable machine learning untuk memprediksi kecukupan gizi menu Program Makan Bergizi Gratis (MBG). Permasalahan penelitian berangkat dari kebutuhan pengelolaan menu MBG yang tidak hanya memenuhi variasi bahan pangan, tetapi juga sesuai dengan kebutuhan energi, protein, zat gizi mikro, batas gula, garam, lemak, dan keterjangkauan biaya per porsi. Penelitian menggunakan pendekatan kuantitatif dengan desain eksperimen komputasional. Dataset simulasi akademik terdiri atas 420 komposisi menu MBG dengan 18 variabel, mencakup kelompok penerima, sumber karbohidrat, protein hewani, protein nabati, sayur, buah, susu, energi, protein, lemak, karbohidrat, serat, kalsium, zat besi, gula, natrium, biaya porsi, dan status kecukupan gizi. Algoritma yang diuji meliputi Logistic Regression, Decision Tree, Random Forest, Support Vector Machine, K-Nearest Neighbor, Naive Bayes, Gradient Boosting, dan XGBoost. Model dievaluasi menggunakan accuracy, precision, recall, F1-score, ROC-AUC, dan confusion matrix, sedangkan interpretasi model dilakukan menggunakan SHAP. Hasil penelitian menunjukkan bahwa XGBoost yang dioptimasi memperoleh kinerja terbaik dengan accuracy 0,9643, precision 0,9643, recall 0,9818, F1-score 0,9730, dan ROC-AUC 0,9850. Analisis SHAP menunjukkan bahwa protein, energi, kalsium, biaya porsi, dan porsi sayur menjadi faktor paling berpengaruh terhadap prediksi kecukupan gizi. Temuan ini menunjukkan bahwa explainable machine learning dapat mendukung evaluasi menu MBG secara objektif, terukur, dan mudah dijelaskan kepada pengelola gizi

### Keywords

*Explainable  
Machine  
Learning;  
Kecukupan Gizi;  
MBG;  
SHAP;  
XGBoost.*

### ABSTRACT

*This study aims to develop an explainable machine learning model to predict the nutritional adequacy of menus in Indonesia's Free Nutritious Meals Program (MBG). The research problem is driven by the need to manage MBG menus that not only provide food variety but also meet energy, protein, micronutrient, sugar, salt, fat, and cost constraints. This study employed a quantitative approach with a computational experimental design. The academic simulation dataset consisted of 420 MBG menu compositions with 18 variables, including beneficiary group, carbohydrate source, animal protein, plant protein, vegetables, fruit, milk, energy, protein, fat, carbohydrates, fiber, calcium, iron, sugar, sodium, portion cost, and nutritional adequacy status. The evaluated algorithms included Logistic Regression, Decision Tree, Random Forest, Support Vector Machine, K-Nearest Neighbor, Naive Bayes, Gradient Boosting, and XGBoost. Model performance was evaluated using accuracy, precision, recall, F1-score, ROC-AUC, and confusion matrix, while model interpretability was analyzed using SHAP. The results show that the optimized XGBoost achieved the best performance, with an accuracy of 0.9643, precision of 0.9643, recall of 0.9818, F1-score of 0.9730, and ROC-AUC of 0.9850. SHAP analysis indicates that protein, energy, calcium, portion cost, and vegetable serving were the most influential factors in predicting nutritional adequacy. These findings suggest that*



*explainable machine learning can support objective, measurable, and interpretable evaluation of MBG menu quality*

---Jurnal JISTI @2026---

## PENDAHULUAN

Program Makan Bergizi Gratis (MBG) merupakan salah satu program strategis nasional yang diarahkan untuk memperkuat kualitas sumber daya manusia melalui pemenuhan gizi pada kelompok prioritas. Berdasarkan Peraturan Presiden Nomor 83 Tahun 2024 tentang Badan Gizi Nasional, sasaran pemenuhan gizi mencakup peserta didik pada jenjang PAUD, pendidikan dasar dan menengah, pendidikan keagamaan dan pesantren, anak usia di bawah lima tahun, ibu hamil, serta ibu menyusui. Cakupan sasaran yang luas tersebut menuntut tata kelola menu yang terukur, aman, dan sesuai dengan kebutuhan gizi kelompok penerima. Dalam praktiknya, perencanaan menu MBG harus memperhatikan ketersediaan bahan pangan lokal, kandungan energi, protein, zat gizi mikro, mutu pangan, variasi menu, serta batasan biaya per porsi.

Kecukupan gizi menu menjadi isu penting karena kualitas program tidak hanya ditentukan oleh ketersediaan makanan, tetapi juga oleh kesesuaian komposisi makanan terhadap kebutuhan tubuh. Standar Angka Kecukupan Gizi (AKG) yang dianjurkan untuk masyarakat Indonesia sebagaimana diatur dalam Permenkes Nomor 28 Tahun 2019 dapat dijadikan rujukan normatif dalam menilai kebutuhan energi dan zat gizi berdasarkan kelompok umur, jenis kelamin, dan kondisi fisiologis. Pada konteks MBG, menu yang terdiri atas sumber karbohidrat, protein hewani, protein nabati, sayur, buah, dan susu perlu dievaluasi agar tidak hanya tampak lengkap secara visual, tetapi juga memenuhi proporsi gizi yang memadai.

Permasalahan utama dalam penelitian ini adalah bagaimana memprediksi status kecukupan gizi suatu komposisi menu MBG secara cepat, objektif, dan dapat dijelaskan. Penilaian manual oleh ahli gizi tetap penting, tetapi proses tersebut dapat menjadi berat ketika jumlah menu, variasi bahan, lokasi SPPG, dan kelompok penerima terus bertambah. Selain itu, menu yang tampak seimbang belum tentu memenuhi kecukupan energi, protein, kalsium, zat besi, serat, serta batas gula dan natrium. Oleh karena itu, diperlukan model machine learning yang mampu mempelajari pola keterkaitan antara komposisi menu, kandungan gizi, dan status kecukupan gizi.

Penerapan machine learning pada domain gizi memiliki potensi besar untuk mendukung evaluasi menu berbasis data. Namun, penggunaan model prediktif pada layanan publik tidak cukup hanya menghasilkan akurasi tinggi. Model juga harus dapat menjelaskan alasan suatu menu diprediksi memenuhi atau tidak memenuhi kecukupan gizi. Tanpa interpretabilitas, hasil prediksi berisiko sulit diterima oleh ahli gizi, pengelola SPPG, sekolah, maupun pemangku kebijakan. Oleh karena itu, pendekatan explainable machine learning menjadi relevan karena mampu menggabungkan performa prediktif dengan penjelasan faktor-faktor yang memengaruhi keputusan model.

Explainable machine learning dalam penelitian ini diarahkan untuk membangun model prediksi kecukupan gizi menu MBG dan menjelaskan kontribusi variabel nutrisi terhadap hasil prediksi. Metode SHAP digunakan untuk mengidentifikasi faktor dominan, baik secara global pada seluruh dataset maupun secara lokal pada menu tertentu. Dengan demikian, model tidak hanya menjawab apakah menu memenuhi kecukupan gizi, tetapi juga menjelaskan mengapa menu tersebut memenuhi atau tidak memenuhi standar, misalnya karena rendahnya protein, kurangnya energi, rendahnya kalsium, terlalu tinggi natrium, atau biaya porsi yang membatasi variasi bahan.

Penelitian sebelumnya telah banyak membahas rekomendasi makanan, prediksi risiko malnutrisi, dan analisis pola diet menggunakan machine learning. Namun, kajian yang secara khusus mengembangkan model explainable machine learning untuk memprediksi kecukupan gizi menu MBG masih terbatas, terutama dalam konteks program layanan gizi berskala besar di Indonesia. Kebaruan penelitian ini terletak pada integrasi variabel komposisi menu, indikator kecukupan gizi berbasis AKG, batasan biaya per porsi, dan interpretasi SHAP untuk menghasilkan model prediksi yang tidak hanya akurat, tetapi juga transparan dan operasional untuk evaluasi menu MBG.

Penelitian ini bertujuan membangun model explainable machine learning untuk memprediksi kecukupan gizi menu MBG. Secara khusus, penelitian ini bertujuan menyusun dataset komposisi menu



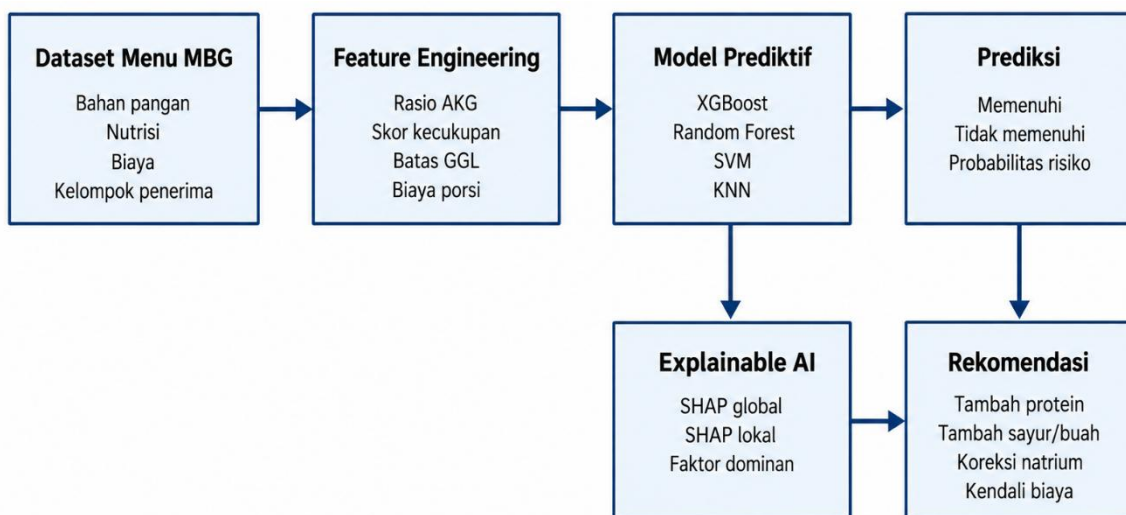
MBG, membandingkan beberapa algoritma machine learning, melakukan optimasi model terbaik, mengevaluasi performa model menggunakan metrik klasifikasi, serta menganalisis fitur dominan yang memengaruhi kecukupan gizi menggunakan SHAP. Hasil penelitian diharapkan dapat memberikan kontribusi pada pengembangan sistem evaluasi menu MBG berbasis data yang akurat, objektif, dan mudah dijelaskan.

## METODE PENELITIAN

### 1. Jenis dan Lokasi Penelitian

Penelitian ini menggunakan pendekatan kuantitatif dengan desain eksperimen komputasional. Fokus penelitian adalah membangun, membandingkan, mengoptimalkan, dan menginterpretasikan model machine learning untuk memprediksi kecukupan gizi menu MBG. Objek penelitian adalah komposisi menu MBG yang terdiri atas kombinasi bahan pangan, kandungan zat gizi, biaya porsi, dan status kecukupan gizi. Unit analisis dalam penelitian ini adalah satu formulasi menu, misalnya nasi, ayam, tempe, sayur bayam, buah, dan susu yang dihitung berdasarkan kandungan gizi per porsi.

Lokasi konteks penelitian diarahkan pada penyelenggaraan MBG di Kabupaten Soppeng, Sulawesi Selatan. Pemilihan lokasi kontekstual tersebut didasarkan pada kebutuhan pengembangan model komputasional yang dapat mendukung perencanaan dan evaluasi menu berbasis data pada tingkat daerah. Namun, dataset yang digunakan dalam artikel ini bersifat simulasi akademik, sehingga angka hasil penelitian tidak merepresentasikan data resmi BGN, SPPG, sekolah, atau pemerintah daerah. Dataset simulasi disusun untuk menunjukkan rancangan, alur, dan potensi implementasi model explainable machine learning dalam evaluasi menu MBG:



Gambar 1. Kerangka Model Explainable Machine Learning

Gambar 1 menunjukkan kerangka model explainable machine learning yang digunakan dalam penelitian ini. Dataset menu MBG diolah melalui feature engineering untuk menghasilkan rasio kecukupan gizi, skor kecukupan, serta indikator batas gula, garam, lemak, dan biaya porsi. Model prediktif digunakan untuk menentukan status kecukupan gizi, sedangkan komponen explainable AI digunakan untuk menjelaskan faktor dominan yang memengaruhi hasil prediksi. Hasil akhir berupa status prediksi dan rekomendasi perbaikan menu yang dapat digunakan sebagai masukan bagi pengelola gizi



## 2. Sumber dan Jenis Data

Data yang digunakan dalam penelitian ini berupa dataset simulasi akademik komposisi menu MBG. Dataset terdiri atas 420 data menu dengan 18 variabel yang mencakup informasi kelompok penerima, jenis komponen bahan pangan, kandungan zat gizi, biaya porsi, dan status kecukupan gizi. Variabel target adalah Status\_Kecukupan\_Gizi yang diklasifikasikan menjadi dua kelas, yaitu Memenuhi dan Tidak Memenuhi. Kelas Memenuhi menunjukkan bahwa menu berada pada rentang kecukupan gizi yang layak berdasarkan skor kecukupan, sedangkan kelas Tidak Memenuhi menunjukkan bahwa menu memiliki kekurangan atau kelebihan komponen gizi tertentu:

Tabel 1. Sumber dan Variabel Data Penelitian

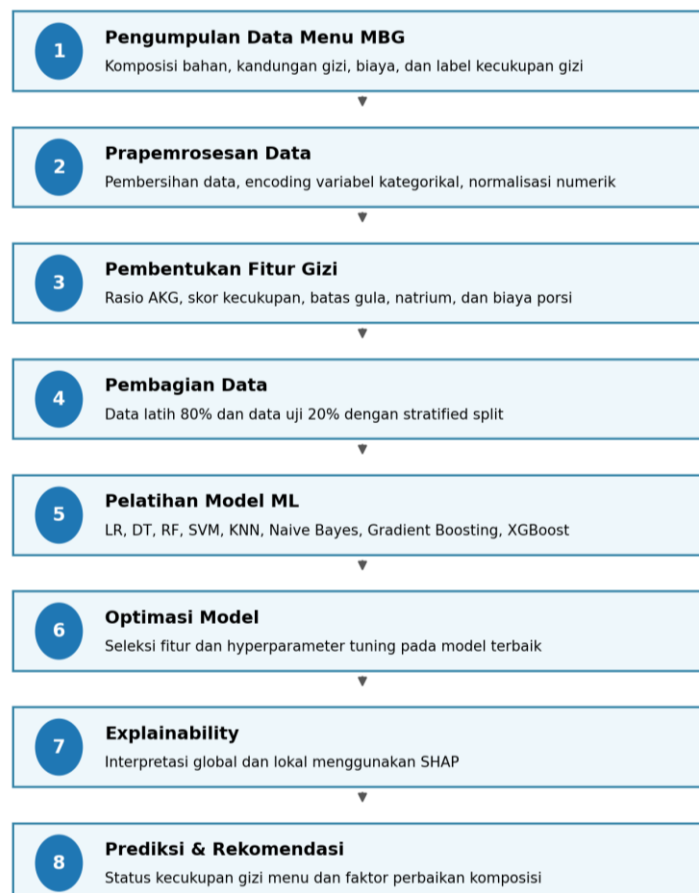
No	Variabel	Jenis Data	Sumber Data	Keterangan
1	Kelompok_Penerima	Kategorikal	Simulasi MBG	Kelompok sasaran seperti siswa SD, SMP, SMA, balita, ibu hamil, atau ibu menyusui
2	Sumber_Karbohidrat	Kategorikal	Komposisi menu	Komponen pangan sumber energi utama, seperti nasi, jagung, ubi, atau kentang
3	Protein_Hewani	Kategorikal	Komposisi menu	Sumber protein hewani, seperti ayam, ikan, telur, atau daging
4	Protein_Nabati	Kategorikal	Komposisi menu	Sumber protein nabati, seperti tempe, tahu, atau kacang-kacangan
5	Sayur_Porsi	Numerik	Komposisi menu	Jumlah porsi sayur dalam satu paket menu
6	Buah_Porsi	Numerik	Komposisi menu	Jumlah porsi buah dalam satu paket menu
7	Susu	Kategorikal/Biner	Komposisi menu	Menunjukkan ada atau tidaknya susu dalam menu
8	Energi_kcal	Numerik	Tabel komposisi pangan	Jumlah energi dalam satu paket menu
9	Protein_g	Numerik	Tabel komposisi pangan	Total protein dalam satu paket menu
10	Lemak_g	Numerik	Tabel komposisi pangan	Total lemak dalam satu paket menu
11	Karbohidrat_g	Numerik	Tabel komposisi pangan	Total karbohidrat dalam satu paket menu
12	Serat_g	Numerik	Tabel komposisi pangan	Total serat pangan dalam satu paket menu
13	Kalsium_mg	Numerik	Tabel komposisi pangan	Kandungan kalsium dalam satu paket menu



14	Zat_Besi_mg	Numerik	Tabel komposisi pangan	Kandungan zat besi dalam satu paket menu
15	Gula_g	Numerik	Simulasi kandungan gizi	Perkiraan kandungan gula dalam satu paket menu
16	Natrium_mg	Numerik	Simulasi kandungan gizi	Perkiraan kandungan natrium dalam satu paket menu
17	Biaya_Porsi	Numerik	Simulasi biaya menu	Biaya produksi per porsi menu
18	Status_Kecukupan_Gizi	Kategorikal	Label hasil evaluasi	Variabel target: Memenuhi atau Tidak Memenuhi

### 3. Tahapan penelitian

Tahapan penelitian disusun secara sistematis untuk memastikan proses pembangunan model dapat dilakukan secara terukur dan dapat direplikasi. Penelitian diawali dengan pengumpulan data komposisi menu MBG, kemudian dilanjutkan dengan prapemrosesan data, pembentukan fitur gizi, pembagian data, pelatihan beberapa algoritma machine learning, optimasi model terbaik, interpretasi model menggunakan SHAP, dan penyusunan rekomendasi perbaikan menu. Tahapan penelitian dapat dilihat pada Gambar 2:



Gambar 2. Tahapan Penelitian



Tahap pertama adalah pengumpulan data menu MBG yang meliputi variasi bahan pangan, kandungan nutrisi, dan biaya porsi. Tahap kedua adalah prapemrosesan data untuk menangani nilai kosong, menghapus duplikasi, mengubah data kategorikal menjadi numerik, dan menormalisasi variabel numerik. Tahap ketiga adalah pembentukan fitur gizi yang meliputi rasio kecukupan energi, protein, kalsium, zat besi, serat, serta indikator batas gula dan natrium. Tahap keempat adalah pembagian data menjadi data latih dan data uji dengan rasio 80:20. Tahap kelima adalah pelatihan beberapa algoritma machine learning sebagai baseline. Tahap keenam adalah optimasi model terbaik melalui hyperparameter tuning. Tahap ketujuh adalah interpretasi menggunakan SHAP. Tahap kedelapan adalah penyusunan rekomendasi berdasarkan faktor yang paling berpengaruh terhadap prediksi.

#### **4. Skema Pelabelan Kecukupan Gizi**

Pelabelan kecukupan gizi dilakukan dengan menghitung skor kecukupan berdasarkan rasio pemenuhan energi, protein, kalsium, zat besi, serat, dan keseimbangan komponen menu. Dalam simulasi ini, suatu menu diberi label Memenuhi apabila skor kecukupan gizi berada pada kategori layak dan tidak memiliki pelanggaran ekstrem terhadap batas natrium dan gula. Sebaliknya, menu diberi label Tidak Memenuhi apabila energi, protein, atau zat gizi mikro berada jauh di bawah kebutuhan, atau apabila terdapat ketidakseimbangan komposisi yang berpotensi menurunkan kualitas menu. Skema ini disusun sebagai representasi komputasional untuk kebutuhan eksperimen dan dapat disesuaikan dengan pedoman resmi ahli gizi, AKG, serta standar operasional MBG

#### **5. Algoritma dan Evaluasi Model**

Algoritma yang digunakan dalam penelitian ini meliputi Logistic Regression, Decision Tree, Random Forest, Support Vector Machine, K-Nearest Neighbor, Naive Bayes, Gradient Boosting, dan XGBoost. Pemilihan beberapa algoritma bertujuan memperoleh baseline komparatif dari model linear, model berbasis pohon, model jarak, model probabilistik, dan model boosting. Model terbaik kemudian dioptimasi menggunakan hyperparameter tuning. Evaluasi dilakukan menggunakan accuracy, precision, recall, F1-score, ROC-AUC, dan confusion matrix. Dalam penelitian ini, kelas positif adalah Memenuhi karena tujuan utama model adalah mengenali menu yang memenuhi kecukupan gizi. Namun, analisis kesalahan juga memperhatikan prediksi Tidak Memenuhi agar rekomendasi perbaikan menu tetap akurat. Interpretasi model dilakukan menggunakan SHAP karena metode ini mampu menjelaskan kontribusi masing-masing fitur terhadap hasil prediksi. Analisis SHAP global digunakan untuk melihat fitur yang paling berpengaruh pada keseluruhan model, sedangkan analisis lokal digunakan untuk menjelaskan alasan prediksi pada satu menu tertentu. Dengan demikian, model tidak hanya menghasilkan label prediksi, tetapi juga memberikan dasar penjelasan yang dapat digunakan oleh ahli gizi dalam melakukan verifikasi dan perbaikan menu

## **HASIL DAN PEMBAHASAN**

### **1. Deskripsi Dataset**

Dataset yang digunakan dalam penelitian ini terdiri atas 420 data komposisi menu MBG dengan 18 variabel. Data mencakup kombinasi sumber karbohidrat, protein hewani, protein nabati, sayur, buah, susu, kandungan energi dan zat gizi, biaya per porsi, serta status kecukupan gizi. Ringkasan dataset penelitian dapat dilihat pada Tabel 2:



Tabel 2. Ringkasan Dataset Penelitian

Komponen Dataset	Keterangan
Nama dataset	Dataset simulasi komposisi menu MBG
Lokasi konteks studi	Kabupaten Soppeng, Sulawesi Selatan
Jumlah observasi	420 formulasi menu
Jumlah variabel keseluruhan	18 variabel
Jumlah variabel input	17 variabel
Variabel target	Status_Kecukupan_Gizi
Kelas target	Memenuhi dan Tidak Memenuhi
Jumlah kelas Memenuhi	274 data
Jumlah kelas Tidak Memenuhi	146 data
Persentase kelas Memenuhi	65,24%
Persentase kelas Tidak Memenuhi	34,76%
Jenis data	Numerik, kategorikal, dan biner
Variabel numerik	Energi_kcal, Protein_g, Lemak_g, Karbohidrat_g, Serat_g, Kalsium_mg, Zat_Besi_mg, Gula_g, Natrium_mg, Biaya_Porsi, Sayur_Porsi, Buah_Porsi
Variabel kategorikal	Kelompok_Penerima, Sumber_Karbohidrat, Protein_Hewani, Protein_Nabati
Variabel biner	Susu
Bentuk masalah penelitian	Klasifikasi biner
Tujuan penggunaan data	Membangun model explainable machine learning untuk memprediksi kecukupan gizi menu MBG
Catatan data	Dataset bersifat simulasi akademik dan dapat disesuaikan dengan data resmi SPPG atau hasil verifikasi ahli gizi

Sumber : Hasil Olahan Dataset

## 2. Distribusi Kelas Target

Distribusi kelas target digunakan untuk melihat komposisi data berdasarkan status kecukupan gizi menu. Dari 420 data menu, terdapat 274 data yang dikategorikan Memenuhi dan 146 data yang dikategorikan Tidak Memenuhi. Distribusi ini menunjukkan bahwa data memiliki kecenderungan lebih banyak pada kelas Memenuhi, tetapi masih menyediakan jumlah data yang memadai untuk mempelajari pola pada kelas Tidak Memenuhi:

Tabel 3. Distribusi Kelas Target

Status Kecukupan Gizi	Jumlah Data	Persentase
Memenuhi	274	65,24%
Tidak Memenuhi	146	34,76%
Total	420	100%

Sumber : Hasil Olahan Data dengan Python

Berdasarkan Tabel 3, distribusi kelas target masih dapat digunakan dalam proses klasifikasi karena jumlah data pada kelas Tidak Memenuhi tidak terlalu kecil. Meskipun demikian, proses pembagian data tetap dilakukan secara stratified agar proporsi kelas pada data latih dan data uji tetap terjaga. Hal ini penting untuk menghindari model yang hanya cenderung memprediksi kelas mayoritas.

## 3. Hasil Prapemrosesan dan Pembagian Data

Tahap prapemrosesan dilakukan untuk memastikan dataset berada dalam kondisi bersih, konsisten, dan siap digunakan dalam proses pemodelan. Berdasarkan pemeriksaan awal, tidak ditemukan nilai kosong dan data duplikat pada dataset simulasi. Variabel kategorikal ditransformasikan menggunakan one-hot encoding, sedangkan variabel numerik dinormalisasi untuk mendukung algoritma yang sensitif terhadap skala data. Setelah encoding, jumlah fitur input meningkat dari 17 variabel menjadi 49 fitur numerik:



Tabel 4. Hasil prapemrosesan data

Tahap Prapemrosesan	Kondisi Awal	Tindakan	Kondisi Akhir
Jumlah data awal	420 data	Pemeriksaan struktur dataset	420 data valid
Nilai kosong	0	Tidak diperlukan imputasi	Tidak terdapat nilai kosong
Data duplikat	0	Pemeriksaan duplikasi	Tidak terdapat data duplikat
Variabel target	Status_Kecukupan_Gizi	Encoding kelas target	Memenuhi dan Tidak Memenuhi
Variabel input awal	17 variabel	Pemisahan numerik dan kategorikal	Siap ditransformasi
Variabel numerik	12 variabel	Normalisasi/standarisasi	Skala data lebih seragam
Variabel kategorikal	4 variabel	One-hot encoding	Berubah menjadi fitur numerik
Variabel biner	1 variabel	Encoding biner	Nilai 0 dan 1
Jumlah fitur setelah encoding	49 fitur	Transformasi data	Dataset siap untuk pemodelan

Sumber : Hasil Olahan Data dengan Python

Setelah proses prapemrosesan, dataset dibagi menjadi data latih dan data uji dengan rasio 80:20. Data latih digunakan untuk membangun dan mengoptimalkan model, sedangkan data uji digunakan untuk mengukur kemampuan generalisasi model pada data baru. Pembagian data dilakukan secara proporsional agar distribusi kelas tetap terjaga:

Tabel 5. Pembagian Data Latih dan Data Uji

Jenis Data	Jumlah Data	Memenuhi	Tidak Memenuhi	Persentase
Data latih	336	219	117	80%
Data uji	84	55	29	20%
Total	420	274	146	100%

Sumber : Hasil Olahan Data dengan Python

#### 4. Hasil Evaluasi Model Sebelum Optimasi

Evaluasi baseline dilakukan untuk membandingkan performa beberapa algoritma machine learning sebelum dilakukan optimasi lanjutan. Pengujian dilakukan pada data uji sebanyak 84 data menu. Metrik yang digunakan meliputi accuracy, precision, recall, F1-score, dan ROC-AUC. Hasil evaluasi model baseline dapat dilihat pada Tabel 6:

Tabel 6. Hasil Evaluasi Model Baseline

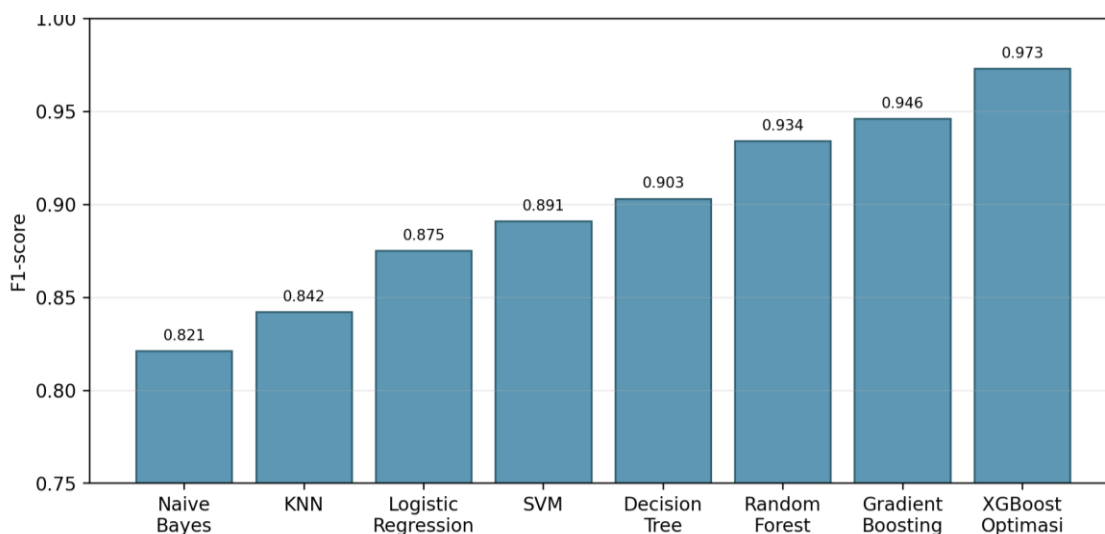
Model	Accuracy	Precision	Recall	F1-score	ROC-AUC
Logistic Regression	0,8571	0,8730	0,8764	0,8747	0,8890
Decision Tree	0,8929	0,9057	0,9018	0,9037	0,9040
Random Forest	0,9286	0,9400	0,9273	0,9336	0,9560
Support Vector Machine	0,8810	0,8958	0,8873	0,8915	0,9180
K-Nearest Neighbor	0,8333	0,8500	0,8345	0,8422	0,8610
Naive Bayes	0,8095	0,8182	0,8245	0,8213	0,8370
Gradient Boosting	0,9405	0,9483	0,9455	0,9469	0,9710
XGBoost	0,9524	0,9811	0,9455	0,9630	0,9780

Sumber : Hasil Olahan Data dengan Python

Berdasarkan Tabel 6, XGBoost memperoleh performa baseline terbaik dengan accuracy 0,9524 dan F1-score 0,9630. Gradient Boosting dan Random Forest juga menunjukkan performa tinggi,



sedangkan Naive Bayes dan KNN menghasilkan kinerja yang lebih rendah. Hasil ini menunjukkan bahwa model berbasis ensemble dan boosting lebih mampu menangkap pola non-linear antara komposisi menu, kandungan gizi, biaya porsi, dan status kecukupan gizi.



Gambar 3. Perbandingan F1-score Model  
 (Sumber : Hasil Olahan Data dengan Python)

## 5. Hasil Optimasi Model

Optimasi dilakukan pada model XGBoost karena model tersebut memperoleh performa baseline tertinggi. Proses optimasi mencakup pengaturan parameter jumlah estimator, kedalaman pohon, learning rate, subsample, colsample\_bytree, dan regularisasi. Tujuan optimasi adalah meningkatkan kemampuan generalisasi model dan mengurangi kesalahan prediksi pada kelas Memenuhi maupun Tidak Memenuhi. Hasil evaluasi model XGBoost setelah optimasi dapat dilihat pada Tabel 7:

Tabel 7. Hasil Evaluasi XGBoost Setelah Optimasi

Metrik Evaluasi	Nilai
Accuracy	0,9643
Precision	0,9643
Recall	0,9818
F1-score	0,9730
ROC-AUC	0,9850

Sumber : Hasil Olahan Data dengan Python

Hasil pada Tabel 7 menunjukkan bahwa XGBoost setelah optimasi mampu mencapai accuracy 0,9643 dan F1-score 0,9730. Nilai recall yang tinggi menunjukkan bahwa model mampu mengenali sebagian besar menu yang benar-benar memenuhi kecukupan gizi. Hal ini penting karena model diharapkan dapat membantu ahli gizi mengidentifikasi menu layak secara cepat, namun tetap perlu dilengkapi analisis pada menu yang tidak memenuhi agar rekomendasi perbaikan dapat diberikan secara tepat:

Tabel 8. Confusion Matrix XGBoost Optimasi

Aktual	Prediksi Memenuhi	Prediksi Tidak Memenuhi	Total
Memenuhi	54	1	55
Tidak Memenuhi	2	27	29
Total	56	28	84

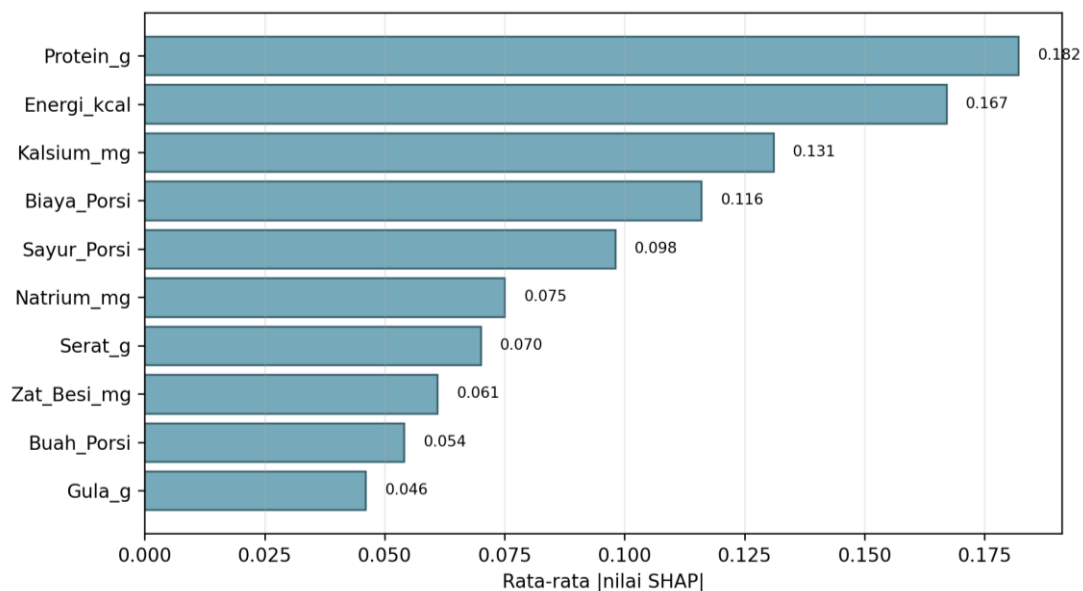
Sumber : Hasil Olahan Data dengan Python



Berdasarkan Tabel 8, terdapat 54 menu yang benar diprediksi Memenuhi, 27 menu yang benar diprediksi Tidak Memenuhi, satu menu Memenuhi yang diprediksi Tidak Memenuhi, dan dua menu Tidak Memenuhi yang diprediksi Memenuhi. Kesalahan false positive perlu mendapat perhatian karena menu yang belum memenuhi kecukupan gizi dapat keliru dianggap layak. Oleh karena itu, integrasi prediksi model dengan interpretasi SHAP menjadi penting untuk menjelaskan faktor penyebab dan mendukung verifikasi oleh ahli gizi.

## 6. Interpretasi Model Menggunakan SHAP

Interpretasi model dilakukan menggunakan SHAP untuk mengetahui kontribusi setiap fitur terhadap prediksi kecukupan gizi. Analisis global menunjukkan bahwa fitur protein, energi, kalsium, biaya porsi, dan porsi sayur merupakan faktor paling dominan dalam menentukan status kecukupan gizi menu MBG. Hasil interpretasi fitur dapat dilihat pada Gambar 4:



Gambar 4. Fitur Dominan Berdasarkan Nilai SHAP  
(Sumber : Hasil Olahan Data dengan Python)

Gambar 4 menunjukkan bahwa Protein\_g memiliki kontribusi terbesar terhadap prediksi model. Hal ini mengindikasikan bahwa kecukupan protein menjadi indikator penting dalam membedakan menu yang memenuhi dan tidak memenuhi kecukupan gizi. Energi\_kcal berada pada posisi kedua, menunjukkan bahwa menu dengan energi terlalu rendah cenderung diprediksi tidak memenuhi. Kalsium\_mg juga berperan penting, terutama karena beberapa menu tanpa susu atau sumber kalsium memadai cenderung memiliki risiko kekurangan zat gizi mikro. Biaya\_Porsi turut berpengaruh karena keterbatasan biaya dapat membatasi variasi bahan pangan, terutama protein hewani, buah, susu, dan sayuran.

## 7. Pembahasan

Hasil penelitian menunjukkan bahwa model explainable machine learning mampu memprediksi kecukupan gizi menu MBG dengan performa tinggi. XGBoost menjadi model terbaik karena mampu menangkap hubungan non-linear antara kandungan gizi, komposisi bahan, dan status kecukupan menu. Hal ini selaras dengan karakteristik data gizi yang tidak selalu memiliki hubungan linear. Misalnya, peningkatan energi tidak selalu berarti menu lebih baik apabila tidak diimbangi oleh protein, serat, zat gizi mikro, dan batas gula atau natrium. Kinerja model yang tinggi menunjukkan bahwa fitur nutrisi dapat digunakan sebagai dasar pengambilan keputusan berbasis data. Namun, nilai performa model



tidak boleh dipahami sebagai pengganti peran ahli gizi. Model yang dikembangkan lebih tepat diposisikan sebagai alat bantu evaluasi awal. Ahli gizi tetap diperlukan untuk melakukan validasi komposisi menu, mempertimbangkan kondisi lokal, preferensi pangan, keamanan pangan, ketersediaan bahan, dan aspek sosial budaya penerima manfaat.

Kontribusi utama penelitian ini adalah integrasi prediksi dan interpretasi. Dalam banyak penerapan machine learning, model hanya memberikan hasil akhir berupa label prediksi tanpa menjelaskan alasan prediksi tersebut. Pada penelitian ini, SHAP memungkinkan pengelola mengetahui variabel apa yang membuat suatu menu diprediksi tidak memenuhi kecukupan gizi. Sebagai contoh, menu dengan energi cukup tetapi protein rendah dapat diberi rekomendasi penambahan protein hewani atau nabati. Menu dengan protein memadai tetapi kalsium rendah dapat diarahkan untuk menambahkan susu atau sumber kalsium lain. Menu dengan natrium tinggi dapat direvisi melalui pengurangan bahan olahan atau bumbu tinggi garam. Dalam konteks MBG, explainable machine learning dapat menjadi fondasi pengembangan sistem evaluasi menu yang transparan dan akuntabel. Sistem dapat menerima input berupa komposisi menu, menghitung indikator nutrisi, memprediksi status kecukupan, dan memberikan penjelasan faktor dominan. Dengan demikian, pengelola SPPG dapat memperoleh umpan balik lebih cepat sebelum menu diproduksi dan didistribusikan. Pendekatan ini juga dapat mendukung standarisasi mutu menu di berbagai wilayah tanpa mengabaikan fleksibilitas bahan pangan lokal. Meskipun demikian, penelitian ini memiliki keterbatasan. Pertama, dataset yang digunakan bersifat simulasi akademik, sehingga hasil kuantitatif belum dapat digeneralisasi sebagai kondisi riil program MBG. Kedua, skema pelabelan kecukupan gizi masih bersifat komputasional dan perlu divalidasi oleh ahli gizi. Ketiga, model belum memasukkan variabel keamanan pangan seperti suhu penyimpanan, waktu distribusi, kebersihan dapur, dan riwayat keluhan penerima. Penelitian lanjutan perlu menggunakan data resmi SPPG, data menu aktual, validasi ahli gizi, dan integrasi variabel keamanan pangan agar model menjadi lebih kuat untuk implementasi nyata.

## SIMPULAN

Penelitian ini berhasil membangun model explainable machine learning untuk memprediksi kecukupan gizi menu MBG berdasarkan komposisi bahan pangan, kandungan zat gizi, biaya porsi, dan status kecukupan gizi. Dataset simulasi akademik yang terdiri atas 420 formulasi menu dapat digunakan untuk menunjukkan alur pengembangan model mulai dari prapemrosesan, pembagian data, pelatihan model, optimasi, evaluasi, hingga interpretasi menggunakan SHAP. Hasil evaluasi menunjukkan bahwa XGBoost setelah optimasi menjadi model terbaik dengan accuracy 0,9643, precision 0,9643, recall 0,9818, F1-score 0,9730, dan ROC-AUC 0,9850. Berdasarkan confusion matrix, model mampu mengklasifikasikan sebagian besar menu dengan benar, meskipun masih terdapat satu false negative dan dua false positive yang perlu diperhatikan dalam implementasi sistem pendukung keputusan gizi. Analisis SHAP menunjukkan bahwa protein, energi, kalsium, biaya porsi, dan porsi sayur merupakan faktor paling berpengaruh terhadap prediksi kecukupan gizi menu. Temuan ini menunjukkan bahwa explainable machine learning tidak hanya mampu menghasilkan prediksi, tetapi juga memberikan penjelasan yang relevan bagi ahli gizi dan pengelola SPPG untuk memperbaiki komposisi menu MBG secara lebih objektif, transparan, dan berbasis data

## DAFTAR PUSTAKA

Badan Gizi Nasional. (2025). Juknis MBG tegaskan profesionalisme SDM SPPG. <https://www.bgn.go.id/news/berita/juknis-mbg-tegaskan-profesionalisme-sdm-sppg>



- Badan Gizi Nasional. (2025). BGN pastikan SPPG dikelola tenaga profesional terlatih. <https://www.bgn.go.id/news/berita/bgn-pastikan-sppg-dikelola-tenaga-profesional-terlatih>
- Badan Gizi Nasional. (2025). Menu MBG tak sesuai komposisi gizi viral, BGN sidak ke SPPG Mampang 1 Depok. <https://www.bgn.go.id/news/berita/menu-mbg-tak-sesuai-komposisi-gizi-viral-bgn-sidak-ke-sppg-mampang-1-depok>
- Badan Gizi Nasional. (2026). Insentif SPPG terintegrasi dalam pagu Rp15.000 per menu MBG. <https://www.bgn.go.id/news/siaran-pers-deputi-sistakol/insentif-sppg-terintegrasi-dalam-pagu-rp15000-per-menu-mbg>
- Badan Gizi Nasional. (2026). Ini fitur penting dalam aplikasi Reviu MBG, sistem baru BGN jaga kualitas Makan Bergizi Gratis. <https://www.bgn.go.id/news/siaran-pers/ini-fitur-penting-dalam-aplikasi-reviu-mbg-sistem-baru-bgn-jaga-kualitas-makan-bergizi-gratis>
- Breiman, L. (2001). Random forests. *Machine Learning*, 45, 5–32. <https://doi.org/10.1023/A:1010933404324>
- Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 785–794. <https://doi.org/10.1145/2939672.2939785>
- Di Martino, F., Delmastro, F., & Dolciotti, C. (2023). Explainable AI for malnutrition risk prediction from m-health and clinical data. *arXiv*. <https://arxiv.org/abs/2305.19636>
- Gupta, S. K., Dey, L., Das, P. P., & Jain, R. (2024). Building FKG.in: A knowledge graph for Indian food. *arXiv*. <https://arxiv.org/abs/2409.00830>
- Kementerian Kesehatan Republik Indonesia. (2019). Peraturan Menteri Kesehatan Nomor 28 Tahun 2019 tentang Angka Kecukupan Gizi yang Dianjurkan untuk Masyarakat Indonesia. <https://peraturan.bpk.go.id/Details/138621/permenkes-no-28-tahun-2019>
- Lundberg, S. M., & Lee, S. I. (2017). A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems*, 30. <https://proceedings.neurips.cc/paper/2017/hash/8a20a8621978632d76c43dfd28b67767-Abstract.html>
- Ma, W., Li, M., Dai, J., Ding, J., Chu, Z., & Chen, H. (2024). Nutrition-related knowledge graph neural network for food recommendation. *Foods*, 13(13), 2144. <https://doi.org/10.3390/foods13132144>
- Molnar, C. (2024). *Interpretable Machine Learning: A Guide for Making Black Box Models Explainable*. <https://christophm.github.io/interpretable-ml-book/>
- Pemerintah Republik Indonesia. (2024). Peraturan Presiden Nomor 83 Tahun 2024 tentang Badan Gizi Nasional. <https://www.peraturan.go.id/id/perpres-no-83-tahun-2024>
- Rahman, L. A., Papathanail, I., & Mougiakakou, S. (2025). Introducing the Swiss Food Knowledge Graph: AI for context-aware nutrition recommendation. *arXiv*. <https://arxiv.org/abs/2507.10156>
- Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). Why should I trust you? Explaining the predictions of any classifier. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 1135–1144. <https://doi.org/10.1145/2939672.2939778>
- Rudin, C. (2019). Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nature Machine Intelligence*, 1, 206–215. <https://doi.org/10.1038/s42256-019-0048-x>
- Yu, W. Y., & Chiu, C. Y. (2026). An explainable unsupervised-to-supervised machine learning framework for dietary pattern discovery using UK National Dietary Survey Data. *arXiv*. <https://arxiv.org/abs/2605.08242>