

Model Ensemble Stacking untuk Klasifikasi Big Data Stunting Berbasis XGBoost dan MLP

^{1*}Khairul Hawani Rambe, ²Frans Mikael Sinaga, ³Leni Anggraini Susanti, ⁴Moh. Erkamim
^{1,3,4}Politeknik Negeri Bali
²Universitas Pelita Harapan
^{1,3,4}Bali, Indonesia
²Medan, Indonesia

^{1*}khairulhawani@pnb.ac.id, ²frans.sinaga@uph.edu, ³lenissnti@pnb.ac.id, ⁴erkamim@pnb.ac.id

*Penulis Korespondensi

Diajukan : 13/12/2025
Diterima : 30/12/2025
Dipublikasi : 20/01/2026

ABSTRAK

Klasifikasi status gizi balita berbasis data berskala besar memerlukan model yang akurat, stabil, dan mampu menangani kompleksitas serta ketidakseimbangan data. Penelitian ini bertujuan mengevaluasi kinerja model *ensemble stacking* berbasis *XGBoost* dan *Multi-Layer Perceptron* (MLP) dalam klasifikasi status gizi balita. Populasi penelitian mencakup seluruh data rekam medis balita periode 2023–2024 yang diperoleh dari rumah sakit, dengan teknik pengambilan sampel menggunakan total sampling terhadap dataset yang tersedia. Sampel penelitian terdiri dari 94.291 observasi yang memuat variabel jenis kelamin, usia, berat badan, tinggi atau panjang badan, nilai *Z-score*, serta label status gizi. Metode penelitian menggunakan pendekatan kuantitatif berbasis *machine learning* dengan tahapan pra-pemrosesan data, pelatihan model, dan evaluasi performa. Pra-pemrosesan meliputi pembersihan data, *encoding* variabel kategorikal, normalisasi fitur numerik, pembagian data latih dan uji dengan rasio 80:20, serta validasi silang 5-fold. Model *stacking* dibangun dengan *XGBoost* sebagai *base learner* dan MLP sebagai *meta learner*. Evaluasi dilakukan menggunakan confusion matrix, akurasi, *precision*, *recall*, *F1-score*, dan ROC-AUC. Hasil eksperimen menunjukkan bahwa model *stacking* mencapai akurasi 99,64% dengan nilai ROC-AUC mendekati 1, serta kesalahan prediksi yang sangat rendah pada kedua kelas. Temuan ini menunjukkan bahwa integrasi algoritma *boosting* dan *neural network* mampu meningkatkan stabilitas dan kemampuan generalisasi model. Dengan demikian, pendekatan *stacking XGBoost-MLP* efektif untuk klasifikasi status gizi balita dan berpotensi diterapkan sebagai sistem pendukung keputusan berbasis big data dalam deteksi dini masalah gizi.

Kata Kunci: Klasifikasi, Mitra Medika, Multi-Layer Perceptron, Stunting, XGBoost

I. PENDAHULUAN

Stunting merupakan permasalahan kesehatan kronis yang berdampak jangka panjang terhadap pertumbuhan fisik, perkembangan kognitif, dan kualitas sumber daya manusia (Cahyani et al., 2024). Kompleksitas faktor penyebab stunting yang meliputi aspek kesehatan, sosial, dan demografis menghasilkan data berskala besar dengan karakteristik heterogeny (Hariguna et al., 2024a; Syahrial et al., 2022). Dalam penelitian ini, dataset yang digunakan terdiri atas 94.291 data observasi dengan tujuh variabel utama yang diperoleh setelah melalui tahap pembersihan data. Jumlah dan karakteristik data tersebut mencerminkan sifat big data, khususnya dari sisi volume dan keragaman, sehingga pendekatan konvensional menjadi kurang optimal. Pemanfaatan big data analytics dan machine learning menjadi solusi yang relevan untuk mengekstraksi pola tersembunyi dan menghasilkan prediksi yang lebih akurat dalam konteks klasifikasi stunting. (Nirsal, S.Kom. et al., n.d.; Samuel Aleksander Mandowen et al., 2025).

Berbagai penelitian sebelumnya menunjukkan bahwa *machine learning* telah banyak

diterapkan dalam bidang kesehatan, termasuk untuk klasifikasi dan prediksi risiko stunting. Algoritma seperti *Decision Tree*, *Random Forest*, *Support Vector Machine*, *Multi-Layer Perceptron* (MLP), dan *XGBoost* menunjukkan performa yang menjanjikan dalam mengolah data kesehatan berdimensi tinggi (Renaningtias & Prihatiningrum, 2025). MLP unggul dalam memodelkan hubungan non-linear antarvariabel (ASHURI et al., 2024; Cahyani et al., 2024), sedangkan *XGBoost* efektif dalam menangani data besar, kompleks, dan tidak seimbang (Le et al., 2022; Noh et al., 2021). Meskipun demikian, sebagian besar penelitian masih mengandalkan model tunggal seperti *XGBoost* atau MLP secara terpisah masih memiliki keterbatasan dalam menjaga stabilitas dan generalisasi performa, terutama ketika dihadapkan pada data kesehatan berskala besar dan tidak seimbang.

Untuk mengatasi keterbatasan tersebut, pendekatan *ensemble learning* dikembangkan untuk mengatasi keterbatasan model tunggal dengan mengombinasikan beberapa algoritma guna meningkatkan performa prediksi (Samuel Aleksander Mandowen et al., 2025; Sudirwo, S, E. et al., 2025). Teknik *stacking*, sebagai salah satu metode ensemble, menggabungkan keluaran beberapa base learner melalui meta-learner dan terbukti mampu meningkatkan akurasi klasifikasi pada berbagai domain, termasuk bidang kesehatan (Yousefi et al., 2024; Zhang et al., 2024). Meskipun demikian, penerapan *stacking* yang secara khusus mengombinasikan *XGBoost* dan MLP untuk klasifikasi stunting berbasis *big data* masih relatif terbatas.

Penelitian sebelumnya yang melibatkan penulis menunjukkan bahwa optimasi *Random Forest* mampu meningkatkan kinerja klasifikasi stunting, namun masih terdapat ruang pengembangan untuk menangani kompleksitas data secara lebih optimal (Hadini Marpaung et al., 2025). Studi lain juga menegaskan bahwa pendekatan *hybrid* dan *ensemble* memiliki potensi besar dalam meningkatkan akurasi klasifikasi stunting, khususnya pada konteks negara berkembang (Hariguna et al., 2024; Yaqoob et al., 2024).

Gap penelitian terletak pada belum optimalnya penggunaan model *stacking* yang mengintegrasikan algoritma boosting dan neural network dalam klasifikasi big data stunting (Cahyani et al., 2024; Hasdyna et al., 2024). Berdasarkan gap tersebut, pertanyaan penelitian dalam studi ini dirumuskan sebagai berikut:

1. Bagaimana performa model *stacking* berbasis XGBoost dan MLP dalam mengklasifikasikan data stunting berskala besar?
2. Apakah model *stacking* XGBoost–MLP mampu meningkatkan akurasi dan stabilitas klasifikasi dibandingkan model tunggal XGBoost dan MLP?

Oleh karena itu, penelitian ini bertujuan mengoptimalkan klasifikasi stunting menggunakan model *stacking* berbasis XGBoost dan *Multi-Layer Perceptron*. Kontribusi utama penelitian ini adalah:

1. Pengembangan model *stacking* yang mengombinasikan algoritma *boosting* (XGBoost) dan *neural network* (MLP) untuk klasifikasi stunting berbasis big data kesehatan.
2. Penerapan machine learning pada dataset rekam medis berskala besar sebagai pendekatan analitik dalam pemantauan status gizi.
3. Penyediaan model berperforma tinggi yang berpotensi diimplementasikan sebagai sistem pendukung keputusan untuk deteksi dini masalah gizi balita.

Hipotesis penelitian menyatakan bahwa model *stacking* XGBoost–MLP mampu meningkatkan akurasi dan stabilitas klasifikasi dibandingkan model tunggal, sehingga berpotensi mendukung pengambilan keputusan berbasis data dalam program intervensi kesehatan anak.

II. STUDI LITERATUR

1. Penelitian Terdahulu

Berbagai penelitian menunjukkan bahwa *machine learning* efektif dalam klasifikasi dan prediksi stunting pada anak (Bitew et al., 2022). Di Indonesia, algoritma MLP dan *Random Forest* telah berhasil diterapkan untuk klasifikasi stunting dengan hasil yang menjanjikan, meskipun masih memiliki keterbatasan dalam generalisasi pada data kompleks (Hadini Marpaung et al., 2025). Penelitian sebelumnya membuktikan bahwa XGBoost memiliki akurasi tinggi dalam prediksi stunting dan unggul dalam menangani data kesehatan tabular (Renaningtias &

Prihatiningrum, 2025). Namun, sebagian besar studi masih menggunakan model tunggal dan belum mengintegrasikan pendekatan ensemble yang lebih robust untuk meningkatkan stabilitas dan kemampuan generalisasi. Selain itu, evaluasi pada banyak penelitian masih terbatas pada dataset kecil hingga menengah, sementara karakteristik data stunting bersifat heterogen, non-linear, dan tidak seimbang. Kondisi ini menunjukkan perlunya pendekatan stacking yang mengombinasikan XGBoost dan MLP guna mengoptimalkan akurasi serta stabilitas klasifikasi pada skala big data.

Tabel 1. Komparasi Penelitian Terdahulu Klasifikasi Stunting

Peneliti	Metode	Skala Data	Keterbatasan
Bitew et al. (2022)	RF, SVM	Kecil	Generalisasi rendah
Hadini Marpaung et al. (2025)	Random Forest	Besar	Tanpa boosting
Cahyani et al. (2024)	MLP	Nasional	Sensitif parameter
Renaningtias et al. (2025)	XGBoost	Besar	Tidak stacking
Anhar & Soewito (2025)	XGBoost	Menengah	Belum ensemble
Penelitian ini	Stacking XGBoost-MLP	Big Data	Mengatasi keterbatasan sebelumnya

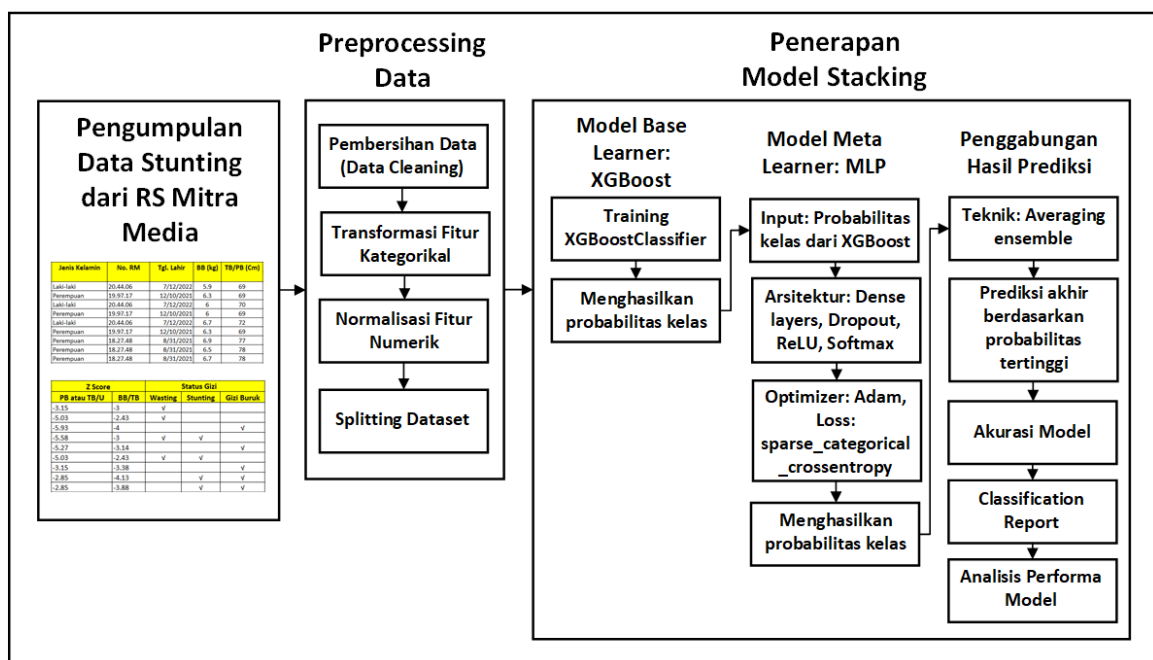
Sumber tabel : Hasil sintesis literatur

2. XGBoost, MLP, dan Ensemble Learning

XGBoost efektif dalam menangani data tabular besar dan tidak seimbang melalui mekanisme boosting dan regularisasi (Renaningtias & Prihatiningrum, 2025), sedangkan MLP unggul dalam memodelkan hubungan non-linear kompleks (Cahyani et al., 2024). Penggabungan XGBoost dan MLP melalui stacking didasarkan pada prinsip diversitas model, di mana keunggulan model berbasis pohon dan neural network saling melengkapi untuk meningkatkan akurasi, stabilitas, dan kemampuan generalisasi pada big data stunting.

III. METODE

Penelitian ini menggunakan pendekatan kuantitatif berbasis *machine learning* untuk mengoptimalkan klasifikasi status gizi balita melalui model stacking ensemble yang mengombinasikan XGBoost sebagai base learner dan MLP sebagai *meta learner* (Cahyani et al., 2024b; Vipin et al., 2023). Alur penelitian meliputi pengumpulan data, pra-pemrosesan, pembangunan model, serta evaluasi kinerja, sebagaimana ditunjukkan pada Gambar 1.



Gambar 1. Pendekatan Pemecahan Masalah
Sumber gambar : Hasil perancangan metode penelitian (2026)

1. Pengumpulan Data

Data yang digunakan merupakan data status gizi balita yang diperoleh dari Rumah Sakit Mitra Medika Tanjung Mulia pada periode 2023–2024. Dataset terdiri atas atribut jenis kelamin, usia, berat badan (BB), tinggi atau panjang badan (TB/PB), nilai *Z-score*, serta label status gizi yang mencakup kategori stunting, *wasting*, dan gizi buruk. Setelah melalui tahap pembersihan data, dataset akhir yang digunakan dalam penelitian ini berjumlah 94.291 data observasi dengan 7 atribut, yang memenuhi karakteristik data berskala besar (*big data*) pada konteks analitik kesehatan.

Distribusi kelas pada dataset menunjukkan ketidakseimbangan antar kategori status gizi, di mana kelas stunting memiliki proporsi data terbesar dibandingkan *wasting* dan gizi buruk. Kondisi ini mencerminkan karakteristik nyata data kesehatan dan menjadi tantangan utama dalam proses klasifikasi. Variabel penelitian terdiri atas:

- 1) Variabel independen (fitur): jenis kelamin, usia, berat badan (BB), tinggi/panjang badan (TB/PB), serta nilai *Z-score* (PB/TB dan BB/TB).
- 2) Variabel dependen (label): status gizi balita yang diklasifikasikan ke dalam kategori stunting, *wasting*, dan gizi buruk.

Dataset yang digunakan merupakan klasifikasi biner dengan dua kelas (0 dan 1). Berdasarkan hasil eksplorasi data, distribusi kelas menunjukkan adanya ketidakseimbangan proporsi antar kelas, sehingga pendekatan evaluasi model tidak hanya mengandalkan akurasi tetapi juga metrik lain seperti precision, recall, dan ROC-AUC.

2. Pra-pemrosesan Data

Tahap pra-pemrosesan bertujuan meningkatkan kualitas data sebelum pemodelan. Langkah-langkah yang dilakukan meliputi:

- 1) Pembersihan data, yaitu penghapusan data duplikat dan penanganan nilai kosong.
- 2) Transformasi fitur kategorikal menggunakan teknik encoding agar dapat diproses oleh algoritma pembelajaran mesin.
- 3) Normalisasi fitur numerik menggunakan *MinMaxScaler* dengan persamaan:

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

Keterangan:

X_{scaled} : nilai hasil normalisasi,

X : nilai asli,

X_{min} : nilai minimum,

X_{max} : nilai maksimum.

- 4) Pembagian dataset menjadi data latih sebesar 80% dan data uji sebesar 20%.

Selain pembagian data latih dan data uji dengan rasio 80:20, penelitian ini juga menerapkan 5-Fold Cross Validation pada data latih untuk menguji stabilitas model. Dataset latih dibagi menjadi 5 subset, di mana 4 subset digunakan untuk pelatihan dan 1 subset untuk validasi secara bergantian. Nilai performa akhir diperoleh dari rata-rata hasil kelima fold. Proses ini bertujuan untuk:

- 1) Mengurangi bias akibat pembagian data tunggal.
- 2) Meningkatkan generalisasi model.
- 3) Mengukur konsistensi performa antar fold.

3. Penerapan Model Stacking

Pendekatan *stacking ensemble* diterapkan untuk meningkatkan akurasi klasifikasi melalui penggabungan dua model *machine learning*:

- 1) Model Dasar (*Base Learner*)
Model XGBoost digunakan sebagai *base learner* untuk menghasilkan probabilitas kelas awal. Parameter utama yang digunakan meliputi $n_estimators=100$, $learning_rate=0.1$, dan $max_depth=6$. Model ini dilatih menggunakan fungsi log loss untuk memaksimalkan kemampuan diskriminasi antar kelas.
- 2) Model Meta (*Meta Learner*)
Probabilitas keluaran dari XGBoost digunakan sebagai masukan bagi model MLP. Arsitektur MLP terdiri dari dua *hidden layer* dengan 64 dan 32 neuron, fungsi aktivasi ReLU, serta dropout sebesar 30% untuk mengurangi risiko overfitting. Lapisan keluaran menggunakan fungsi softmax untuk klasifikasi multikelas. Proses optimasi dilakukan menggunakan algoritma Adam dengan fungsi *loss sparse categorical crossentropy*.
- 3) Penggabungan Model
Pendekatan yang digunakan dalam penelitian ini adalah true stacking, di mana XGBoost berperan sebagai base learner yang menghasilkan probabilitas kelas, dan probabilitas tersebut digunakan sebagai input bagi MLP sebagai meta learner. Pada tahap akhir, dilakukan averaging probabilitas antara keluaran XGBoost dan MLP untuk meningkatkan stabilitas prediksi sebelum penentuan kelas akhir, sebagaimana dirumuskan pada persamaan:

$$P_{final} = \frac{P_{XGBoost} + P_{MLP}}{2}$$

Keterangan:

P_{final} : probabilitas akhir hasil penggabungan model,

$P_{XGBoost}$: probabilitas prediksi yang dihasilkan oleh model XGBoost,

P_{MLP} : probabilitas prediksi yang dihasilkan oleh MLP.

4. Evaluasi Model

Evaluasi kinerja model dilakukan menggunakan confusion matrix untuk mengukur akurasi prediksi terhadap data uji. Metrik evaluasi yang digunakan meliputi:

1. *Precision* yaitu menunjukkan proporsi prediksi positif yang benar dari keseluruhan prediksi positif.

$$Precision = (TP / (TP + FP)) * 100\%$$

2. *Recall* yaitu mengukur kemampuan model dalam menemukan seluruh data positif yang relevan.

$$Recall = (TP / (TP + FN)) * 100\%$$

3. Akurasi yaitu mengukur proporsi prediksi yang benar (baik positif maupun negatif) terhadap total data.

$$Accuracy = (TP + TN) / (TP + TN + FP + FN) * 100\%$$

Selain metrik akurasi, precision, recall, dan F1-score, penelitian ini juga menambahkan evaluasi menggunakan Receiver Operating Characteristic – Area Under Curve (ROC-AUC) untuk mengukur kemampuan diskriminatif model. ROC-AUC dihitung menggunakan pendekatan One-vs-Rest (OvR) untuk klasifikasi multikelas dengan persamaan:

$$ROC - AUC = \int_0^1 TPR(FPR^{-1}(x))dx$$

Keterangan:

TPR : True Positive Rate,

FPR : False Positive Rate,

Implementasi model dilakukan menggunakan Python 3.12 pada platform Google Colab, dengan library utama TensorFlow (Keras) untuk MLP, XGBoost untuk model boosting, dan Scikit-learn untuk evaluasi model. Proses pelatihan dijalankan pada lingkungan berbasis GPU yang tersedia di Google Colab untuk meningkatkan efisiensi komputasi.

IV. HASIL DAN PEMBAHASAN

Hasil dan pembahasan pada penelitian ini diperoleh melalui serangkaian tahapan yang meliputi pengumpulan dataset, preprocessing data, pelatihan model, serta evaluasi performa model klasifikasi status gizi balita. Dataset yang digunakan merupakan data rekam medis balita yang diperoleh dari RS Mitra Medika pada periode 2023–2024. Contoh sepuluh data awal dari dataset penelitian yang digunakan ditampilkan pada Tabel 2 untuk memberikan gambaran struktur dan karakteristik data.

Tabel 2. Sample Dataset Penelitian

Jenis Kelamin	Umur	TB	BB	Status Gizi	Wasting	Z-Score
Laki-laki	20	77.7	8.5	1	Underweight	-0.73
Laki-laki	10	79.0	10.3	0	Risk of Overweight	-0.60
Perempuan	2	50.3	8.3	1	Risk of Overweight	-3.47
Perempuan	5	56.4	10.9	1	Risk of Overweight	-2.86
Laki-laki	11	76.3	5.9	0	Severely Underweight	-0.87
Laki-laki	16	80.7	9.9	0	Normal weight	-0.43
Perempuan	15	72.6	6.5	0	Severely Underweight	-1.24
Perempuan	18	78.4	15.6	0	Risk of Overweight	-0.66
Laki-laki	2	63.4	7.0	0	Risk of Overweight	-2.16
Laki-laki	6	60.4	11.5	1	Risk of Overweight	-2.46

Sumber tabel : Data rekam medis RSU Mitra Medika Tanjung Mulia (2023–2024)

Dataset selanjutnya diproses melalui beberapa tahapan preprocessing untuk memastikan kualitas data sebelum digunakan pada proses pelatihan model. Tahapan pertama adalah transformasi variabel kategorikal. Fitur jenis kelamin dikonversi menjadi numerik (Laki-laki = 1, Perempuan = 0), sedangkan variabel wasting dikodekan menjadi biner (Normal weight = 0, selain itu = 1). Sementara itu, fitur numerik seperti TB, BB, dan Z-score dinormalisasi menggunakan metode MinMaxScaler agar berada pada rentang 0–1 sehingga tidak terjadi dominasi skala pada proses pelatihan model. Setelah proses transformasi, seluruh fitur dipastikan bertipe numerik. Setelah melalui tahapan preprocessing, dataset dibagi menjadi dua bagian, yaitu data latih dan data uji dengan rasio 80% untuk pelatihan dan 20% untuk pengujian (`random_state = 42`). Contoh sepuluh data awal setelah melalui proses *preprocessing* ditampilkan pada Tabel 3.

Tabel 3. Dataset Hasil Preprocessing

Jenis Kelamin	Umur	TB	BB	Status Gizi	Wasting	Z-Score
1	20	0.650000	0.462963	1	1	0.650000
1	10	0.674074	0.574074	0	1	0.674074
0	2	0.142593	0.450617	1	1	0.142593
0	5	0.255556	0.611111	1	1	0.255556
1	11	0.624074	0.302469	0	1	0.624074
1	16	0.705556	0.549383	0	0	0.705556
0	15	0.555556	0.339506	0	1	0.555556
0	18	0.662963	0.901235	0	1	0.662963
1	2	0.385185	0.370370	0	1	0.385185
1	6	0.329630	0.648148	1	1	0.329630

Sumber tabel : Hasil preprocessing data menggunakan Google Colab

Sebelum pelatihan model, dilakukan analisis korelasi antar fitur numerik menggunakan correlation matrix berbasis koefisien Pearson untuk mengidentifikasi hubungan linear dan potensi

multikolinearitas. Hasil analisis menunjukkan bahwa TB dan BB memiliki korelasi positif yang cukup kuat. *Z-score* juga berkorelasi signifikan dengan TB dan BB karena dihitung berdasarkan parameter antropometri tersebut. Sementara itu, Jenis_Kelamin dan Wasting memiliki korelasi rendah terhadap fitur lain, sehingga memberikan informasi tambahan yang tidak redundan. Tidak ditemukan korelasi sangat tinggi (>0,90), sehingga tidak terdapat indikasi multikolinearitas ekstrem. Dengan demikian, seluruh fitur dinilai relevan untuk digunakan dalam pelatihan model klasifikasi.



Gambar 2. Correlation Matrix Antar Fitur Penelitian
 Sumber: Hasil olahan data menggunakan Google Colab

Tahap berikutnya adalah pelatihan model klasifikasi menggunakan pendekatan *stacking*. Model XGBoost digunakan sebagai *base learner* dengan parameter *n_estimators* = 100, *learning_rate* = 0.1, dan *max_depth* = 6 untuk mempelajari pola awal hubungan antara variabel antropometri dan status gizi balita. Model ini menghasilkan keluaran berupa probabilitas kelas untuk setiap data. Probabilitas tersebut kemudian digunakan sebagai fitur masukan tambahan bagi model MLP yang berperan sebagai *meta learner*.

Model MLP dilatih menggunakan arsitektur dua *hidden layer* dengan jumlah neuron 64 dan 32 menggunakan fungsi aktivasi ReLU serta output layer softmax. Proses optimasi dilakukan menggunakan algoritma Adam dengan *fungsi loss sparse categorical crossentropy*. Pelatihan dilakukan selama 20 epoch dengan batch size 32. Pada epoch terakhir diperoleh *validation accuracy* sebesar 0.9964 dengan *validation loss* sebesar 0.0091, yang menunjukkan konvergensi model yang stabil.

Hasil 5-Fold Cross Validation pada data latih menunjukkan performa yang stabil dengan rata-rata akurasi sebesar 0.9958 dan standar deviasi sebesar 0.0007. Hasil 5-Fold Cross Validation ditampilkan pada Tabel 4.

Tabel 4. Hasil 5-Fold Cross Validation

Fold	Accuracy
Fold 1	0.99708358
Fold 2	0.99655332
Fold 3	0.99642052
Fold 4	0.99642052
Fold 5	0.99569137

Rata-rata 0.9964338640208676

Sumber tabel : Hasil cross_val_score pada Google Colab

Penggabungan keluaran model XGBoost dan MLP dilakukan menggunakan teknik averaging ensemble, di mana probabilitas akhir diperoleh dari rata-rata probabilitas kedua model sebelum dikonversi menjadi label kelas. Evaluasi performa model dilakukan menggunakan data uji dengan metrik akurasi, precision, recall, dan F1-score. Hasil evaluasi model stacking XGBoost-MLP ditampilkan pada Gambar 3.

```

Confusion Matrix:
[[14387  40]
 [  28 4404]]

Evaluation Metrics:
Akurasi : 0.9964
Presisi : 0.9964
Recall  : 0.9964
F1-Score : 0.9964

Classification Report (Per Kelas):
              precision    recall  f1-score   support

     0           1.00         1.00         1.00     14427
     1           0.99         0.99         0.99      4432

 accuracy                1.00         1.00     18859
 macro avg              0.99         1.00         0.99     18859
 weighted avg           1.00         1.00         1.00     18859
    
```

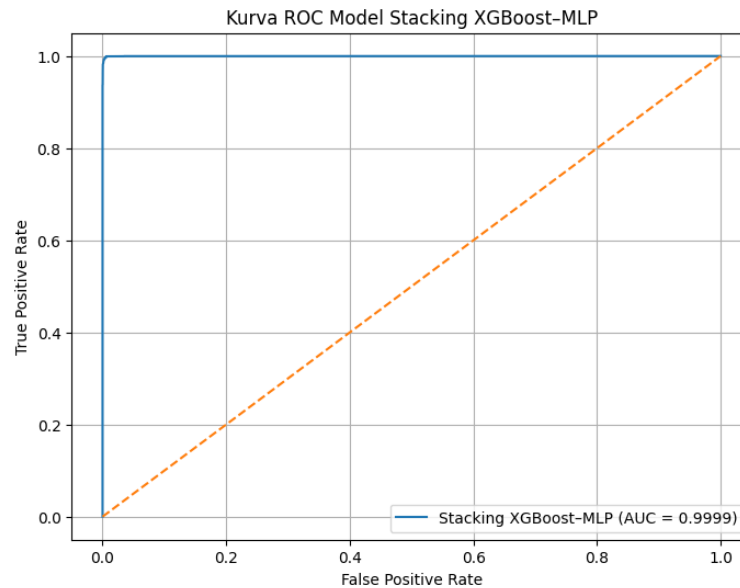
Gambar 3. Hasil Evaluasi Model Stacking XGBoost-MLP
 Sumber tabel : Hasil evaluasi model menggunakan Google Colab

Berdasarkan hasil evaluasi tersebut, model stacking menunjukkan performa yang sangat tinggi dengan akurasi sebesar 0.9964. *Confusion matrix* menunjukkan hanya 40 kesalahan prediksi pada kelas 0 dan 28 kesalahan pada kelas 1 dari total 18.859 data uji, yang menandakan kemampuan model dalam mengklasifikasikan kedua kelas secara konsisten. Nilai ROC-AUC yang diperoleh dari model stacking ditampilkan pada Tabel 5 dan hasil kurva ROC ditampilkan pada Gambar 4.

Tabel 5. Hasil Evaluasi ROC-AUC

Model	ROC-AUC (OvR)
Stacking XGBoost-MLP	0.9999223731000764

Sumber tabel : Hasil perhitungan roc_auc_score pada Google Colab



Gambar 4. Kurva ROC Model Stacking XGBoost-MLP

Sumber tabel : Visualisasi roc_curve menggunakan matplotlib pada Google Colab

Kurva ROC menunjukkan bahwa model memiliki kemampuan diskriminasi yang sangat baik dengan nilai AUC mendekati 1.0. Hal ini memperkuat hasil confusion matrix yang menunjukkan tingkat kesalahan prediksi yang sangat rendah. Peningkatan stabilitas prediksi pada model stacking menunjukkan bahwa kombinasi dua model dengan karakteristik berbeda mampu meningkatkan generalisasi model. Secara keseluruhan, pendekatan ini memiliki potensi yang sangat baik untuk diterapkan dalam sistem pendukung keputusan deteksi status gizi balita berbasis data besar.

V. KESIMPULAN

Penelitian ini bertujuan untuk mengoptimalkan klasifikasi status gizi balita berbasis big data menggunakan pendekatan stacking ensemble yang mengintegrasikan XGBoost sebagai base learner dan Multi-Layer Perceptron (MLP) sebagai meta learner, serta mengevaluasi apakah model tersebut mampu meningkatkan akurasi dan stabilitas dibandingkan model tunggal. Berdasarkan hasil eksperimen pada dataset berskala besar yang terdiri dari 94.291 observasi, model stacking XGBoost-MLP terbukti memberikan performa yang sangat tinggi dengan akurasi sebesar 99,64%, nilai ROC-AUC mendekati 1 (0,9999), serta precision, recall, dan F1-score yang seimbang pada kedua kelas. Hasil 5-Fold Cross Validation juga menunjukkan rata-rata akurasi 0,9964 dengan standar deviasi yang sangat rendah, yang mengindikasikan stabilitas dan konsistensi performa model.

Temuan ini membuktikan bahwa kombinasi algoritma boosting dan neural network melalui teknik stacking mampu meningkatkan kemampuan generalisasi dan diskriminasi model pada data kesehatan yang besar, heterogen, dan tidak seimbang, sekaligus mengatasi keterbatasan model tunggal sebagaimana diidentifikasi dalam penelitian sebelumnya. Dengan demikian, hipotesis penelitian yang menyatakan bahwa model stacking XGBoost-MLP lebih unggul dibandingkan model tunggal dapat diterima, dan pendekatan ini berpotensi diimplementasikan sebagai sistem pendukung keputusan berbasis big data untuk deteksi dini dan intervensi masalah gizi balita secara lebih akurat, stabil, dan adaptif.

VI. REFERENSI

Ashuri, P. I., Cahyani, I. A., & Kusuma, C. S. (2024). MIND (Multimedia Artificial Intelligent Networking Database Klasifikasi Penyakit Stunting Menggunakan Algoritma Multi-Layer Perceptron. *Journal MIND Journal* | ISSN, 9(1), 52–63. <https://doi.org/10.26760/mindjournal.v9i1.52-63>

- Bitew, F. H., Sparks, C. S., & Nyarko, S. H. (2022). Machine learning algorithms for predicting undernutrition among under-five children in Ethiopia. *Public Health Nutrition*, 25(2), 269–280. <https://doi.org/10.1017/S1368980021004262>
- Cahyani, I. A., Ashuri, P. I., & Aditya, C. S. K. (2024a). Stunting Disease Classification Using Multi-Layer Perceptron Algorithm with GridSearchCV. *Sinkron*, 9(1), 392–401. <https://doi.org/10.33395/sinkron.v9i1.13245>
- Cahyani, I. A., Ashuri, P. I., & Aditya, C. S. K. (2024b). Stunting Disease Classification Using Multi-Layer Perceptron Algorithm with GridSearchCV. *Sinkron*, 9(1), 392–401. <https://doi.org/10.33395/sinkron.v9i1.13245>
- Hadini Marpaung, S., Mikael Sinaga, F., Hawani Rambe, K., & Presly Simamora, F. (2025). Random Forest Optimization Using Recursive Feature Elimination for Stunting Classification. *Indonesian Journal of Artificial Intelligence and Data Mining (IJAIMD)*, 8(1), 281–287. <https://doi.org/10.24014/ijaidm.v8i1.35295>
- Hariguna, T., Sarmini, & Azis, A. (2024a). Health and Socio-Demographic Risk Factors of Childhood Stunting: Assessing the Role of Factor Interactions Through the Development of an AI Predictive Model. *Journal of Applied Data Sciences*, 5(4), 2175–2186. <https://doi.org/10.47738/jads.v5i4.612>
- Hariguna, T., Sarmini, & Azis, A. (2024b). Health and Socio-Demographic Risk Factors of Childhood Stunting: Assessing the Role of Factor Interactions Through the Development of an AI Predictive Model. *Journal of Applied Data Sciences*, 5(4), 2175–2186. <https://doi.org/10.47738/jads.v5i4.612>
- Hasdyna, N., Dinata, R. K., Rahmi, R., & Fajri, T. I. (2024). *Hybrid Machine Learning for Stunting Prevalence: A Novel Comprehensive Approach to Classification, Prediction, and Clustering Optimization in Aceh, Indonesia*. <https://doi.org/10.20944/preprints202409.0485.v1>
- Le, T. T. H., Oktian, Y. E., & Kim, H. (2022). XGBoost for Imbalanced Multiclass Classification-Based Industrial Internet of Things Intrusion Detection Systems. *Sustainability (Switzerland)*, 14(14). <https://doi.org/10.3390/su14148707>
- Nirsal, S.Kom., M. P., Dr. Ir. H. Apriyanto, M.Si., M. M., Frans Mikael Sinaga, S.Kom., M. K., Yuni Marlina Saragih, S.Kom., M. K., Kusumastuti, S. Y., Loso Judijanto, S.Si., M.M., M. S., Kelvin, S.Kom., M. K., Dr. Ir. Ridwang, S.Kom., M.T., I., Sophya Hadini Marpaung, S.Kom., M. M. S. I., Najirah Umar, S.Kom, M., Dr. Ahmad Muhammad Thantawi, S.T., M. M., & Khairul Hawani Rambe, B.Sc., M. E. (n.d.). *BIG DATA (Panduan dan Peluang di Era Digital)*. Retrieved <https://buku.sonpedia.com/2024/12/big-data-panduan-dan-peluang-di-era.html>
- Noh, B., Youm, C., Goh, E., Lee, M., Park, H., Jeon, H., & Kim, O. Y. (2021). XGBoost based machine learning approach to predict the risk of fall in older adults using gait outcomes. *Scientific Reports*, 11(1). <https://doi.org/10.1038/s41598-021-91797-w>
- Renaningtias, N., & Prihatiningrum, A. (2025). *Child Stunting Risk Analysis through Machine Learning Models using XGBoost Algorithm*. 12(2), 135–142.

- Samuel Aleksander Mandowen, S. Si. M. I., Ir. Jarot Budiasto, S.T., M. T., Pulung Hendro Prastyo, S.S.T., M. E., Santi Prayudani, S.Kom., M. K., Khairul Hawani Rambe, B.Sc., M. E., Tri Sandhika Jaya, S.Kom., M. K., Muhammad Ridwan, ST, M., Taufiqurrahman, S.T., M. K., Kuntarto, G. P., & Judijanto, L. (2025). *PENGANTAR MACHINE LEARNING*. PT. Sonpedia Publishing Indonesia. <https://buku.sonpedia.com/2025/11/pengantar-machine-learning-teori.html>
- Sudirwo, S, E., M. M., Abdul Hadi, S.E., M. S., Loso Judijanto, SSi, MM, Ms., Nuraini Purwandari, ST., M., Neni N. L. Ersela Zain, S.Si., M. M., Khairul Hawani Rambe, B.Sc., M. E., Iqbal Ramadhani Mukhlis, S.Kom., M. K., Hilman Jihadi S.Kom., M., Prof. Dr. Hanik Mahliatussikah, M. H., Baskoro, Dr. B. H., & Adnan Yusufi, M. Pd. I. (2025). *ARTIFICIAL INTELLIGENCE*. <https://buku.sonpedia.com/2025/04/artificial-intelligence-teori-konsep.html>
- Syahrial, S., Ilham, R., Asikin, Z. F., & Nurdin, St. S. I. (2022). Stunting Classification in Children's Measurement Data Using Machine Learning Models. *Journal La Multiapp*, 3(2), 52–60. <https://doi.org/10.37899/journallamultiapp.v3i2.614>
- Vipin, V., Trivedi, K., & Koley, S. (2023). Optimization of parameters of the OWC wave energy converter device using MLP and XGBoost models. *Results in Physics*, 55. <https://doi.org/10.1016/j.rinp.2023.107163>
- Yaqoob, A., Verma, N. K., Aziz, R. M., & Shah, M. A. (2024). Optimizing cancer classification: a hybrid RDO-XGBoost approach for feature selection and predictive insights. *Cancer Immunology, Immunotherapy : CII*, 73(12), 261. <https://doi.org/10.1007/s00262-024-03843-x>
- Yousefi, Z., Alesheikh, A. A., Jafari, A., Torktatari, S., & Sharif, M. (2024). Stacking Ensemble Technique Using Optimized Machine Learning Models with Boruta–XGBoost Feature Selection for Landslide Susceptibility Mapping: A Case of Kermanshah Province, Iran. *Information (Switzerland)*, 15(11). <https://doi.org/10.3390/info15110689>
- Zhang, H. W., Wang, Y. R., Hu, B., Song, B., Wen, Z. J., Su, L., Chen, X. M., Wang, X., Zhou, P., Zhong, X. M., Pang, H. W., & Wang, Y. H. (2024). Using machine learning to develop a stacking ensemble learning model for the CT radiomics classification of brain metastases. *Scientific Reports*, 14(1). <https://doi.org/10.1038/s41598-024-80210-x>