

ANALISIS PERFORMA JST BACKPROPAGATION MENGGUNAKAN OPTIMIZER SGD, RMSPROP DAN ADAM UNTUK KLASIFIKASI STUNTING

Adi Putra Sinaga, Naikson Fandier Saragih, Indra Kelana Jaya✉

Program Studi Teknik Informatika, Universitas Methodist Indonesia, Medan, Indonesia

Email: indraikjs@gmail.com

DOI: <https://doi.org/10.46880/jmika.Vol10No1.pp393-401>

ABSTRACT

Stunting is a linear growth disorder in toddlers caused by chronic malnutrition during critical developmental periods and remains a significant national health problem in Indonesia. According to the 2022 Indonesian Nutritional Status Survey (SSGI), national stunting prevalence reached 21.6%, surpassing the World Health Organization (WHO) recommended ceiling of 20%. This study evaluates three optimizers — Stochastic Gradient Descent (SGD), RMSProp, and Adam — within a Backpropagation Artificial Neural Network (ANN) for classifying toddler stunting status. The dataset comprises 1,454 anthropometric records (sex, age, height) collected from UPT Puskesmas Kampung Baru, Medan City, covering 2020–2023. Preprocessing included MinMax scaling (fitted exclusively on training data to prevent leakage) and SMOTETomek resampling to address severe class imbalance. Forty-eight model configurations were evaluated via grid search across a 3-32-1 architecture (ReLU hidden layer, Sigmoid output). Evaluation metrics comprised accuracy, precision, recall, F1-score, specificity, AUC, and Matthews Correlation Coefficient (MCC). RMSProp achieved perfect scores on all metrics (1.000) in the shortest execution time (133.42 s). Adam achieved equivalent classification performance with similarly rapid convergence (135.28 s). SGD attained perfect recall (1.000) only when training data were simultaneously balanced and normalized, empirically demonstrating that non-adaptive optimizers require both interventions to compete with adaptive counterparts.

Keyword: Artificial Neural Network, Backpropagation, Stunting, Classification, Optimizer.

ABSTRAK

Stunting merupakan gangguan pertumbuhan linier pada balita yang disebabkan oleh kekurangan gizi kronis dalam periode perkembangan kritis dan masih menjadi masalah kesehatan nasional yang signifikan. Berdasarkan Survei Status Gizi Indonesia (SSGI) tahun 2022, prevalensi stunting nasional mencapai 21,6%, melampaui ambang batas yang direkomendasikan oleh World Health Organization (WHO) sebesar 20%. Penelitian ini mengevaluasi tiga optimizer — Stochastic Gradient Descent (SGD), RMSProp, dan Adam — dalam algoritma Backpropagation Jaringan Saraf Tiruan (JST) untuk klasifikasi status stunting balita. Dataset mencakup 1.454 rekam antropometri (jenis kelamin, usia, tinggi badan) yang dikumpulkan dari UPT Puskesmas Kampung Baru, Kota Medan, periode 2020–2023. Preprocessing meliputi normalisasi MinMax (fit eksklusif pada data latih untuk mencegah kebocoran data) dan resampling SMOTETomek untuk mengatasi ketidakseimbangan kelas yang ekstrem. Sebanyak 48 konfigurasi dievaluasi melalui grid search pada arsitektur 3-32-1 (lapisan tersembunyi ReLU, keluaran Sigmoid). Metrik evaluasi meliputi akurasi, presisi, recall, F1-Score, spesifisitas, AUC, dan Matthews Correlation Coefficient (MCC). RMSProp mencapai skor sempurna pada semua metrik (1,000) dengan waktu eksekusi terpendek (133,42 detik). Adam mencapai performa setara dengan konvergensi sama cepat (135,28 detik). SGD mencapai recall sempurna (1,000) hanya ketika data latih sekaligus diseimbangkan dan dinormalisasi, yang membuktikan secara empiris bahwa optimizer non-adaptif memerlukan kedua intervensi untuk menyaingi optimizer adaptif.

Kata Kunci: Jaringan Saraf Tiruan, Backpropagation, Stunting, Klasifikasi, Optimizer.

PENDAHULUAN

Stunting merepresentasikan kegagalan pertumbuhan linier pada balita yang diakibatkan oleh malnutrisi kronis dalam jangka panjang, khususnya pada fase 1.000 hari (sekitar 2,7 tahun) pertama kehidupan (Kementerian Kesehatan RI, 2020).

Indikator utamanya adalah postur tinggi badan yang berada di bawah standar normal berdasarkan grafik usia, yang tidak hanya berdampak pada keterbelakangan fisik tetapi juga memicu gangguan perkembangan kognitif jangka panjang. Merujuk pada laporan Survei Status Gizi Indonesia (SSGI) tahun



2022, angka prevalensi stunting di tingkat nasional mencapai 21,6% (Kementerian Kesehatan RI, 2022). Walaupun terdapat tren penurunan dari tahun sebelumnya (24,4%), persentase ini tetap berada di atas ambang batas toleransi maksimal yang direkomendasikan oleh World Health Organization (WHO) sebesar 20% (WHO, 2014). Fakta ini menegaskan perlunya inovasi metode penanganan, dari sekadar intervensi gizi konvensional menjadi pendekatan prediktif berbasis komputasi untuk deteksi dini di fasilitas pelayanan kesehatan dasar.

Implementasi kecerdasan buatan, khususnya arsitektur Jaringan Saraf Tiruan (JST) yang dilatih dengan algoritma Backpropagation, telah teruji keandalannya dalam menangani problem klasifikasi pada sektor kesehatan pediatrik (Rumelhart et al., 1986; Haykin, 2009). Berbagai kajian terdahulu telah membuktikan hal ini. Asriningtias et al. (2025) sukses membangun model Artificial Neural Network (ANN) yang dioptimasi melalui GridSearchCV untuk deteksi dini stunting dengan tingkat akurasi 99% pada data berstandar WHO. Di sisi lain, riset Ndagijimana et al. (2025) mengaplikasikan metode Multilayer Perceptron (MLP) dengan skor Area Under Curve (AUC) mencapai 0,84. Dalam ranah penelitian lokal, Yunus et al. (2025) telah mengomparasikan teknik machine learning tradisional (Naive Bayes dan Decision Tree C4.5) untuk pemodelan stunting. Sayangnya, mayoritas studi berbasis komputasi tersebut masih bertumpu pada satu rancangan hiperparameter statis dan belum melakukan analisis komparatif yang mendalam terkait dampak pemilihan optimizer selama proses training model.

Penentuan fungsi optimizer merupakan tahapan krusial dalam pelatihan algoritma Backpropagation karena secara langsung mendikte laju konvergensi, stabilitas kalkulasi, serta kemampuan generalisasi pemodelan (Ruder, 2016). Saat ini, Stochastic Gradient Descent (SGD), Root Mean Square Propagation (RMSProp), dan Adaptive Moment Estimation (Adam) adalah tiga optimizer yang mendominasi riset deep learning (Kingma & Ba, 2015; Tieleman & Hinton, 2012). Beberapa riset lintas disiplin telah membandingkan ketiganya, seperti pada klasifikasi citra ekspresi wajah (Anggara et al., 2023) maupun estimasi prediktif numerik (Wijaya et al., 2025), yang menyimpulkan bahwa keunggulan tiap optimizer sangat fluktuatif bergantung pada karakteristik dataset dan kedalaman arsitektur. Kendati demikian, pada kasus klasifikasi malnutrisi kronis yang identik dengan fenomena imbalanced data secara ekstrem — kuantitas sampel stunting jauh lebih minim dari sampel balita normal — efek pemilihan optimizer terhadap performa

model, terutama saat dikolaborasikan dengan teknik resampling dan standardisasi fitur, masih menjadi ruang diskursus yang belum tereksplorasi secara sistematis dalam literatur lokal.

Berpijak pada uraian di atas, riset ini dilatarbelakangi oleh dua celah empiris (research gap). Pertama, minimnya analisis komparatif yang menyejajarkan kinerja operasional SGD, RMSProp, dan Adam secara komprehensif pada dataset stunting berkarakteristik kewilayahan di Indonesia. Kedua, urgensi untuk mengisolasi dan mengukur dampak dari dua teknik preprocessing yang kerap digabungkan tanpa evaluasi terpisah: normalisasi fitur melalui Min-Max Scaling dan teknik sintesis penyeimbangan kelas melalui SMOTETomek (Chawla et al., 2002; Batista et al., 2004). Sebagai rujukan tambahan, eksperimen Raihana et al. (2025) mendemonstrasikan bahwa algoritma SMOTE terbukti sukses mengontrol metrik evaluasi pada kelas minoritas, kendati riset tersebut belum mengintegrasikannya ke dalam mesin JST maupun skema uji optimizer.

Untuk menjembatani celah saintifik tersebut, studi ini bertujuan membangun sekaligus mengevaluasi 48 skema konfigurasi model JST Backpropagation berarsitektur 3-32-1 (3 neuron masukan, 32 neuron lapisan tersembunyi dengan fungsi aktivasi Rectified Linear Unit/ReLU, dan 1 neuron keluaran dengan fungsi aktivasi Sigmoid) dalam mengklasifikasikan status stunting menggunakan basis rekam medis dari UPT Puskesmas Kampung Baru, Kota Medan, pada rentang 2020–2023. Tiga kontribusi utama yang ditawarkan penelitian ini meliputi: (1) eksekusi MinMaxScaler yang ketat terhadap potensi kebocoran data (data leakage) dengan mekanisme fit eksklusif pada set data latih; (2) penerapan algoritma SMOTETomek yang dibatasi murni pada data training; serta (3) penelusuran grid search atas lima dimensi hiperparameter untuk mencari kombinasi optimizer terbaik. Output dari penelitian ini diproyeksikan menjadi rujukan metodologis yang praktis bagi peneliti dalam merancang arsitektur komputasi klinis biner yang berhadapan dengan anomali data tidak seimbang.

TINJAUAN PUSTAKA

Jaringan Saraf Tiruan

Jaringan Saraf Tiruan (JST) adalah model komputasi yang terinspirasi oleh struktur dan cara kerja sistem saraf biologis, terdiri atas sekumpulan neuron buatan yang saling terhubung melalui bobot yang dapat dilatih (Haykin, 2009). Sebuah neuron buatan menerima sejumlah masukan, menghitung jumlah terbobot dari masukan tersebut, menambahkannya dengan bias, lalu melewatkan hasil tersebut pada fungsi

aktivasi non-linier. Susunan neuron dalam beberapa lapisan — umumnya satu lapisan masukan, satu atau lebih lapisan tersembunyi, dan satu lapisan keluaran — memungkinkan JST mempelajari representasi non-linier yang kompleks dari data.

Algoritma Backpropagation

Algoritma Backpropagation merupakan metode pelatihan JST berlapis yang diperkenalkan oleh Rumelhart et al. (1986) dan hingga kini menjadi fondasi utama pembelajaran berbasis gradien di bidang deep learning (Goodfellow et al., 2016). Algoritma ini bekerja dalam tiga fase berulang. Pertama, fase propagasi maju (forward pass) menyebarkan sinyal masukan melalui jaringan untuk menghasilkan keluaran prediksi. Kedua, fase propagasi mundur (backward pass) menghitung gradien fungsi kerugian terhadap setiap bobot menggunakan chain rule. Ketiga, fase pembaruan bobot menyesuaikan setiap bobot ke arah yang berlawanan dengan gradien, dengan besaran langkah yang dikendalikan oleh learning rate.

Fungsi Aktivasi

Fungsi aktivasi memperkenalkan non-linearitas ke dalam JST, yang merupakan syarat agar jaringan mampu mempelajari pola non-linier. Fungsi Rectified Linear Unit (ReLU) didefinisikan sebagai $f(x) = \max(0, x)$, banyak digunakan pada lapisan tersembunyi karena mempercepat konvergensi pelatihan dan mengurangi masalah vanishing gradient (Nair & Hinton, 2010). Pada lapisan keluaran untuk klasifikasi biner, fungsi sigmoid biner dengan formulasi $f(x) = 1/(1 + e^{-x})$ digunakan karena memetakan keluaran ke rentang [0, 1] yang dapat ditafsirkan langsung sebagai probabilitas kelas positif (Goodfellow et al., 2016). Pada penelitian ini, ReLU digunakan pada lapisan tersembunyi (32 neuron) dan Sigmoid pada lapisan keluaran (1 neuron).

Optimizer SGD, RMSProp dan Adam

Optimizer merupakan algoritma yang menentukan bagaimana bobot model diperbarui berdasarkan gradien kerugian (Ruder, 2016). Stochastic Gradient Descent (SGD) adalah optimizer non-adaptif yang memperbarui bobot dengan persamaan $w \leftarrow w - \alpha \nabla L(w)$, dengan α merupakan learning rate global. SGD bersifat sederhana namun sensitif terhadap skala fitur dan memerlukan tuning hati-hati. RMSProp (Tieleman & Hinton, 2012) adalah optimizer adaptif yang menormalisasi gradien menggunakan rata-rata bergerak dari kuadrat gradien, sehingga setiap parameter mendapatkan learning rate efektif yang berbeda. Adam (Kingma & Ba, 2015) memperluas pendekatan ini dengan menggabungkan

momentum orde pertama dan momentum orde kedua dengan koreksi bias inisialisasi.

Karena kemampuan adaptifnya, Adam dan RMSProp umumnya lebih robust terhadap perbedaan skala fitur dibandingkan SGD. Bukti empiris mengenai variasi kinerja ini telah dilaporkan pada beragam domain: Anggara et al. (2023) menemukan bahwa Adam unggul dalam akurasi (68,61%) dibanding SGD (57,68%) dan RMSProp (54,83%) pada klasifikasi citra, sementara Wijaya et al. (2025) memperoleh hasil serupa pada ANN untuk prediksi harga saham. Temuan-temuan tersebut mengindikasikan bahwa pemilihan optimizer terbaik bersifat dataset-dependent dan memerlukan eksperimen empiris pada domain target.

Normalisasi Fitur dengan MinMaxScaler

MinMaxScaler adalah teknik normalisasi yang mentransformasikan setiap fitur ke rentang [0, 1] dengan formula $x' = (x - x_{\min}) / (x_{\max} - x_{\min})$. Untuk mencegah kebocoran data (data leakage), parameter x_{\min} dan x_{\max} hanya dihitung (fit) dari data latih, kemudian diterapkan (transform) pada data latih dan data uji secara terpisah (Goodfellow et al., 2016). Normalisasi sangat penting bagi optimizer non-adaptif seperti SGD: tanpa normalisasi, gradien akan didominasi oleh fitur dengan rentang terluas. Dalam penelitian ini, fitur tinggi badan memiliki rentang sekitar 50-120 cm, sehingga arah pembaruan bobot menjadi suboptimal.

Penanganan Kelas Tidak Seimbang dengan SMOTETomek

Pada kasus klasifikasi stunting, jumlah sampel kelas minoritas (stunting) jauh lebih sedikit dibandingkan kelas mayoritas (tidak stunting), sehingga model cenderung bias ke kelas mayoritas dan menghasilkan recall yang sangat rendah. Synthetic Minority Over-sampling Technique (SMOTE) yang diperkenalkan oleh Chawla et al. (2002) menghasilkan sampel sintesis kelas minoritas dengan menginterpolasi antar tetangga terdekat di ruang fitur. SMOTETomek (Batista et al., 2004) merupakan kombinasi SMOTE dengan algoritma Tomek Links yang membersihkan sampel-sampel yang berada terlalu dekat dengan batas keputusan antar-kelas. Kombinasi ini menghasilkan distribusi data yang seimbang sekaligus lebih bersih secara topologi. Sesuai praktik standar, SMOTETomek hanya diterapkan pada data latih, sedangkan data uji tetap mempertahankan distribusi alaminya untuk menjamin validitas evaluasi. Dalam konteks deteksi stunting, Raihana et al. (2025) mendemonstrasikan bahwa kombinasi SMOTE dengan algoritma KNN

mampu meningkatkan akurasi deteksi kelas minoritas hingga 99,17%, mengonfirmasi relevansi teknik oversampling untuk masalah klasifikasi stunting yang inheren imbalanced.

Metrik Evaluasi pada Data Tidak Seimbang

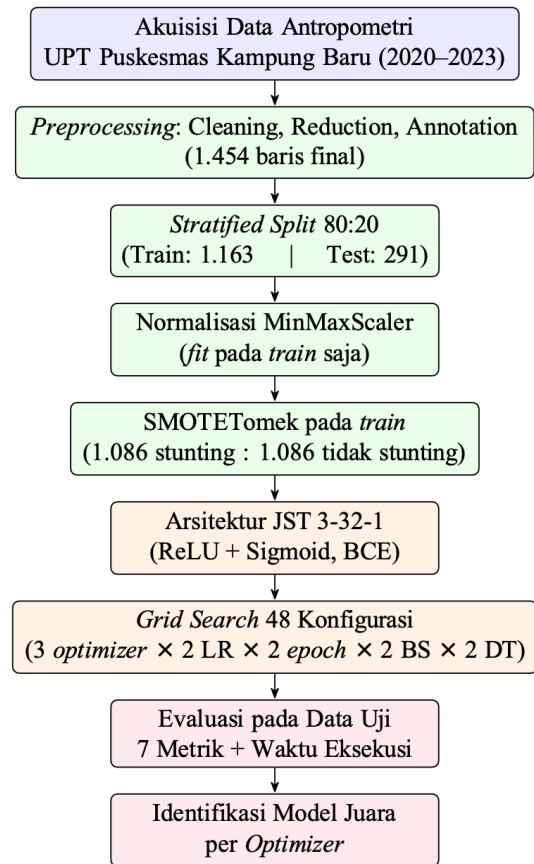
Pada data imbalanced, akurasi tunggal dapat menyesatkan karena nilai tinggi dapat dicapai hanya dengan memprediksi seluruh sampel sebagai kelas mayoritas. Evaluasi dalam penelitian ini menggunakan tujuh metrik yang diturunkan dari confusion matrix: $accuracy = (TP+TN)/N$; $precision = TP/(TP+FP)$; $recall = TP/(TP+FN)$; $specificity = TN/(TN+FP)$; $F1-Score = 2 \cdot Prec \cdot Rec / (Prec+Rec)$; AUC sebagai area di bawah kurva ROC; serta MCC sebagai ϕ -koefisien Matthews. Matthews Correlation Coefficient (MCC) dilaporkan sebagai metrik prioritas karena Chicco & Jurman (2020) menunjukkan bahwa MCC merupakan metrik yang lebih informatif dibanding akurasi maupun F1 pada klasifikasi biner dengan kelas tidak seimbang. Dalam konteks deteksi dini stunting, recall menjadi metrik klinis paling kritis karena setiap balita stunting yang tidak terdeteksi (false negative) berarti kehilangan kesempatan intervensi gizi tepat waktu.

METODE PENELITIAN

Alur Penelitian

Penelitian ini dirancang sebagai studi eksperimental kuantitatif yang membandingkan performa tiga optimizer pada arsitektur Jaringan Saraf Tiruan (JST) berbasis algoritma Backpropagation untuk klasifikasi status stunting balita. Alur penelitian secara keseluruhan mencakup tahapan dari akuisisi data, preprocessing, perancangan arsitektur, pelaksanaan grid search, hingga evaluasi metrik klasifikasi. Tahapan utama meliputi: (1) akuisisi data antropometri UPT Puskesmas Kampung Baru periode 2020–2023; (2) preprocessing yang mencakup cleaning, reduction, dan annotation sehingga diperoleh 1.454 baris data final; (3) stratified split 80:20 menghasilkan data latih 1.163 baris dan data uji 291 baris; (4) normalisasi menggunakan MinMaxScaler yang di-fit hanya pada data latih; (5) penerapan SMOTETomek pada data latih yang menghasilkan 1.086 sampel per kelas; (6) pelatihan model JST 3-32-1 dengan ReLU dan Sigmoid; (7) grid search 48 konfigurasi (3 optimizer \times 2 learning rate \times 2 epoch \times 2 batch size \times 2 tipe data); (8) evaluasi pada data uji menggunakan tujuh metrik dan waktu eksekusi; serta (9) identifikasi model juara per optimizer. Diagram alur lengkap disajikan pada Gambar 1. Rangkaian metodologi ini dirancang secara berurutan guna memastikan tidak adanya kebocoran data (data

leakage) selama proses pra-pemrosesan, serta menjaga validitas evaluasi komparatif dari ketiga optimizer yang diuji.



Gambar 1. Diagram Alur Penelitian Klasifikasi Stunting Menggunakan JST Backpropagation dengan Tiga Optimizer

Sumber dan Karakteristik Data

Data yang digunakan adalah rekam antropometri balita dari UPT Puskesmas Kampung Baru, Kota Medan, yang dikumpulkan pada rentang 2020–2023. Setelah melalui tahap cleaning (pembersihan nilai hilang dan outlier menggunakan metode IQR dengan $k=1,4$ per kelas), annotation (pelabelan status stunting berdasarkan indeks TB/U sesuai Peraturan Menteri Kesehatan Nomor 2 Tahun 2020), dan reduction (seleksi atribut), diperoleh 1.454 baris data unik dengan empat atribut: jenis kelamin (kategorik: Laki-laki=1, Perempuan=0), umur (numerik dalam bulan, rentang 0–59), tinggi badan (numerik dalam cm), dan status sebagai label biner (Stunting=1, Tidak Stunting=0).

Preprocessing dan Pembagian Data

Dataset dibagi menjadi data latih dan data uji dengan rasio 80:20 menggunakan stratified random split (random_state=42). Stratifikasi memastikan

proporsi kelas stunting terjaga di kedua subset. Total data uji adalah 291 baris (18 stunting, 273 tidak stunting), sedangkan data latih berjumlah 1.163 baris (74 stunting, 1.089 tidak stunting). Normalisasi fitur dilakukan menggunakan MinMaxScaler ke rentang [0, 1], di mana parameter scaler hanya dihitung dari data latih (fit on train only) untuk mencegah kebocoran informasi dari set uji ke proses pelatihan.

Penanganan Kelas Tidak Seimbang

Pada data latih asli, rasio stunting:tidak-stunting adalah $74:1.089 \approx 1:15$ (sekitar 6,4% kelas minoritas). Rasio yang sangat tidak seimbang ini berpotensi menyebabkan model belajar untuk mengabaikan kelas minoritas. Untuk mengatasinya, diterapkan teknik SMOTETomek pada data latih setelah scaling. Setelah resampling, distribusi kelas menjadi seimbang dengan 1.086 sampel per kelas (total 2.172 sampel). Penelitian ini mengevaluasi dua tipe data latih: DT1 (original) dengan 1.163 baris dan distribusi asli tidak seimbang, serta DT2 (balanced) dengan 2.172 baris setelah SMOTETomek.

Arsitektur Model JST

Model JST menggunakan arsitektur 3-32-1: tiga neuron masukan, satu lapisan tersembunyi berisi 32 neuron dengan fungsi aktivasi Rectified Linear Unit (ReLU), dan satu neuron keluaran dengan fungsi aktivasi Sigmoid. Fungsi kerugian yang digunakan adalah binary cross-entropy. Seluruh model diinisialisasi dengan bobot awal yang sama (hardcoded) dan seed yang sama (42) untuk memastikan reproduktifitas eksperimen lintas seluruh 48 konfigurasi. ModelCheckpoint digunakan selama pelatihan untuk menyimpan bobot model terbaik berdasarkan validation accuracy (save_best_only=True), sehingga model yang dievaluasi adalah model pada epoch dengan performa validasi optimal — bukan epoch terakhir.

Skema Grid Search 48 Eksperimen

Eksperimen melibatkan 48 konfigurasi model dari kombinasi lima hiperparameter sebagaimana dirinci pada Tabel 1. Penamaan model mengikuti format DT[1/2]_O[1/2/3]_LR[1/2]_E[1/2]_BS[1/2], dengan kode DT untuk tipe data, O untuk optimizer, LR untuk learning rate, E untuk epoch, dan BS untuk batch size. Sebagai contoh, DT2_O3_LR1_E1_BS2

merepresentasikan model dengan data balanced, optimizer RMSProp, learning rate 0,01, 500 epoch, dan batch size 32.

Tabel 1. Konfigurasi Grid Search Lima Hiperparameter (Total 48 Model)

Hiperparameter	Nilai yang Diuji	Kode	Jumlah
Tipe data latih	original, balanced	DT1, DT2	2
Optimizer	SGD, Adam, RMSProp	O1, O2, O3	3
Learning rate	0,01 ; 0,001	LR1, LR2	2
Epoch	500 ; 1.000	E1, E2	2
Batch size	16 ; 32	BS1, BS2	2
Total kombinasi	$2 \times 3 \times 2 \times 2 \times 2$	—	48

Metrik Evaluasi

Evaluasi dilakukan pada data uji yang tidak tersentuh selama pelatihan (n=291; 18 stunting, 273 tidak stunting) menggunakan tujuh metrik: accuracy, precision, recall, specificity, F1-Score, AUC, dan MCC. Recall digunakan sebagai prioritas tertinggi karena meminimalkan balita stunting yang tidak terdeteksi (false negative), dengan F1-Score sebagai tiebreaker utama, accuracy sebagai tiebreaker sekunder, dan MCC sebagai metrik utama untuk data tidak seimbang. Selain metrik klasifikasi, waktu eksekusi pelatihan (detik) dicatat untuk setiap model sebagai indikator efisiensi komputasi.

HASIL DAN PEMBAHASAN

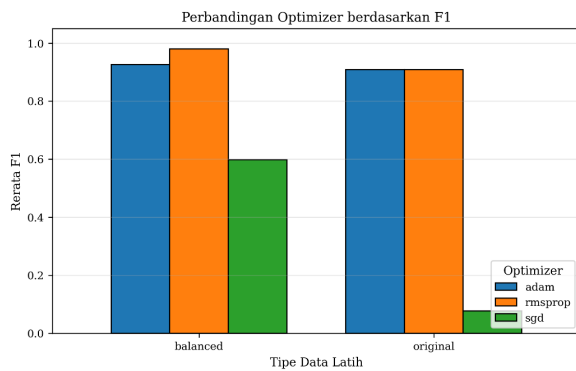
Identifikasi Model Juara per Optimizer

Dari 48 konfigurasi yang dievaluasi pada data uji, dipilih satu model terbaik untuk masing-masing optimizer berdasarkan kombinasi recall tertinggi (sebagai prioritas klinis), diikuti oleh F1-Score, akurasi, dan waktu eksekusi sebagai tiebreaker. Seluruh tiga model juara berasal dari skenario data balanced (DT2), mengindikasikan bahwa penyeimbangan kelas melalui SMOTETomek merupakan prasyarat mutlak untuk performa optimal pada dataset yang sangat tidak seimbang ini. Tabel 2 menyajikan komparasi lengkap ketiga model juara, sementara Gambar 2 dan Gambar 3 menampilkan perbandingan agregat lintas optimizer pada metrik F1-Score dan Recall.

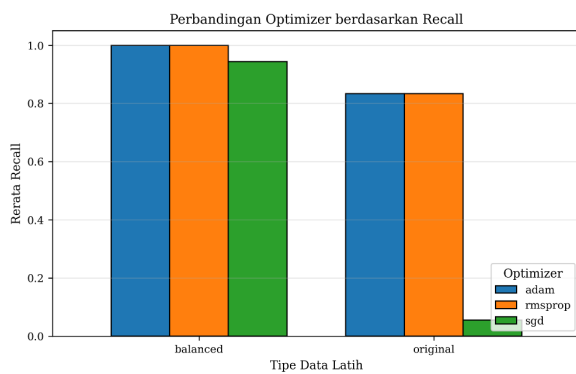
Tabel 2. Komparasi Metrik Tiga Model Juara untuk Masing-Masing Optimizer Pada Data Uji (n=291)

Atribut	SGD	Adam	RMSProp
Nama Model	DT2_O1_LR1_E2_BS2	DT2_O2_LR2_E1_BS2	DT2_O3_LR1_E1_BS2

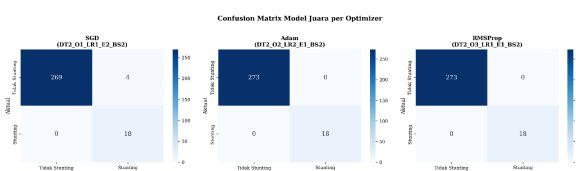
Atribut	SGD	Adam	RMSProp
Learning rate	0,01	0,001	0,01
Epoch	1.000	500	500
Batch size	32	32	32
TP / TN / FP / FN	18 / 269 / 4 / 0	18 / 273 / 0 / 0	18 / 273 / 0 / 0
Accuracy	0,9863	1,0000	1,0000
Precision	0,8182	1,0000	1,0000
Recall	1,0000	1,0000	1,0000
Specificity	0,9853	1,0000	1,0000
F1-Score	0,9000	1,0000	1,0000
AUC	1,0000	1,0000	1,0000
MCC	0,8979	1,0000	1,0000
Waktu eksekusi (detik)	251,66	135,28	133,42



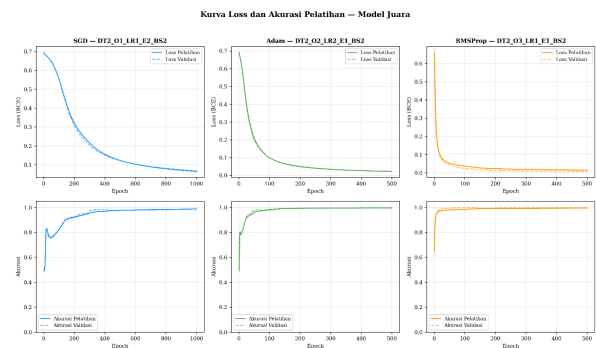
Gambar 2. Perbandingan Agregat F1-Score Antar Optimizer pada Skenario Data Original dan Balanced



Gambar 3. Perbandingan Agregat Recall Antar Optimizer pada Skenario Data Original dan Balanced



Gambar 4. Confusion Matrix Model Juara Untuk Masing-Masing Optimizer



Gambar 5. Kurva Loss dan Akurasi Pelatihan Untuk Ketiga Model Juara

Performa RMSProp: Stabilitas dan Efisiensi

Hasil pada Tabel 2 memperlihatkan bahwa RMSProp muncul sebagai optimizer dengan kombinasi performa dan efisiensi terbaik. Selain mencapai metrik klasifikasi sempurna (accuracy, precision, recall, F1-Score, AUC, dan MCC seluruhnya bernilai 1,000), model juara RMSProp (DT2_O3_LR1_E1_BS2) merupakan model dengan waktu eksekusi terpendek di antara ketiga juara, yaitu 133,42 detik.

Stabilitas RMSProp tercermin pula pada level agregat: dari 8 konfigurasi yang dievaluasi pada data balanced (24 konfigurasi balanced dibagi 3 optimizer = 8 per optimizer), seluruh konfigurasi RMSProp mencapai recall sempurna (1,000), dan empat di antaranya, seluruhnya pada learning rate 0,01 (LR1), mencapai semua metrik sempurna pada accuracy, F1, dan MCC. Empat konfigurasi dengan learning rate 0,001 (LR2) mencapai recall sempurna namun masih menghasilkan beberapa false positive (FP=1-3).

Konsistensi ini disebabkan oleh karakteristik mekanisme RMSProp yang menormalisasi gradien menggunakan rata-rata bergerak dari kuadrat gradien (Tieleman & Hinton, 2012), sehingga setiap parameter

mendapatkan learning rate efektif yang sesuai dengan magnitudo riwayat gradiennya. Pada dataset dengan tiga fitur sederhana yang sudah dinormalisasi seperti dalam penelitian ini, kombinasi tersebut memungkinkan konvergensi yang cepat tanpa perlu tuning learning rate yang ketat.

Adam menunjukkan performa yang setara dengan RMSProp dalam hal metrik klasifikasi pada model juaranya, namun dengan waktu eksekusi sedikit lebih lama (135,28 detik). Dari 8 konfigurasi Adam pada DT2, dua konfigurasi mencapai semua metrik sempurna: DT2_O2_LR2_E1_BS2 dan DT2_O2_LR2_E2_BS2, keduanya pada learning rate rendah 0,001 dan batch size 32. Konvergensi Adam yang cepat pada learning rate rendah dijelaskan oleh penggabungan momentum orde pertama dan orde kedua dengan koreksi bias inisialisasi (Kingma & Ba, 2015), yang efektif menstabilkan pembaruan bobot di fase awal pelatihan. Temuan ini selaras dengan laporan Wijaya et al. (2025) yang menempatkan RMSProp dan Adam pada peringkat teratas dibandingkan SGD pada arsitektur ANN dengan data terstruktur. Secara komparatif, performa metrik sempurna (1,000) yang dicapai oleh model juara ini menunjukkan peningkatan stabilitas yang signifikan bila disandingkan dengan penelitian terdahulu; misalnya riset Asriningtias et al. (2025) yang mencapai akurasi 99% dengan model ANN dasar, serta Ndagijimana et al. (2025) yang memperoleh AUC 0,84 menggunakan Multilayer Perceptron standar.

Perilaku SGD: Bukti Pentingnya Dua Intervensi Bersamaan

Hasil paling instruktif dalam penelitian ini terdapat pada perilaku SGD. Pada skenario data original (DT1), data yang telah dinormalisasi namun masih tidak seimbang, SGD gagal mendeteksi sampel stunting di hampir seluruh konfigurasi: tujuh dari delapan kombinasi menghasilkan recall 0,000, dengan rerata recall hanya 0,056. Satu-satunya konfigurasi yang berhasil mendeteksi sebagian stunting adalah DT1_O1_LR1_E2_BS1 (1.000 epoch, batch=16) dengan recall = 0,4444.

Sebaliknya, pada data balanced (DT2), seluruh konfigurasi SGD dengan learning rate 0,01 berhasil mencapai recall sempurna 1,000, dan model juaranya (DT2_O1_LR1_E2_BS2) mencapai F1-Score 0,900 dengan nol false negative. Rerata recall SGD pada skenario DT2 adalah 0,944, peningkatan 88,9 poin persentase dibandingkan skenario DT1 (0,056).

Perbandingan ini secara empiris mengonfirmasi bahwa SGD, sebagai optimizer non-adaptif, tidak memiliki mekanisme normalisasi gradien internal

seperti yang dimiliki Adam dan RMSProp, sehingga bergantung pada dua kondisi sekaligus: (1) normalisasi fitur untuk menyeimbangkan kontribusi gradien antar-dimensi, dan (2) keseimbangan kelas pada data latih untuk mencegah gradien didominasi oleh kelas mayoritas. Tanpa penyeimbangan kelas (skenario DT1), bahkan dengan normalisasi yang sudah benar, SGD masih gagal mendeteksi stunting secara konsisten. Implikasi praktis dari temuan ini adalah bahwa praktisi yang memilih SGD untuk klasifikasi medis berdistribusi tidak seimbang harus memastikan kedua intervensi tersebut diterapkan sebelum membuat keputusan mengenai kelayakan optimizer ini.

Catatan Kritis terhadap Skor Sempurna

Enam konfigurasi pada penelitian ini mencapai metrik sempurna (accuracy = F1 = MCC = 1,000) pada data uji: dua konfigurasi Adam (DT2_O2_LR2_E1_BS2 dan DT2_O2_LR2_E2_BS2) dan empat konfigurasi RMSProp (DT2_O3_LR1_E1_BS1 hingga DT2_O3_LR1_E2_BS2). Interpretasi atas hasil ini memerlukan beberapa pertimbangan statistik dan metodologis.

Pertama, ukuran kelas minoritas pada data uji relatif kecil (18 sampel stunting dari 291 total). Dengan $n = 18$, estimasi recall sempurna memiliki ketidakpastian yang cukup besar. Menggunakan Wilson score interval 95%, rentang sebenarnya recall berada pada [0,82; 1,00] — artinya, dengan populasi yang lebih besar, performa aktual berpotensi turun hingga 82% meski tetap sangat baik. Kedua, model juara Adam dan RMSProp menghasilkan vektor prediksi yang identik pada data uji (TP=18, TN=273, FP=0, FN=0), sehingga superioritas relatif keduanya tidak dapat dipisahkan secara empiris pada split ini. Keunggulan RMSProp atas Adam dalam penelitian ini hanya bersumber dari selisih waktu eksekusi 1,86 detik (133,42 vs 135,28 detik) — perbedaan yang tidak signifikan secara praktis.

Ketiga, mekanisme ModelCheckpoint dengan `save_best_only=True` memitigasi overfitting dengan menyimpan model pada epoch ber-validation accuracy terbaik. Ketiadaan data leakage pun terjamin karena MinMaxScaler dan SMOTETomek hanya disesuaikan pada data latih. Keempat, hasil yang dilaporkan merepresentasikan validasi internal (internal validation) pada satu random split tunggal (`random_state=42`), bukan model yang telah teruji secara eksternal. Generalisasi ke populasi yang lebih besar dan distribusi yang berbeda masih perlu dievaluasi melalui validasi eksternal pada dataset puskesmas lain atau dengan teknik stratified k-fold

cross-validation. Penelitian lanjutan yang direkomendasikan meliputi penerapan 5-fold stratified cross-validation, early stopping, serta pengujian pada dataset gabungan multi-puskesmas untuk memperkuat klaim robustness model.

Secara praktis, hasil penelitian ini memberikan implikasi penting bagi pengambil kebijakan kesehatan, khususnya di tingkat fasilitas pelayanan kesehatan dasar seperti UPT Puskesmas Kampung Baru Kota Medan. Model prediksi berbasis JST dengan optimizer RMSProp ini dapat direkomendasikan sebagai algoritma inti (engine) pada sistem informasi posyandu untuk keperluan deteksi dini (early warning system). Dengan kemampuan klasifikasi yang sangat presisi pada kondisi data minoritas, intervensi gizi spesifik dapat dialokasikan oleh tenaga kesehatan secara lebih proaktif dan tepat sasaran sebelum gangguan pertumbuhan pada balita menjadi permanen.

KESIMPULAN

Penelitian ini mengevaluasi 48 konfigurasi JST Backpropagation (arsitektur 3-32-1) untuk klasifikasi stunting balita menggunakan tiga optimizer: SGD, Adam, dan RMSProp. RMSProp terbukti memberikan performa terbaik dengan metrik klasifikasi sempurna (akurasi, presisi, recall, F1-Score, AUC, MCC bernilai 1,000) dan waktu eksekusi tercepat (133,42 detik). Adam menunjukkan metrik yang setara namun dengan waktu konvergensi sedikit lebih lambat. Kontribusi krusial penelitian ini adalah pembuktian empiris bahwa optimizer non-adaptif seperti SGD mutlak membutuhkan dua intervensi bersamaan—normalisasi MinMaxScaler (fit murni pada data latih) dan penyeimbangan kelas SMOTETomek—untuk bisa menyaingi performa optimizer adaptif. Tanpa keduanya, SGD terbukti gagal mendeteksi kelas minoritas secara konsisten.

Mengingat skor sempurna pada penelitian ini diestimasi melalui single test split dengan jumlah kelas minoritas yang terbatas (18 sampel uji), keterbatasan empiris model ini masih perlu divalidasi lebih lanjut. Sebagai implikasi praktis, model RMSProp ini layak diadopsi oleh fasilitas pelayanan kesehatan sebagai mesin deteksi dini stunting. Untuk penelitian selanjutnya, disarankan melakukan validasi eksternal pada dataset gabungan multi-puskesmas dengan distribusi demografis yang lebih luas, serta menerapkan metode stratified 5-fold cross-validation untuk menguji ketahanan (robustness) model secara lebih komprehensif.

DAFTAR PUSTAKA

- Anggara, D., Suarna, N., & Wijaya, Y. A. (2023). Analisa perbandingan performa optimizer Adam, SGD, dan RMSProp pada model H5. *Networking Engineering Research Operation*, 8(1), 53–64. <https://doi.org/10.21107/nero.v8i1.19226>
- Asriningtias, S. R., Megawati, C. D., Kusumaningtyas, D., & Surya, D. U. (2025). Hyperparameter-tuned artificial neural networks for early stunting prediction in toddlers. *Sinkron*, 9(2), 832–840. <https://doi.org/10.33395/sinkron.v9i2.14695>
- Batista, G. E. A. P. A., Prati, R. C., & Monard, M. C. (2004). A study of the behavior of several methods for balancing machine learning training data. *ACM SIGKDD Explorations Newsletter*, 6(1), 20–29. <https://doi.org/10.1145/1007730.1007735>
- Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: Synthetic minority over-sampling technique. *Journal of Artificial Intelligence Research*, 16, 321–357. <https://doi.org/10.1613/jair.953>
- Chicco, D., & Jurman, G. (2020). The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. *BMC Genomics*, 21(1), 6. <https://doi.org/10.1186/s12864-019-6413-7>
- Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT Press.
- Haykin, S. (2009). *Neural networks and learning machines* (3rd ed.). Pearson Education.
- Kementerian Kesehatan Republik Indonesia. (2020). *Peraturan Menteri Kesehatan Republik Indonesia Nomor 2 Tahun 2020 tentang Standar Antropometri Anak* [Berita Negara Republik Indonesia Tahun 2020 Nomor 113].
- Kementerian Kesehatan Republik Indonesia. (2022). *Buku saku hasil Survei Status Gizi Indonesia (SSGI) 2022* [Laporan Survei]. Kementerian Kesehatan Republik Indonesia. <https://www.badankebijakan.kemkes.go.id/>
- Kingma, D. P., & Ba, J. (2015). Adam: A method for stochastic optimization. *3rd International Conference on Learning Representations (ICLR)*. <https://arxiv.org/abs/1412.6980>
- Nair, V., & Hinton, G. E. (2010). Rectified linear units improve restricted Boltzmann machines. *Proceedings of the 27th International Conference on Machine Learning (ICML-10)*, 807–814.
- Ndagijimana, S., Kabano, I., Masabo, E., & Ntaganda, J. M. (2025). Predicting stunting status among under-5 children in Rwanda using neural network model: Evidence from 2020 Rwanda demographic and health survey. *F1000Research*, 13, 128.

<https://doi.org/10.12688/f1000research.141458.2>

- Raihana, H. A., Gunawan, P. H., & Aquarini, N. (2025). Breaking class imbalance: Machine learning solutions for stunting detection. *Lontar Komputer: Jurnal Ilmiah Teknologi Informasi*, 16(2), 03. <https://doi.org/10.24843/LKJITI.2025.v16.i02.p03>
- Ruder, S. (2016). An overview of gradient descent optimization algorithms. *arXiv preprint arXiv:1609.04747*. <https://arxiv.org/abs/1609.04747>
- Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors. *Nature*, 323(6088), 533–536. <https://doi.org/10.1038/323533a0>
- Tieleman, T., & Hinton, G. (2012). *Lecture 6.5 — RMSProp: Divide the gradient by a running average of its recent magnitude* [Course lecture notes]. COURSE: Neural Networks for Machine Learning, University of Toronto.
- World Health Organization. (2014). *Global nutrition targets 2025: Stunting policy brief* (WHO/NMH/NHD/14.3) [Policy Brief]. World Health Organization. <https://www.who.int/publications/>
- Wijaya, E., Soeleman, M. A., & Andono, P. N. (2025). Comparative performance analysis of optimization algorithms in artificial neural networks for stock price prediction. *Journal of Applied Informatics and Computing*, 9(1), 23–30. <https://doi.org/10.30871/jaic.v9i1.8820>
- Yunus, M., Biddinika, M. K., & Fadlil, A. (2025). Comparison of machine learning algorithms for stunting classification. *Scientific Journal of Engineering Research*, 1(2), 64–70. <https://doi.org/10.64539/sjer.v1i2.2025.9>