

**KLASIFIKASI STATUS GIZI BERBASIS SISTEM SPASIAL
MENGUNAKAN *ARTIFICIAL NEURAL NETWORK*
DAN *RANDOM FOREST***

TESIS

**Oleh:
DESI ANIS ANGGRAINI
220605220006**



**PROGRAM STUDI MAGISTER INFORMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM
MALANG
2025**

**KLASIFIKASI STATUS GIZI BERBASIS SISTEM SPASIAL
MENGUNAKAN *ARTIFICIAL NEURAL NETWORK*
DAN *RANDOM FOREST***

TESIS

**Diajukan Kepada:
Universitas Islam Negeri Maulana Malik Ibrahim Malang
Untuk Memenuhi Salah Satu Persyaratan Dalam
Memperoleh Gelar Magister Komputer (M.Kom)**

**Oleh:
DESI ANIS ANGGRAINI
NIM. 220605220006**

**PROGRAM STUDI MAGISTER INFORMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM
MALANG
2025**

**KLASIFIKASI STATUS GIZI BERBASIS SISTEM SPASIAL
MENGUNAKAN *ARTIFICIAL NEURAL NETWORK*
DAN *RANDOM FOREST***

TESIS

**Diajukan Kepada:
Fakultas Sains dan Teknologi
Universitas Islam Negeri Maulana Malik Ibrahim Malang
Untuk Memenuhi Salah Satu Persyaratan Dalam
Memperoleh Gelar Magister Komputer (M.Kom)**

**Oleh:
DESI ANIS ANGGRAINI
NIM. 220605220006**

**PROGRAM STUDI MAGISTER INFORMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM
MALANG
2025**

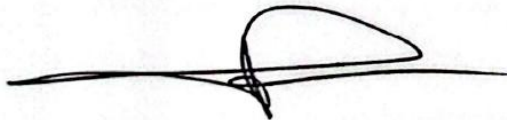
**KLASIFIKASI STATUS GIZI BERBASIS SISTEM SPASIAL
MENGUNAKAN *ARTIFICIAL NEURAL NETWORK*
DAN *RANDOM FOREST***

TESIS

Oleh:
DESI ANIS ANGGRAINI
NIM. 220605220006

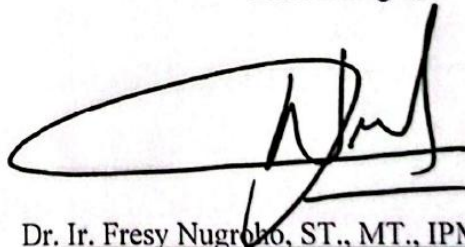
Telah diperiksa dan disetujui untuk diuji:
Tanggal: 06 November 2025

Pembimbing I,



Dr. Ir. Fachrul Kurniawan ST., M.MT., IPU
NIP. 19771020 200912 1 001

Pembimbing II,



Dr. Ir. Fresy Nugroho, ST., MT., IPM., ASEAN Eng
NIP. 19710722 201101 1 001

Mengetahui,
Ketua Program Studi Magister Informatika
Fakultas Sains dan Teknologi
Universitas Islam Negeri Maulana Malik Ibrahim Malang



Prof. Dr. Ir. Muhammad Faisal, M.T
NIP. 19740510 200501 1 007

**KLASIFIKASI STATUS GIZI BERBASIS SISTEM SPASIAL
MENGUNAKAN *ARTIFICIAL NEURAL NETWORK*
DAN *RANDOM FOREST***

TESIS

**Oleh:
DESI ANIS ANGGRAINI
NIM. 220605220006**

Telah dipertahankan di Depan Dewan Penguji Thesis
dan Dinyatakan Diterima Sebagai Salah Satu Persyaratan
Untuk Memperoleh Gelar Magister Komputer (M.Kom)
Tanggal: 06 November 2025

Susunan Dewan Penguji

Penguji I : Prof. Dr. Ir. Muhammad Faisal, M.T
NIP. 19740510 200501 1 007

Penguji II : Dr. Cahyo Crysdian, M.Cs
NIP. 19740424 200901 1 008

Pembimbing I : Dr. Ir. Fachrul Kurniawan ST., M.MT.,
IPU
NIP. 19771020 200912 1 001

Pembimbing II : Dr. Ir. Fresy Nugroho, ST., MT., IPM.,
ASEAN Eng
NIP. 19710722 201101 1 001

Tanda tangan

()

()

()

Mengetahui,
Ketua Program Studi Magister Informatika
Fakultas Sains dan Teknologi
Universitas Islam Negeri Maulana Malik Ibrahim Malang



Prof. Dr. Ir. Muhammad Faisal, M.T
NIP. 19740510 200501 1 007

PERNYATAAN KEASLIAN TULISAN

Saya yang bertanda tangan di bawah ini:

Nama : Desi Anis Anggraini
NIM : 220605220006
Program Studi : Magister Informatika
Fakultas : Sains dan Teknologi

Menyatakan dengan sebenarnya bahwa Thesis yang saya tulis ini benar-benar merupakan hasil karya saya sendiri, bukan merupakan pengambilalihan data, tulisan atau pikiran orang lain yang saya akui sebagai tulisan atau pikiran saya sendiri, kecuali dengan mencantumkan sumber cuplikan pada daftar pustaka.

Apabila di kemudian hari terbukti atau dapat dibuktikan Thesis ini hasil jiplakan, maka saya bersedia menerima sanksi atas perbuatan tersebut.

Malang, 15 Desember 2025

Yang Menyatakan,



Desi Anis Anggraini
NIM. 220605220006

MOTTO

“Keberhasilan bukan sekedar akhir dari perjalanan, melainkan hasil dari keteguhan hati dalam menghadapi ujian dan pengorbanan”

PERSEMBAHAN

Dengan Segenap rasa Syukur Allah SWT, tesis ini saya persembahkan kepada:

1. Yang pertama, saya berterima kasih kepada diri saya sendiri, karena sudah berjuang sejauh ini. Perjalanan ini sungguh sangat tidak mudah, banyak sekali yang dikorbankan, waktu, tenaga, pikiran dan materi. Akan tetapi diri ini bisa melewati semua dengan baik.
2. Seluruh keluarga saya (ibu, bapak, suami, anak, ibu mertua, dan adek, serta seluruh keluarga besar saya) yang selalu memberikan doa, dukungan, dan semangat tanpa henti.
3. Dosen pembimbing dan seluruh pengajar, yang telah sabar membimbing, memberi ilmu, dan berbagi pengalaman yang begitu berharga, sehingga saya bisa belajar lebih banyak dalam bidang ilmu ini.
4. Rekan-rekan Angkatan 7 Magister Informatika Fakultas Sains dan Teknologi Universitas Islam Negeri Maulana Malik Ibrahim Malang yang telah mendukung, berbagi pengetahuan, ide, dan dukungan moral selama ini.
5. Rektor Universitas Bahaudin Mudhary Madura yang telah memberikan kesempatan untuk melanjutkan pendidikan serta mendukung penelitian ini.

Semoga karya ini dapat bermanfaat dan menjadi langkah awal bentuk kontribusi nyata bagi lingkungan dan ilmu pengetahuan.

KATA PENGANTAR

Assalamu 'alaikum warahmatullah wabarakatuh

Puji syukur, alhamdulillah, kehadiran Allah SWT atas limpahan rahmat dan kasih sayang-Nya, sehingga penulis dapat menyelesaikan studi di Program Studi Magister Informatika Fakultas Sains dan Teknologi Universitas Islam Negeri Maulana Malik Ibrahim Malang sekaligus menyelesaikan Thesis ini dengan baik. Penulis mengucapkan terima kasih sebesar-besarnya kepada semua pihak yang telah membantu terselesaikannya Thesis ini. Ucapan terima kasih ini penulis sampaikan kepada:

1. Bapak Dr. Fachrul Kurniawan dan Bapak Dr. Fresy Nugroho selaku dosen pembimbing Thesis atas arahan dan motivasi yang telah diberikan,
2. Segenap civitas akademika Program Studi Magister Informatika atas ilmu dan bimbingannya,
3. Keluarga tercinta atas dukungan dan doa yang selalu dilantarkan, dan
4. Semua rekan-rekan Magister Informatika, baik angkatan 7 maupun senior junior yang berkenan berbagi pengalaman dan memberi masukan.

Penulis menyadari bahwa penyusunan Thesis ini masih mengandung kekurangan. Meski demikian, penulis berharap semoga Thesis ini dapat memberi manfaat kepada para pembaca, khususnya pihak-pihak yang menjadi tujuan penyelenggaraan penelitian ini, dan penulis secara pribadi.

Wassalamu 'alaikum warahmatullah wabarakatuh.

Malang, 15 Desember 2025

Penulis

DAFTAR ISI

HALAMAN SAMPUL	i
HALAMAAN JUDUL	ii
HALAMAN PENGAJUAN	iii
HALAMAN PENGESAHAN	v
HALAMAN PERNYATAAN	vi
MOTTO	vii
HALAMAN PERSEMBAHAN	viii
KATA PENGANTAR	ix
DAFTAR ISI	x
DAFTAR GAMBAR	xii
DAFTAR TABEL	xiv
ABSTRAK	xvii
ABSTRACT	xviii
مستخلص البحث	xix
BAB 1 PENDAHULUAN	1
1.1. Latar Belakang	1
1.2. Pernyataan Masalah	2
1.3. Tujuan Penelitian.....	3
1.4. Batasan Masalah.....	3
1.5. Manfaat Penelitian	3
BAB II STUDI PUSTAKA	5
2.1. Pengukuran Status Gizi	5
BAB III METODOLOGI PENELITIAN	15
3.1. Desain Penelitian.....	15
3.2. Instrumen Penelitian.....	26
3.3. Miniatur Penelitian (Simulasi)	27
BAB IV KLASIFIKASI STATUS GIZI DENGAN ANN	36
4.1. Arsitektur ANN dan Parameter Pelatihan	36
4.2. Konfigurasi Pelatihan Jaringan Neural (Training Configuration).....	37
4.3. Hyperparameter Tuning Grid	39
4.4. Skenario Pelatihan Model ANN.....	41

BAB V KLASIFIKASI STATUS GIZI DENGAN RANDOM FOREST	49
5.1. Arsitektur Random Forest.....	49
5.2. Hyperparameter Random Forest	50
5.3. Hyperparameter Tuning Grid	52
5.4. Random Forest Optimal Configuration.....	54
5.5. Skenario Pelatihan Model	57
BAB VI ANALISIS DATA SPASIAL	65
6.1. Analisis Status Gizi Per-Kecamatan	65
6.2. Distribusi Status Gizi per Kecamatan	83
6.3. Persentase Gizi Baik per Kecamatan	84
6.4. Pemetaan Zona Status Gizi Kabupaten Sumenep	84
BAB VII PENUTUP	87
7.1. Kesimpulan	87
7.2. Saran.....	88
DAFTAR PUSTAKA.....	89

DAFTAR GAMBAR

Gambar 2. 2 Model <i>Perceptron</i>	5
Gambar 2. 3 Kerangka Teoritis	13
Gambar 3. 1 Desain Penelitian.....	15
Gambar 3. 2 Validasi Data	16
Gambar 3. 3 Desain Sistem.....	18
Gambar 3. 4 Diagram alir algoritma metode ANN	21
Gambar 3. 5 Flowchart Klasifikasi RF	22
Gambar 4. 1 Hyperparamter ANN	41
Gambar 6. 1 Peta hasil klasifikasi status gizi kec Sumenep	65
Gambar 6. 2 Peta hasil klasifikasi status gizi kec. Kalianget.....	66
Gambar 6. 3 Peta hasil klasifikasi status gizi kec. Manding.....	67
Gambar 6. 4 Peta hasil klasifikasi status gizi kec. Talango.....	68
Gambar 6. 5 Peta hasil klasifikasi status gizi kec. Bluto	68
Gambar 6. 6 Peta hasil klasifikasi status gizi kec. Saronggi.....	69
Gambar 6. 7 Peta hasil klasifikasi status gizi kec. Lenteng	70
Gambar 6. 8 Peta hasil klasifikasi status gizi kec. Giligenting	70
Gambar 6. 9 Peta hasil klasifikasi status gizi kec. Guluk-guluk.....	71
Gambar 6. 10 Peta hasil klasifikasi status gizi kec. Ganding.....	71
Gambar 6. 11 Peta hasil klasifikasi status gizi kec. Pragaan.....	72
Gambar 6. 12 Peta hasil klasifikasi status gizi kec. Ambunten.....	73
Gambar 6. 13 Peta hasil klasifikasi status gizi kec. Pasongsongan	73
Gambar 6. 14 Peta hasil klasifikasi status gizi kec. Dasuk	74
Gambar 6. 15 Peta hasil klasifikasi status gizi kec. Rubaru.....	75

Gambar 6. 16 Peta hasil klasifikasi status gizi kec. Batang-batang.....	75
Gambar 6. 17 Peta hasil klasifikasi status gizi kec. Batuputih.....	76
Gambar 6. 18 Peta hasil klasifikasi status gizi kec. Dungkek.....	77
Gambar 6. 19 Peta hasil klasifikasi status gizi kec. Gapura.....	77
Gambar 6. 20 Peta hasil klasifikasi status gizi kec. Gayam.....	78
Gambar 6. 21 Peta hasil klasifikasi status gizi kec. Nonggunong	79
Gambar 6. 22 Peta hasil klasifikasi status gizi kec. Masalembu.....	80
Gambar 6. 23 Peta hasil klasifikasi status gizi kec. Arjasa	80
Gambar 6. 24 Peta hasil klasifikasi status gizi kec. Sapeken.....	81
Gambar 6. 25 Peta hasil klasifikasi status gizi kec. Sapeken.....	82
Gambar 6. 26 Peta hasil klasifikasi status gizi kec. Kangayan	82
Gambar 6. 27 Distribusi Gizi Per Kecamatan	83
Gambar 6. 28 Persentase Gizi Baik Per Kecamatan	84
Gambar 6. 29 Peta Resiko Gizi Kurang (Model Terbaik).....	85

DAFTAR TABEL

Tabel 3. 1 Data Gizi	16
Tabel 3. 2 Data Awal	17
Tabel 3. 3 Parameter Validasi Data	17
Tabel 3. 4 Prosedur Penyelesaian Permasalahan Data	17
Tabel 3. 5 Klasifikasi Zona	24
Tabel 3. 6 Instrumen Penelitian.....	27
Tabel 3. 7 Data Status Gizi.....	27
Tabel 3. 8 Normalisasi dan Percobaan Iterasi ke-1	28
Tabel 3. 9 Looping Iterasi ke-2	29
Tabel 3. 10 Looping Iterasi ke-3	29
Tabel 3. 11 Looping Iterasi ke -4	30
Tabel 3. 12 Looping Iterasi ke-5	31
Tabel 3. 13 Pohon ke-1 berdasarkan IMT	32
Tabel 3. 14 Pohon ke-2 berdasarkan Umur/Usia	32
Tabel 3. 15 Pohon ke-3 berdasarkan Berat Badan (BB)	33
Tabel 3. 16 Ringkasan Klasifikasi.....	33
Tabel 3. 17 Zonasi Resiko Gizi Balita	34
Table 4. 1 Tabel Arsitektur Jaringan Neural Network.....	36
Table 4. 2 Konfigurasi Optimizer Model ANN.....	38
Table 4. 3 Parameter Pelatihan Model ANN	38
Table 4. 4 Teknik Regularisasi Model ANN	38
Table 4. 5 Variasi Learning Rate yang diuji	39
Table 4. 6 Variasi Dropout Rate yang diuji	40

Table 4. 7 Arsitektur Hidden layer yang diuji	40
Table 4. 8 Fungsi Aktivasi yang diuji.....	41
Table 4. 9 parameter Terbaik ANN Skenario Awal (80-20).....	42
Table 4. 10 Hasil pelatihan model ANN per-epoch	42
Table 4. 11 Hasil Klasifikasi ANN Skenario Awal (80-20).....	43
Table 4. 12 Parameter Terbaik ANN Skenario-1 (60-40).....	44
Table 4. 13 Hasil pelatihan model ANN per-epoch Skenario-1	44
Table 4. 14 Hasil Klasifikasi ANN Skenario-1 (60-30)	45
Table 4. 15 parameter Terbaik ANN Skenario-2 (70-30).....	46
Table 4. 16 Hasil pelatihan model ANN per-epoch Skenario-2	46
Table 4. 17 Hasil Klasifikasi ANN Skenario-2	47
Table 4. 18 Perbandingan performa setiap skenario	47
Tabel 5. 1 Parameter Base Estimator pada Model Random Forest.....	49
Tabel 5. 2 Parameter Ensemble Random Forest	49
Tabel 5. 3 Parameter Struktur Pohon yang Diuji	50
Tabel 5. 4 Opsi Bobot Kelas Random Forest	51
Tabel 5. 5 Parameter Pelatihan Random Forest	51
Tabel 5. 6 Jumlah Estimator yang Diuji.....	52
Tabel 5. 7 Max Depth yang Diuji.....	52
Tabel 5. 8 Min Samples Split yang Diuji	53
Tabel 5. 9 Min Samples Leaf yang Diuji	53
Tabel 5. 10 Max Features yang Diuji	54
Tabel 5. 11 Opsi Bootstrap	54
Tabel 5. 12 Konfigurasi Optimal Model Random Forest.....	55

Tabel 5. 13 Skenario Pelatihan Model Random Forest.....	57
Tabel 5. 14 Hasil Evaluasi Cross-Validation Random Forest	58
Tabel 5. 15 Tabel Hasil Klasifikasi ANN per Kecamatan.....	58
Tabel 5. 16 Parameter Terbaik Random Forest Skenario - 1	59
Tabel 5. 17 Hasil Evaluasi Cross-Validation Random Forest	60
Tabel 5. 18 Tabel Hasil Klasifikasi Random Forest per Kecamatan.....	61
Tabel 5. 19 Parameter Terbaik Random Forest Skenario - 1	62
Tabel 5. 20 Hasil Evaluasi Cross-Validation Random Forest Skenario – 2	62
Tabel 5. 21 Tabel Hasil Klasifikasi Random Forest per Kecamatan.....	63
Tabel 5. 22 Perbandingan Hasil Kinerja ANN dan Random Forest.....	64

ABSTRAK

Anggraini, Desi Anis, 2025. **Klasifikasi Status Gizi Berbasis Sistem Spasial Menggunakan Artificial Neural Network Dan Random Forest.** Thesis. Program Studi Magister Informatika Fakultas Sains dan Teknologi Universitas Islam Negeri Maulana Malik Ibrahim Malang, Pembimbing: (I) Dr. Ir. Fachrul Kurniawan ST., M.MT., IPU, (II) Dr. Ir. Fresy Nugroho, ST., MT, IPM.

Kata kunci: Klasifikasi status gizi, sistem spasial, Artificial Neural Network, Random Forest.

Masalah gizi pada anak balita masih menjadi tantangan kesehatan masyarakat yang besar. Penelitian ini bertujuan untuk mengembangkan sistem berbasis spasial untuk menilai dan memetakan status gizi menggunakan algoritma Artificial Neural Network (ANN) dan Random Forest (RF). Data yang diperoleh dari Dinas Kesehatan Kabupaten Sumenep meliputi variabel usia, berat badan, tinggi badan, dan jenis kelamin. Kedua model dilatih menggunakan rasio data 70:30 dan dievaluasi dengan metrik akurasi, presisi, recall, dan skor F1. Model JST mencapai akurasi 95,8%, sedangkan model RF mencapai 98,7%. Hasil klasifikasi divisualisasikan melalui Sistem Informasi Geografis (SIG) untuk menggambarkan distribusi spasial dan mengidentifikasi zona berisiko tinggi. Integrasi pembelajaran mesin dan analisis spasial terbukti efektif dalam meningkatkan akurasi klasifikasi, memperbaiki interpretasi data, dan mendukung kebijakan gizi berbasis data serta pengambilan keputusan kesehatan daerah.

ABSTRACT

Anggraini, Desi Anis, 2025. **Spatial System-Based Nutritional Status Classification Using Artificial Neural Network and Random Forest**. Theses. Program Studi Magister Informatika Fakultas Sains dan Teknologi Universitas Islam Negeri Maulana Malik Ibrahim Malang. Promotor: (I) Dr. Ir. Fachrul Kurniawan ST., M.MT., IPU, (II) Dr. Ir. Fresy Nugroho, ST., MT, IPM.

Nutritional problems among children under five remain a major public health challenge. This study aims to develop a spatial-based system for assessing and mapping nutritional status using Artificial Neural Network (ANN) and Random Forest (RF) algorithms. Data obtained from the Sumenep District Health Office included age, weight, height, and gender variables. Both models were trained using a 70:30 data ratio and evaluated with accuracy, precision, recall, and F1-score metrics. The ANN model achieved an accuracy of 95.8%, while the RF model reached 98.7%. Classification results were visualized through a Geographic Information System (GIS) to illustrate spatial distribution and identify high-risk zones. The integration of machine learning and spatial analysis proved effective in enhancing classification accuracy, improving data interpretation, and supporting data-driven nutritional policy and regional health decision-making.

Key words: Nutritional status classification, spatial system, Artificial Neural Network, Random Forest.

مستخلص البحث

أنغريني، ديسي أنيس، 2025. تصنيف حالة الغذائية بناءً على النظام المكاني باستخدام الشبكات العصبية الاصطناعية وغابة عشوائية. رسالة الماجستير. قسم المعلومات، كلية العلوم والتكنولوجيا بجامعة مولانا مالك إبراهيم الإسلامية الحكومية مالانج، المشرف الأول: د. فخر الكورنياوان، الماجستير؛ المشرف الثاني: د. فريسي نوجروهو، الماجستير.

الكلمات الرئيسية: تصنيف حالة غذائية، نظام مكاني، شبكات عصبية اصطناعية، غابة عشوائية.

ظل مشكلة سوء التغذية لدى الأطفال دون سن الخامسة تحديًا كبيرًا للصحة العامة. تهدف هذه الرسالة إلى تطوير نظام قائم على التحليل المكاني لتقييم ورسم خرائط الحالة الغذائية باستخدام خوارزمية الشبكة العصبية الاصطناعية (ANN) والغابة العشوائية (RF). تشمل البيانات المستمدة من مكتب الصحة في محافظة سمنب المتغيرات التالية: العمر، الوزن، الطول، والجنس. تم تدريب كلا النموذجين باستخدام نسبة بيانات 70:30 وتم تقييمهما باستخدام مقاييس الدقة، الوضوح، الاستدعاء، ودرجة ف1. حقق نموذج ANN دقة بلغت 95.8٪، بينما حقق نموذج RF دقة بلغت 98.7٪. تم تصوير نتائج التصنيف من خلال نظام المعلومات الجغرافية (GIS) لتوضيح التوزيع المكاني وتحديد المناطق عالية المخاطر. لقد ثبت أن تكامل التعلم الآلي والتحليل المكاني فعال في تحسين دقة التصنيف، وتحسين تفسير البيانات، ودعم سياسات الغذائية المبنية على البيانات واتخاذ القرارات الصحية على مستوى المنطقة.

BAB 1

PENDAHULUAN

1.1. Latar Belakang

Status gizi balita merupakan elemen penting dalam pencapaian Tujuan Pembangunan Berkelanjutan (SDGs), terutama Good Health and Well-Being yang menekankan kehidupan sehat dan sejahtera bagi semua usia (PBB, 2015). Kondisi gizi yang tidak optimal seperti stunting, wasting, dan underweight berdampak pada kesehatan dan perkembangan kognitif anak, sehingga diperlukan metode penentuan status gizi yang cepat dan akurat.

Inovasi teknologi telah menjadi pendorong penting dalam mendukung pembangunan berkelanjutan. Konsep Smart Green Campus yang dikembangkan oleh Fachrul Kurniawan et al. (2024) menunjukkan bahwa pemanfaatan teknologi dapat memperkuat pencapaian tujuan SDGs, termasuk di bidang kesehatan. Dalam konteks gizi balita, kecerdasan buatan dapat digunakan untuk meningkatkan ketepatan klasifikasi status gizi.

Beberapa studi membuktikan efektivitas kecerdasan buatan dalam klasifikasi status gizi. Shahriar et al. (2019) melaporkan akurasi tinggi pada Artificial Neural Network (93%) dan Random Forest (89,58%), menunjukkan kedua algoritma ini layak digunakan sebagai model prediksi. Namun, penelitian sebelumnya masih jarang mengintegrasikan analisis spasial, padahal faktor geografis turut memengaruhi variasi kondisi gizi.

Untuk mengisi gap tersebut, penelitian ini mengembangkan model klasifikasi status gizi balita menggunakan algoritma ANN dan Random Forest yang

kemudian divisualisasikan dalam Sistem Informasi Geografis (SIG). Dengan menggunakan dataset asli dari Dinas Kesehatan sebanyak 1.260 balita (tahun 2023–2024), penelitian ini bertujuan memberikan hasil klasifikasi yang akurat sekaligus representasi spasial yang bermanfaat bagi pengambil kebijakan kesehatan.

Dalam perspektif Islam, upaya penentuan status gizi balita dan pemetaan kondisi kesehatan menggunakan Sistem Informasi Geografis (SIG) merupakan bagian dari pelaksanaan nilai-nilai dasar syariat seperti keadilan, amanah, dan ihsan. Hal ini sejalan dengan firman Allah dalam QS. An-Nahl ayat 90:

إِنَّ اللَّهَ يَأْمُرُ بِالْعَدْلِ وَالْإِحْسَانِ وَإِيتَاءِ ذِي الْقُرْبَىٰ وَيَنْهَىٰ عَنِ الْفَحْشَاءِ وَالْمُنْكَرِ وَالْبَغْيِ ۚ يَعِظُكُمْ لَعَلَّكُمْ تَذَكَّرُونَ

yang memerintahkan untuk berlaku adil dan berbuat kebaikan kepada sesama. Hadis Nabi Muhammad SAW juga menegaskan tanggung jawab setiap pemimpin terhadap kesejahteraan yang dipimpinnya (HR. Bukhari dan Muslim) serta pentingnya menjaga kelompok lemah dan tidak menelantarkannya (HR. Abu Dawud; an-Nasai). Dengan demikian, penggunaan teknologi untuk pemantauan gizi tidak hanya mendukung peningkatan kualitas layanan kesehatan, tetapi juga mencerminkan komitmen terhadap prinsip moral Islam dalam mewujudkan kesejahteraan masyarakat.

1.2. Pernyataan Masalah

Pernyataan masalah yang akan di bahas pada penelitian ini adalah:

1. Bagaimana menentukan status gizi balita secara lebih akurat dengan memanfaatkan data spasial menggunakan metode Artificial Neural Network (ANN) dan Random Forest?
2. Bagaimana mengevaluasi hasil klasifikasi ke dalam sistem spasial tingkat kecamatan?

1.3. Tujuan Penelitian

Tujuan dari penelitian ini adalah:

1. Mengembangkan model penentuan status gizi balita berbasis data spasial menggunakan metode Artificial Neural Network (ANN) dan Random Forest.
2. Mengevaluasi hasil kinerja akurasi algoritma dalam merepresentasikan data hasil pengelompokan daerah rawan gizi kurang.

1.4. Batasan Masalah

Batasan masalah dari penelitian ini adalah sebagai berikut:

1. Data yang digunakan berasal dari Dinas Kesehatan kab. Sumenep.
2. Proses data dibatasi pada pembersihan, normalisasi, dan penggabungan atribut-spasial.
3. Pelatihan ANN dan Random Forest terdapat tiga skenario: split data (80-20), split data (70-30), dan split data (60-40).
4. Variabel input untuk pemetaan zona meliputi jumlah balita dengan status gizi kurang dan gizi baik di setiap wilayah.

1.5. Manfaat Penelitian

Manfaat dari penelitian ini adalah sebagai berikut:

1. Bagi peneliti, ini bisa menjadi referensi untuk menerapkan metode ANN dan *Random Forest* dalam sistem penentuan status gizi berbasis spasial.
2. Bagi pengguna, sistem penentuan status gizi berbasis spasial, memberikan metode alternatif serta akses untuk menganalisis dan memetakan pengelompokan wilayah.

3. Bagi pemerintah, sistem penentuan status gizi berbasis spasial dapat membantu dalam menentukan status gizi dan memantau daerah-daerah yang berisiko mengalami kekurangan gizi, serta membantu menemukan titik wilayah yang benar-benar memerlukan perhatian khusus, guna menunjang kualitas SDM di masa depan.

BAB II

STUDI PUSTAKA

2.1. Pengukuran Status Gizi

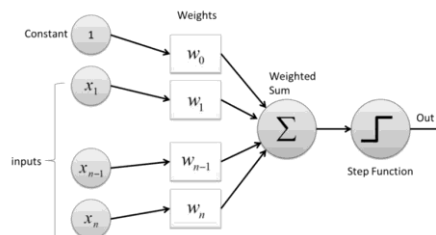
Status gizi balita ditentukan berdasarkan indikator antropometri yang digunakan oleh WHO, meliputi berat badan menurut umur (BB/U), tinggi badan menurut umur (TB/U), dan berat badan menurut tinggi badan (BB/TB). Penilaian menggunakan Z-score, yaitu selisih nilai pengukuran individu terhadap median populasi referensi dibagi simpangan baku.

2.1.1 ANN (*Artificial Neural Network*)

Ghwanmeh et al. (2013) menyatakan bahwa ANN meniru kerja saraf biologis melalui neuron yang saling terhubung. Setiap neuron memproses input menggunakan penjumlahan berbobot dan fungsi non-linier untuk menghasilkan output. Model dasar ANN yang umum digunakan adalah perceptron.

Model perceptron memberikan bobot pada setiap input, menjumlahkannya, lalu memproses hasilnya melalui fungsi aktivasi nonlinier. Kelebihan utamanya adalah mampu merepresentasikan pola nonlinier yang sering muncul pada berbagai fenomena fisik.

. Model *perceptron* dapat dilihat pada ilustrasi di Gambar 2.1.



Gambar 2. 1 Model *Perceptron*

Awalnya, perceptron dianggap menjanjikan karena mampu mempelajari pola fisik. Namun, ia memiliki keterbatasan, misalnya tidak dapat menyelesaikan masalah XOR. Dalam ANN, setiap lapisan berfungsi memecahkan masalah atau menghasilkan fitur baru dari kombinasi fitur lapisan sebelumnya.

Kishore et al. (2023) meneliti identifikasi dini malnutrisi pada bayi baru lahir menggunakan data UNICEF dengan indikator WAZ (weight-for-age) dan HAZ (height-for-age). Lima algoritma ML diuji: SVM, KNN, Logistic Regression, Naïve Bayes, dan Bilayered Neural Network. Hasil menunjukkan SVM paling akurat untuk WAZ (97,4%) dan HAZ (96,2%), sedangkan Bilayered Neural Network dengan aktivasi sigmoid mencapai 96,4% (WAZ) dan 97,9% (HAZ). Temuan ini menegaskan efektivitas machine learning untuk deteksi dini malnutrisi dan potensi penggunaannya dalam mendukung keputusan kesehatan.

Nel et al. (2022) mengembangkan alat skrining otomatis berbasis ANN untuk mendeteksi risiko malnutrisi akut berat (SAM) menggunakan kurva pertumbuhan WFA. ANN dilatih dengan penilaian 30 ahli dan divalidasi pada 185 anak (<5 tahun). Hasil menunjukkan sensitivitas 73,0%, spesifisitas 86,1%, dan ROC-AUC 0,795, lebih unggul dibanding metode konvensional. Temuan ini menunjukkan ANN efektif untuk deteksi dini anak berisiko SAM.

Mutammimul et al. (2022) mengklasifikasi status gizi balita di Aceh menggunakan K-Nearest Neighbor (K-NN) berdasarkan usia, berat badan, tinggi badan, dan jenis kelamin. Data dari RSUD Cut Meutia diolah melalui tahapan KDD, dengan klasifikasi menggunakan nilai Euclidean distance untuk menentukan malnutrisi dan gizi baik. Model K-NN berhasil mengelompokkan W/U, TB/U, dan

W/TB dengan akurasi 86%, menunjukkan efektivitasnya dalam mendukung deteksi dini malnutrisi di tingkat lokal.

Fenta et al. (2021) meneliti klasifikasi status gizi kurang pada balita di Ethiopia menggunakan enam algoritma ML: Logistic Regression, Ridge, LASSO, Elastic Net, ANN, dan Random Forest. Data berasal dari survei nasional 2000–2016. Hasil menunjukkan Random Forest terbaik dengan akurasi 68,8% dan AUC 76,1%. Faktor utama yang memengaruhi status gizi adalah lokasi tempat tinggal, literasi orang tua, dan IMT ibu, menunjukkan potensi ML untuk mendukung kebijakan intervensi gizi.

2.1.2 RF (*Random Forest*)

Random Forest adalah teknik machine learning yang menggabungkan banyak pohon keputusan untuk meningkatkan akurasi klasifikasi (Mantero & Ishwaran, 2021). Metode ini merupakan pengembangan dari CART, di mana setiap pohon dibangun dari kombinasi data dan atribut acak. Keputusan akhir ditentukan melalui voting mayoritas dari seluruh pohon. Pembentukan pohon dimulai dengan pemilihan atribut akar berdasarkan nilai gain dari entropi, dan berlanjut hingga semua data termasuk dalam kelas yang sama.

Yanto et al. (2024) meneliti klasifikasi status gizi anak menggunakan kombinasi pra-pemrosesan K-Means, Pearson Correlation, dan SSE, serta klasifikasi ANN dan Decision Tree. Data berasal dari RSUP M. Djamil Padang dengan parameter jenis kelamin, usia, berat badan, dan tinggi badan. Hasil menunjukkan ANN multi-lapisan (4-5-8-1) terbaik dengan akurasi 99,23% dan MSE 0,77%, menegaskan efektivitas model dalam menentukan status gizi anak secara otomatis.

Yunus et al. (2023) mengembangkan pendekatan hybrid ensemble berbasis majority voting (MVBHE) untuk mengklasifikasi malnutrisi anak <5 tahun di sub-Sahara Afrika menggunakan data DHS. Algoritma ini menggabungkan Random Forest, Decision Tree, XGBoost, dan KNN. Hasil menunjukkan MVBHE mencapai akurasi 96%, lebih tinggi dibandingkan algoritma tunggal, menegaskan efektivitas ensemble dalam meningkatkan klasifikasi status gizi balita.

Jain et al. (2022) meneliti klasifikasi malnutrisi anak <5 tahun di India menggunakan beberapa algoritma machine learning, termasuk Random Forest (RF). Hasil menunjukkan RF mencapai akurasi 68,51%, menegaskan potensinya dalam mendeteksi status malnutrisi anak.

Van et al. (2022) mengembangkan model ML untuk mengklasifikasi risiko malnutrisi anak SD di Filipina berdasarkan data konsumsi makanan dan faktor sosioekonomi. Dari 618 anak, empat algoritma diuji: RF, SVM, LDA, dan Regresi Logistik. Random Forest memberikan akurasi tertinggi 78,55%. Faktor utama meliputi pengeluaran rumah tangga, usia anak dan kepala keluarga, ketahanan pangan, dan keragaman diet. Pendekatan ini mengidentifikasi lebih banyak anak berisiko dibanding metode antropometrik nasional, sehingga lebih efektif untuk intervensi gizi.

Islam et al. (2022) mengembangkan sistem klasifikasi otomatis untuk risiko malnutrisi perempuan menggunakan data BDHS 2014 (15.464 responden). Multinomial Logistic Regression (MLR) digunakan untuk seleksi fitur, dan lima algoritma ML diuji: NB, SVM, DT, ANN, dan RF. Random Forest menunjukkan akurasi tertinggi, 81,4% untuk underweight dan 82,4% untuk overweight/obese, dengan AUC masing-masing 0,837 dan 0,853. Sistem MLR-RF ini efektif untuk

mengidentifikasi perempuan berisiko malnutrisi dan mendukung intervensi kesehatan masyarakat.

Momand et al. (2022) mengembangkan model prediktif untuk mengklasifikasi keparahan malnutrisi edematous pada anak <5 tahun menggunakan data antropometrik (usia, MUAC, berat, tinggi, WHZ). Algoritma yang digunakan adalah Random Forest, J48, dan Naïve Bayes, dengan SMOTE untuk menangani ketidakseimbangan data. Random Forest mencapai akurasi tertinggi 97,14%, diikuti J48 94,51% dan Naïve Bayes 54,17%.

Chilyabanyama et al. (2022) membangun model ML untuk mengklasifikasi stunting pada anak <5 tahun di Zambia menggunakan data ZDHS 2018. Lima algoritma diuji: Logistic Regression, RF, Naïve Bayes, SVM, dan XGBoost, dengan seleksi fitur melalui RF Feature Importance. Random Forest terbaik dengan akurasi 79,2% (train) dan 61,6% (test), F1-score 0,717 (train) dan 0,483 (test), serta Cohen's Kappa 0,554 (train) dan 0,178 (test).

Bitew et al. (2021) mengembangkan model ML untuk mengklasifikasi risiko malnutrisi anak di Ethiopia menggunakan data EDHS 2016. Lima algoritma diuji: xgbTree, RF, NNet, k-NN, dan GLM, untuk indikator stunting, wasting, dan underweight. XgbTree menunjukkan performa terbaik dengan akurasi 67,7% (stunting), 88,0% (wasting), dan 75,7% (underweight), serta sensitivitas tertinggi masing-masing 71,6%, 88,2%, dan 77,5%, menjadikannya model paling efektif untuk deteksi risiko malnutrisi.

Browne et al. (2021) menggunakan Multivariate Random Forest (MRF) untuk mengklasifikasi kemiskinan dan malnutrisi di 11 negara berkembang, meliputi lima indikator: kemiskinan aset, stunting, wasting, berat badan sehat, dan

wanita *underweight*. Akurasi tertinggi $R^2 = 0,58$ untuk kemiskinan, sedangkan malnutrisi $R^2 = 0,17-0,48$, menunjukkan efektivitas terbatas untuk gizi.

Talukder et al. (2020) memklasifikasi status malnutrisi anak <5 tahun di Bangladesh menggunakan LDA, k-NN, SVM, RF, dan LR. Random Forest terbaik dengan akurasi 68,51%, sensitivitas 94,66%, spesifisitas 69,76%, dan Cohen's κ 0,2434, menunjukkan kemampuan diskriminatif moderat.

Rahman et al. (2021) mengklasifikasi malnutrisi balita di Bangladesh menggunakan data BDHS 2014 (7.079 anak). Faktor risiko dianalisis dengan regresi logistik, kemudian SVM, RF, dan LR diterapkan untuk klasifikasi stunting, wasting, dan *underweight*. Random Forest memberikan akurasi tertinggi: 88,3% (stunted), 87,7% (wasted), 85,7% (*underweight*), dengan AUC masing-masing 0,714, 0,523, dan 0,664, mengungguli SVM dan LR.

2.1.3 ASPEK SPASIAL

Santianez et al. (2025) mengembangkan sistem web berbasis SIG untuk memantau malnutrisi di Provinsi Biliran, Filipina. Sistem ini mengintegrasikan indikator gizi (W/A, H/A, W/H) ke peta interaktif dengan filter berdasarkan indikator, waktu, dan wilayah. Hasil menunjukkan sistem efektif membantu petugas kesehatan dan pengambil kebijakan dalam merancang intervensi tepat sasaran serta meningkatkan efisiensi pemantauan dan pelaporan malnutrisi.

Zhang et al. (2024) meneliti klasifikasi stunting anak di Pakistan dengan menggabungkan OLS, SDEM, Random Forest, dan XGBoost, serta Explainable AI (SHAP). Data dari 145 distrik MICS 2017–2018 dianalisis bersama kerentanan iklim. SDEM unggul dengan adjusted R^2 0,71 dan AIC lebih rendah dibanding OLS. RF dan XGBoost menyoroti spatial lag kerentanan iklim dan kemiskinan

multidimensi sebagai fitur utama, serta mengungkap hubungan non-linier antara kemiskinan dan stunting, menegaskan efektivitas kombinasi spasial dan AI untuk analisis malnutrisi berbasis lokasi.

Prayudhy dan Ahyanti (2022) memanfaatkan Sistem Informasi Geografis (SIG) untuk memetakan daerah rawan stunting melalui metode Weighted Sum Overlay, sehingga mampu menampilkan distribusi risiko spasial pada tingkat kecamatan di Kabupaten Pesawaran.

Pendekatan spasial juga digunakan oleh BMC Nutrition (2022) dengan menerapkan model geostatistik Bayesian untuk menghasilkan peta resolusi tinggi (1x1 km) prevalensi malnutrisi akut berat di Papua, yang memungkinkan identifikasi *hotspot* kasus bahkan pada wilayah yang tidak tercakup survei.

Sejalan dengan itu, Utomo et al. (2022) menegaskan bahwa GIS berfungsi sebagai sarana penting dalam menghasilkan peta kerentanan gizi dan mendukung penentuan wilayah prioritas intervensi.

Kontribusi serupa juga terlihat pada penelitian Sari et al. (2023) yang menggunakan GIS untuk memvisualisasikan distribusi spasial stunting di Kabupaten Pasaman, menyoroti kecamatan dengan prevalensi tinggi sebagai fokus penanganan.

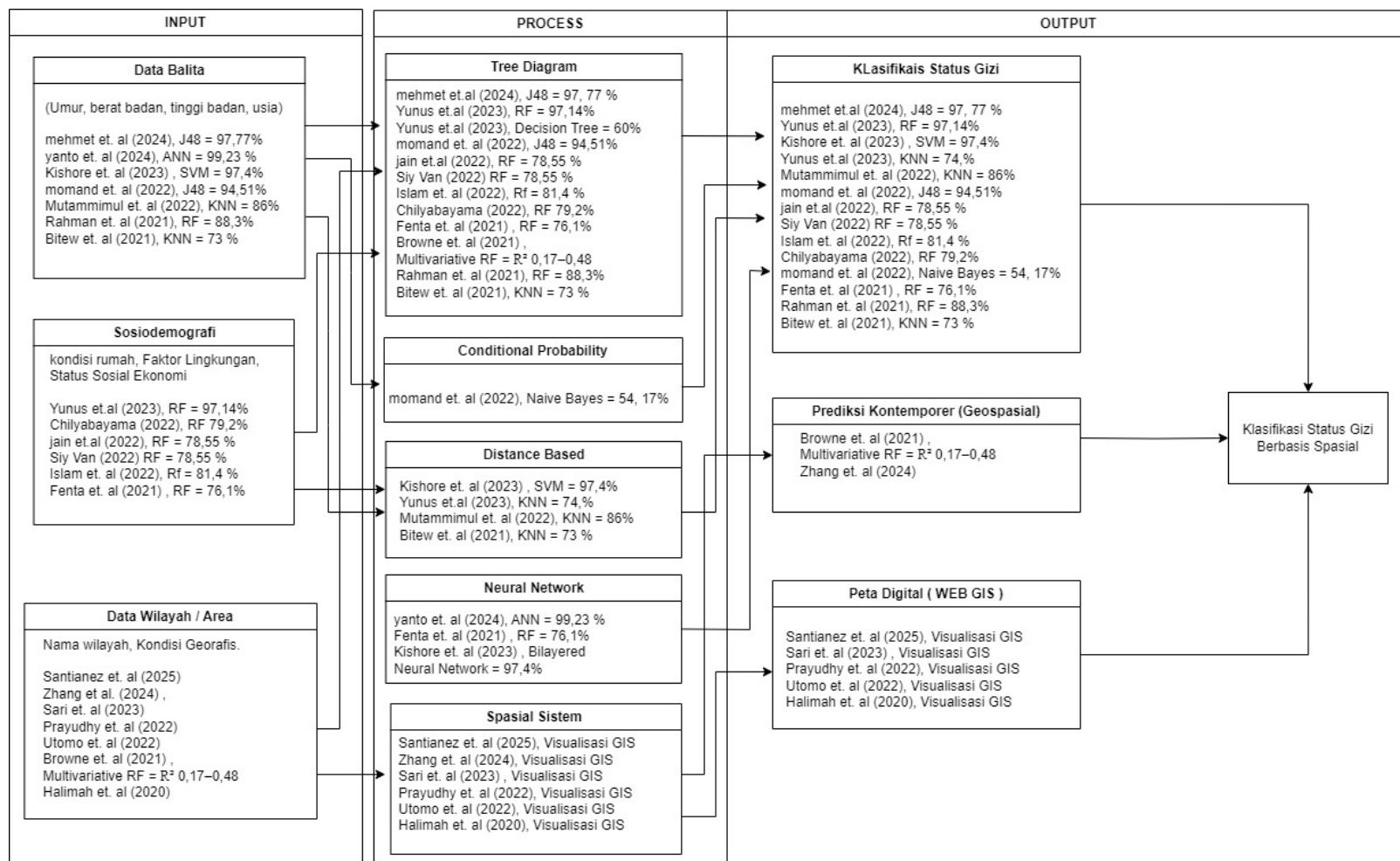
Halimah dan Suntin (2020) juga menggunakan SIG untuk menggambarkan distribusi spasial balita stunting di Makassar, dengan pemetaan kasus per kecamatan yang memperlihatkan area dengan konsentrasi tertinggi.

Sementara itu, Shamsuddin et al. (2022) dalam tinjauannya menekankan peran analisis spasial seperti hotspot detection dan pemodelan geografis sebagai alat untuk memahami pola distribusi malnutrisi anak di berbagai konteks.

WHO menetapkan kategori prevalensi malnutrisi pada balita—very low hingga very high—berdasarkan persentase stunting, wasting, atau underweight, misalnya stunting <10% rendah, 10–20% sedang, dan >20% tinggi hingga sangat tinggi (De Onis et al., 2019).

2.1.4 Kerangka Teoritis / *Theoretical Framework*

Dalam subbagian ini disusun kerangka teori yang merujuk pada berbagai studi pustaka yang relevan sebagai acuan konseptual penelitian. Penyusunan kerangka teori dilakukan melalui tahapan pengelompokan parameter atau variabel yang digunakan sebagai input penelitian, dilanjutkan dengan uraian mendalam mengenai metode analisis yang diterapkan, serta deskripsi terperinci mengenai output atau hasil yang diharapkan. Seluruh keterhubungan antar variabel tersebut divisualisasikan pada Gambar 2.2 sebagai berikut :



Gambar 2. 2 Kerangka Teoritis

Metode ANN dan RF terbukti unggul dalam klasifikasi status gizi. ANN efektif mengenali pola non-linear dengan akurasi tinggi (Yanto et al., 2024) sebesar 99,23%, sementara RF stabil dan akurat untuk data skala besar berbasis lokasi (Momand et al., 2022) sebesar 97,14%. Kombinasi ANN dan RF dipilih untuk menghasilkan klasifikasi gizi yang akurat, komprehensif, dan berbasis spasial.

BAB III

METODOLOGI PENELITIAN

3.1. Desain Penelitian

Penelitian ini dilakukan secara sistematis mulai dari studi literatur, pengumpulan dan validasi data, perancangan dan implementasi sistem, eksperimen metode, evaluasi hasil, hingga penarikan kesimpulan. Prosedur penelitian ditunjukkan pada Gambar 3.1.



Gambar 3. 1 Desain Penelitian

3.1.1. Studi Literatur

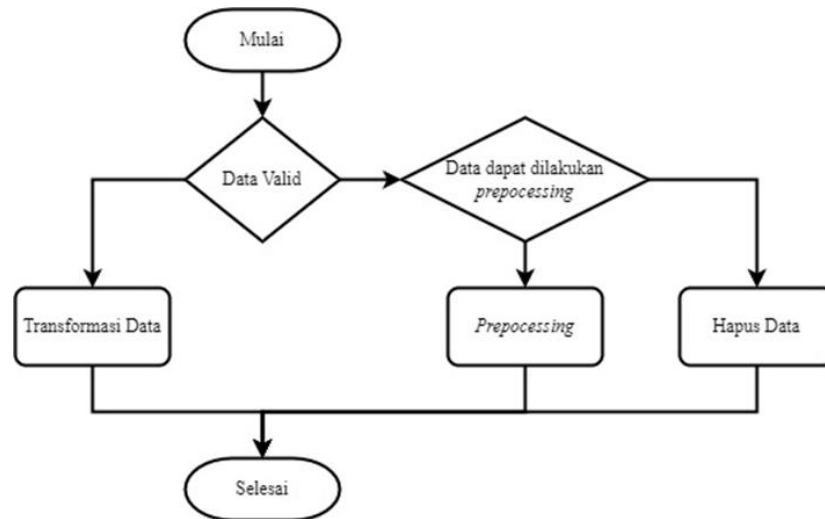
Tahap studi literatur meliputi pengumpulan referensi terkait status gizi, sistem spasial, serta metode ANN dan RF, bertujuan memahami pendekatan yang ada, menemukan research gap, dan menentukan metode terbaik untuk klasifikasi gizi berbasis lokasi.

3.1.2. Pengumpulan Data

Data status gizi berupa umur, tinggi, dan berat badan diperoleh dari Dinas Kesehatan Kabupaten Sumenep (2023–2024). Hasil klasifikasi digunakan untuk membangun sistem spasial berupa peta digital interaktif, memudahkan identifikasi daerah rawan gizi kurang dan evaluasi akurasi model.

3.1.3. Validasi Data

Data gizi divalidasi untuk memastikan keaslian dan kesesuaiannya sebelum digunakan pada tahap pelatihan. Proses validasi ditunjukkan pada Gambar 3.2.



Gambar 3. 2 Validasi Data

Proses validasi memeriksa setiap atribut data gizi. Data valid dilanjutkan ke transformasi, sedangkan data tidak lengkap diproses melalui preprocessing untuk perbaikan. Jika tidak dapat diperbaiki, data dibuang, dan data yang sudah tervalidasi disimpan secara terpisah.

Selanjutnya, dilakukan eksplorasi awal data gizi dari Dinas Kesehatan (2023–2024) untuk memahami struktur dan karakteristiknya.

Dari hasil eksplorasi, data gizi diproses melalui beberapa langkah: menentukan status gizi berdasarkan kondisi tubuh, mengisi nilai kosong dengan imputasi sesuai kolom, dan mengonversi data kategori agar siap diproses. Proses ini ditunjukkan pada Tabel 3.1.

Tabel 3. 1 Data Gizi

ID	NAMA	JK	UMUR (Bln)	BB	TB	STATUS
1	M Alfarisqi	1	13,5	94	45	Gizi Kurang
2	Jesifa Khairunnisa	2	15,4	96	43	Gizi Kurang

3	Davina Nur Latifa	2	12,2	92	42	Gizi Baik
4	Azizah Salsabila	2	15,4	104	56	Gizi Kurang
5	Aisyah Fissilmi	2	13,8	95	49	Gizi Kurang
6	Tanzila Nur F	2	13,8	95	49	Gizi Kurang
7	M Sakbani R	1	14,6	96	46	Gizi Kurang
8	M Ozil Jaris	1	12,5	95	43	Gizi Kurang
9	Raflian Arka Ayubi	1	13	94,2	43	Gizi Kurang

Untuk mencegah overfitting, 1.260 data gizi dibagi rasio 800:460, dengan sebagian besar untuk pelatihan dan sisanya untuk pengujian model, sebagaimana ditunjukkan pada Tabel 3.2.

Tabel 3. 2 Data Awal

Status	Data Latih	Data Uji
Gizi Kurang	416	178
Gizi Baik	384	282

Langkah ini meliputi analisis struktur data dan perencanaan prosedur transformasi serta validasi. Rincian parameter validasi disajikan pada Tabel 3.3.

Tabel 3. 3 Parameter Validasi Data

No	Parameter	Kriteria Data Valid
1	Keberadaan Data	Data harus terisi dan tidak boleh kosong, serta memiliki nilai yang sesuai dengan ketentuan.
2	Tipe Data	Data harus dalam format numerik, bukan <i>string</i> atau format <i>non-numerik</i> lainnya, agar dapat dianggap valid.
3	Jenis Data	Kode kategori digunakan untuk mengonversi data kategori menjadi format yang dapat diolah.

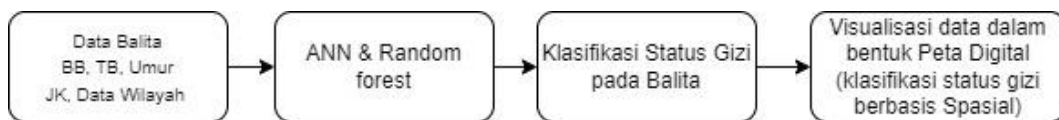
Masalah umum pada data gizi meliputi data hilang, data pribadi, format non-numerik, dan data pilihan. Validasi dilakukan untuk mengatasinya, sebagaimana dijelaskan pada Tabel 3.4.

Tabel 3. 4 Prosedur Penyelesaian Permasalahan Data

No	Permasalahan	Tindakan
1	Data Kosong	Data kosong akan diisi dengan nilai 0
2	Data Pribadi	Informasi yang berkaitan dengan identitas pribadi akan dihapus
3	Data <i>Noo-Numerik</i>	Untuk memungkinkan pemrosesan lebih lanjut, data akan diubah ke dalam bentuk numerik
4	Data dapat dipilih dari berbagai opsi	Data juga akan ditransformasi ke dalam format kategori

3.1.4. Sistem Desain dan Implementasi

Alur proses penelitian mencakup input data gizi, klasifikasi menggunakan ANN dan Random Forest, serta visualisasi hasil dalam peta digital. Desain sistem ditunjukkan pada Gambar 3.3.



Gambar 3. 3 Desain Sistem

A. Data Gizi

Langkah awal penelitian adalah menyiapkan dan memverifikasi data gizi, termasuk pengecekan duplikasi, outlier, dan kelengkapan atribut. Data yang tidak valid dibersihkan atau dikoreksi agar tidak memengaruhi performa model.

B. *Artificial Neural Network* (ANN)

Selanjutnya, data diklasifikasikan untuk menentukan status gizi menggunakan algoritma ANN. Langkah-langkah dan formula pembentukan pohon keputusan dijelaskan sebagai berikut:

1. Normalisasi *Min-Max* dapat dilihat pada persamaan 3.1.

$$X' = \frac{X - X_{\min}}{X_{\max} - X_{\min}} \quad (3.1)$$

Dimana:

X = Nilai Asli

$X_{\min}$ = Nilai Minimum dalam Dataset

$X_{\max}$ = Nilai Maksimum dalam Dataset

X' = Nilai Hasil Normalisasi

Persamaan 3.1 menunjukkan normalisasi Min-Max, yang mengubah skala data agar berada antara 0 dan 1: nilai asli dikurangi $X_{\min}$ lalu dibagi selisih $X_{\max} - X_{\min}$. Metode ini memastikan semua fitur berkontribusi seimbang dalam algoritma pembelajaran mesin.

2. Perhitungan Bobot dan Bias *perceptron*, pada persamaan 3.2.

$$Y = f(\sum(W_i X_i) + b) \quad (3.2)$$

Dimana:

Y = output klasifikasi

W_i = bobot

X_i = input fitur

b = bias

f = fungsi aktivasi

Persamaan 3.2 menjelaskan neuron perceptron: setiap input X_i dikalikan bobot W_i , dijumlahkan dengan bias b , lalu diproses melalui fungsi aktivasi f untuk menghasilkan output klasifikasi Y .

3. Fungsi Aktivasi Sigmoid dapat dilihat pada persamaan 3.3.

$$f(x) = \frac{1}{1 + e^{-x}} \quad (3.3)$$

Dimana:

$f(x)$ = Output fungsi sigmoid

$x = \text{Nilai input}$

$e = \text{Bilangan Euler}$

$-x = \text{Negasi dari input}$

Persamaan 3.3 menunjukkan fungsi sigmoid, yang mengubah input x menjadi output antara 0 dan 1. Fungsi ini cocok untuk memodelkan probabilitas atau keputusan biner, dengan kurva halus berbentuk "S" yang menahan nilai ekstrem agar tetap stabil.

4. *Backpropagation* (Pelatihan model) dapat dilihat pada persamaan 3.4.

$$W_{baru} = W_{lama} + \alpha \times (Y_{aktual} - Y_{prediksi}) \times X \quad (3.4)$$

Dimana:

$\alpha = \text{learning rate}$

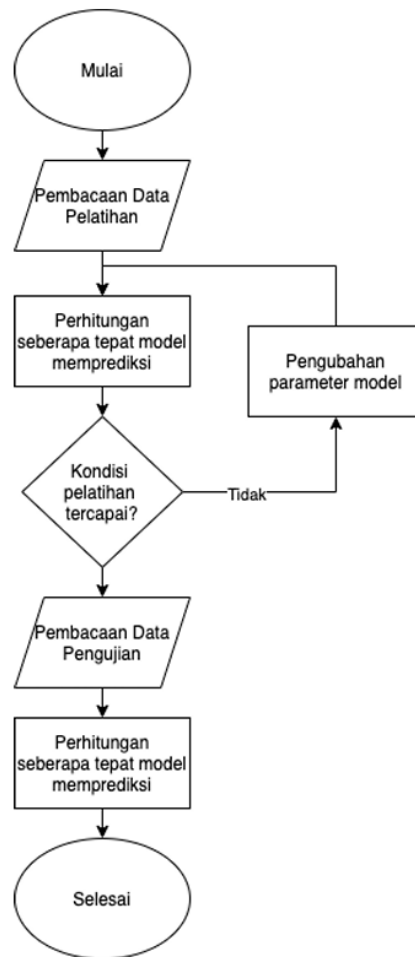
$Y_{aktual} = \text{Nilai Sebenarnya}$

$Y_{klasifikasi} = \text{hasil klasifikasi}$

$X = \text{input fitur}$

Persamaan 3.4 digunakan untuk memperbarui bobot pada model Perceptron. Bobot baru diperoleh dengan menambahkan hasil perkalian antara selisih nilai aktual dan prediksi, learning rate, serta nilai input ke bobot lama, bertujuan mengurangi kesalahan klasifikasi secara bertahap dan meningkatkan akurasi model.

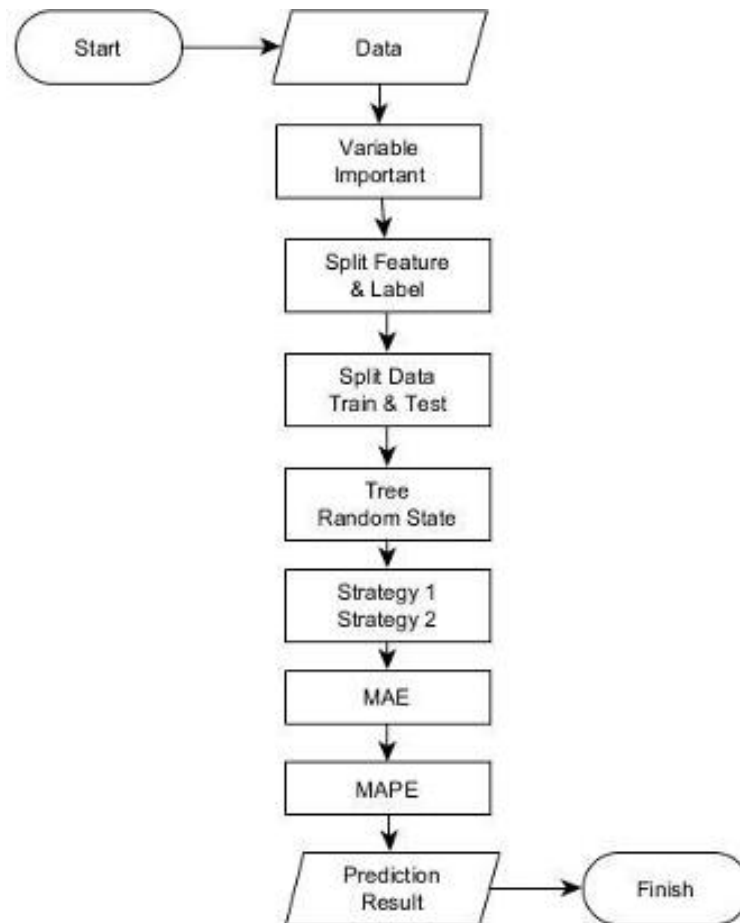
Model ANN akan diimplementasikan menggunakan bahasa Python dengan memanfaatkan library yang mendukung jaringan saraf tiruan. Proses pelatihan mencakup perhitungan output, evaluasi kesalahan, dan pembaruan bobot, sebagaimana ditunjukkan pada Gambar 3.4.



Gambar 3. 4 Diagram alir algoritma metode ANN

C. *Random Forest*

Pada algoritma Random Forest, tahap awal adalah memilih variabel penting yang memengaruhi klasifikasi. Data kemudian diproses dan dioptimalkan untuk menghasilkan klasifikasi yang akurat. Alur kerja metode Random Forest ditunjukkan pada Flowchart Gambar 3.5.



Gambar 3. 5 Flowchart Klasifikasi RF

(Sumber: diolah dari Breiman, L. 2001)

Pada Gambar 3.5, proses pemilihan Variable Importance (VI) digunakan untuk menentukan variabel yang paling berpengaruh berdasarkan perhitungan Random Forest. Menurut Breiman (2001), Random Forest terdiri dari banyak pohon (tree) yang membentuk “forest” dan merupakan pengembangan dari metode Classification and Regression Tree (CART).

Langkah – langkah serta formula yang digunakan dalam pembentukan pohon Keputusan dijelaskan sebagai berikut:

1. Perhitungan entropi dilakukan dengan menggunakan rumus yang ditunjukkan pada Persamaan 3.5.

$$\mathbf{Entropy}(S) = - \sum_{i=1}^c p_i \log_2(p_i) \quad (3.5)$$

Dimana:

Entropy(S) = nilai entropy dari dataset S

c = jumlah total kelas (kategori) dataset

i = indeks yang mewakili kelas ke-i

pi = Probabilitas keunculan elemen dari kelas ke – i

Log₂(Pi) = logaritma basis 2 probabilitas pi

Persamaan 3.5 digunakan dalam algoritma decision tree (ID3 atau C4.5) untuk mengukur entropi, yaitu tingkat ketidakpastian atau impuritas suatu dataset. Semakin merata distribusi kelas dalam data, semakin tinggi nilai entropinya, menunjukkan tingkat ketidakpastian klasifikasi yang lebih besar. *Information Gain* (IG) untuk Seleksi Atribut dalam Pohon Keputusan dapat dilihat pada persamaan 3.6.

$$\mathbf{IG}(S, A) = \mathbf{Entropy}(S) - \sum_{v \in \mathbf{Values}(A)} \frac{|S_v|}{|S|} \times \mathbf{Entropy}(S_v) \quad (3.6)$$

Dimana :

Entropy(S) = nilai entropy dari dataset S

V ∈ Values(A) = setiap nilai v yang mungkin dari atribut A

$\frac{|S_v|}{|S|}$ = Proporsi data pada subset S

Entropy(S_v) = Entropy dari subset data S

Persamaan 3.6 menunjukkan *Information Gain* (IG), yang digunakan dalam algoritma pohon keputusan untuk menilai seberapa efektif suatu atribut memisahkan data. IG dihitung sebagai selisih antara entropi sebelum pembagian dan rata-rata entropi setelah

pembagian; atribut dengan IG lebih tinggi diprioritaskan sebagai node pada pohon keputusan.

D. Visualisasi data klasifikasi ke bentuk Spasial

Menurut WHO, prevalensi malnutrisi balita digunakan untuk mengukur tingkat masalah kesehatan masyarakat dan dikategorikan ke dalam tiga zona warna: hijau (<10%) menunjukkan aman, kuning (10–20%) menandakan waspada, dan merah (>20%) menandakan rawan. Klasifikasi ini memudahkan identifikasi daerah prioritas intervensi dan perencanaan strategi penanggulangan malnutrisi. Tabel 3.5 menyajikan rincian zona WHO tersebut. Tabel 3. 5 Klasifikasi Zona

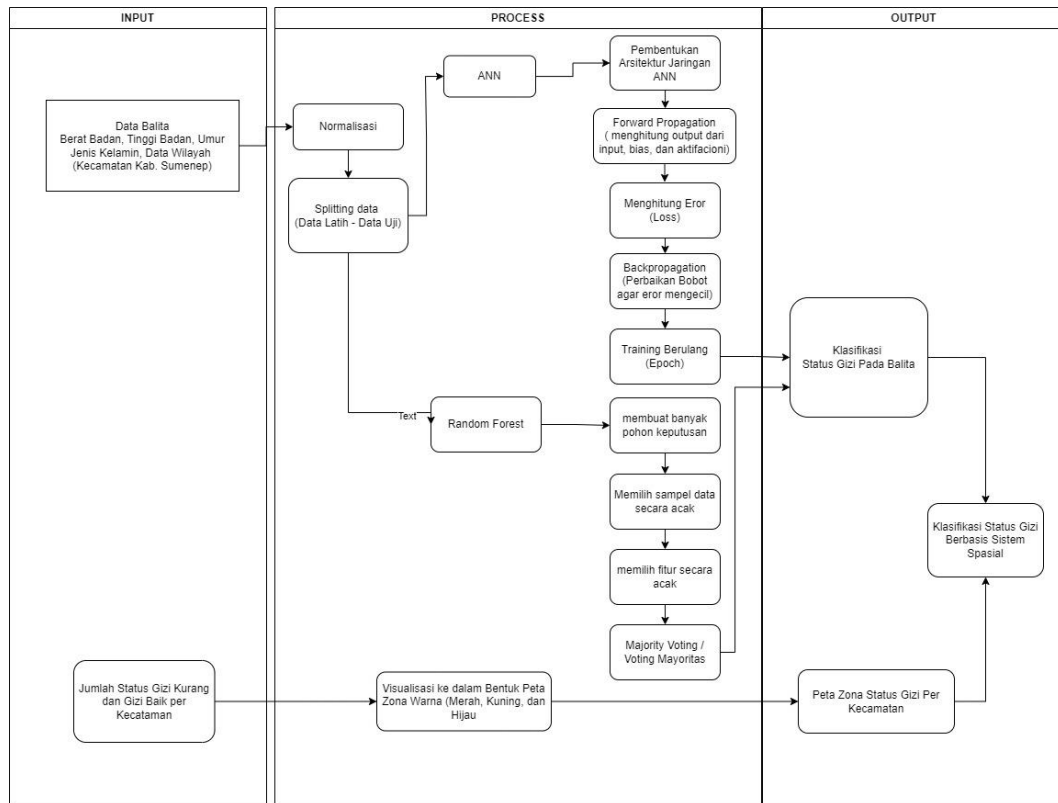
Zona Warna	Ambang Batas	Kategori	Interpretasi
Hijau	<10%	Rendah	Aman
Kuning	10 - 20 %	Sedang	Waspada
Merah	>20%	Tinggi	Rawan

3.1.5. Eksperimen

Tahap uji coba penelitian menggunakan PyCharm dengan bahasa pemrograman Python untuk mengklasifikasi status gizi menggunakan metode ANN dan Random Forest. Hasil klasifikasi kemudian dievaluasi menggunakan Confusion Matrix, dan kesimpulan ditarik berdasarkan analisis hasil tersebut.

3.1.6. Kerangka Konsep Penelitian

Berdasarkan kajian penelitian terdahulu pada studi pustaka, konsep penelitian ini dirumuskan dalam bentuk kerangka konsep yang digambarkan pada Gambar 3.4 berikut:



Gambar 3. 4 Kerangka Konsep Penelitian

Pada tahap pertama, Artificial Neural Network (ANN) memulai proses dengan menormalisasi data agar semua nilai berada pada skala seragam tertentu. Selanjutnya, dilakukan pemisahan data menjadi data latih dan data uji, kemudian Langkah selanjutnya adalah membentuk jaringan ANN, seperti mencari parameter terbaik, dari data input tadi akan d proses *forward propagation* (yaitu menghitung output dari input, bobot, bias, dan activation Function sehingga mendapatkan sebuah kelas prediksi), hasil tersebut akan dihitung error (antara hasil kelas prediksi dengan data sebenarnya, Langkah selanjutnya adalah *backpropagation* (memperbaiki bobot agar error mengecil). Selanjutnya akan dilakukan training secara berulang dengan epoch, untuk mendapatkan hasil yang lebih stabil dan lebih akurat. Hasil akhir dari proses ini adalah klasifikasi status gizi pada balita.

Proses Random Forest dimulai dengan menerima data input yang sudah melewati proses normalisasi dan pembagian data latih-uji, kemudian membuat banyak pohon Keputusan (*Decision Tree*) dengan cara menghasilkan beberapa subset data secara acak melalui teknik *bootstrap*. Pada setiap pohon, fitur-fitur yang digunakan untuk membentuk percabangan juga dipilih secara acak sehingga setiap pohon memiliki struktur yang berbeda. Masing-masing pohon kemudian melakukan proses klasifikasi secara mandiri dan menghasilkan prediksinya. Seluruh hasil prediksi dari masing-masing pohon digabungkan melalui mekanisme voting mayoritas, dan kelas yang mendapatkan suara terbanyak menjadi hasil akhir klasifikasi status gizi.

Proses terakhir adalah memanfaatkan data hasil klasifikasi dituangkan kedalam peta digital. Dengan menghasilkan peta zona status gizi per kecamatan.

3.2. Instrumen Penelitian

Tabel 3.6 menyajikan daftar parameter penelitian, meliputi variabel independen (bebas), dependen (terikat), dan intervening (penghubung). Fokus penelitian ini adalah variabel dependen, yaitu variabel yang dipengaruhi oleh variabel independen atau bebas.

Tabel 3. 6 Instrumen Penelitian

<i>Variabel Independent</i>	<i>Variable Intervening</i>	<i>Main Process</i>	<i>Variable Dependent</i>
Berat Badan (BB), Tinggi Badan (TB), Jenis Kelmain, Umur.	Klasifikasi status gizi	<i>ANN, Random Forest</i>	<i>Accuracy, error, recall, F1 Score</i>

3.3. Miniatur Penelitian (Simulasi)

Pada tahap ini hanya menggunakan data yang terbatas sebagai sampel, dan aplikasi excel untuk simulasi model. Dapat dilihat pada table 3.7.

Tabel 3. 7 Data Status Gizi

ID	NAMA	JENIS KELAMIN	UMUR	BB	TB	STATUS
1	M ZAFRAN AL G	1	57	15,6	106,5	Gizi Kurang
1	RAFAEL JAKFAR	1	39	13,3	92	Gizi Baik
1	DINAR NADIYA W	2	51	15	100	Gizi Kurang
1	ADITYA RISQI ANANTA	1	58	15,5	108	Gizi Baik
1	ZAURA MIQAILA A	2	50	12,2	96	Gizi Kurang
1	AULAN BASIRA F	1	54	15,7	105	Gizi Kurang
1	SILMI AFIFAH	1	52	14,6	101	Gizi Kurang
1	M NUR HAFIZAH	1	59	18,6	106	Gizi Kurang
1	TATU DIIGAHAYU	2	54	15	104	Gizi Kurang
1	M ALFARISQI	1	45	13,5	94	Gizi Kurang
1	JESIFA KHAIRUNNISA	2	43	15,4	96	Gizi Kurang
1	DAVINA NUR LATIFA	2	42	12,2	92	Gizi Baik
1	AZIZAH SALSABILA	2	56	15,4	104	Gizi Kurang
1	AISYAH FISSILMI	2	49	13,8	95	Gizi Kurang
1	TANZILA NUR F	2	49	13,8	95	Gizi Kurang
1	M SAKBANI R	1	46	14,6	96	Gizi Kurang
1	M OZIL JARIS	1	43	12,5	95	Gizi Kurang
1	RAFLIAN ARKA AYUBI	1	43	13	94,2	Gizi Kurang

Gambar di atas menunjukkan data mentah yang digunakan dalam simulasi, terdiri dari variabel prediktor X1 (jenis kelamin), X2 (umur), X3 (BB/TB), serta variabel target Klasifikasi berupa Status Gizi.

3.4.1 Pengukuran Status Gizi menggunakan ANN

Simulasi ini menggunakan sebagian data untuk memfokuskan pada langkah-langkah metode ANN dalam mengklasifikasi status gizi, yang dilakukan melalui aplikasi Excel, sebagaimana ditampilkan pada Tabel 3.8.

Tabel 3. 8 Normalisasi dan Percobaan Iterasi ke-1

DATA MENTAH					NORMALISASI			Fungsi Aktivasi (Step Function)	prediksi aktual
ID	NAMA	BB	TB	UMUR	Hitung Output Neuron (z)	X1	X2	X3	
1	M ZAFRA	15,6	106,5	57	0,764375	0,53125	0,9063	0,9	1
1	RAFAEL	13,3	92	39	0,16875	0,171875	0	0	0
1	DINAR N	15	100	51	0,545	0,4375	0,5	0,6	1
1	ADITYA R	15,5	108	58	0,79625	0,515625	1	0,95	1
1	ZAURA N	12,2	96	50	0,285	0	0,25	0,55	0
1	AULAN B	15,7	105	54	0,7125	0,546875	0,8125	0,75	1
1	SILMI AF	14,6	101	52	0,54875	0,375	0,5625	0,65	1
1	M NUR H	18,6	106	59	0,9625	1	0,875	1	1
1	TATU DII	15	104	54	0,65	0,4375	0,75	0,75	1
1	M ALFAR	13,5	94	45	0,27875	0,203125	0,125	0,3	0
1	JESIFA K	15,4	96	43	0,415	0,5	0,25	0,2	0
1	DAVINA	12,2	92	42	0,13	0	0	0,15	0
1	AZIZAH S	15,4	104	56	0,695	0,5	0,75	0,85	1
1	AISYAH R	13,8	95	49	0,35625	0,25	0,1875	0,5	0
1	TANZILA	13,8	95	49	0,35625	0,25	0,1875	0,5	0
1	M SAKBA	14,6	96	46	0,395	0,375	0,25	0,35	0
1	M OZIL J	12,5	95	43	0,215	0,046875	0,1875	0,2	0
1	RAFLIAN	13	94,2	43	0,23125	0,125	0,1375	0,2	0
				inisialisasi	W1	0,4			
				Bobot dan	W2	0,3			
				Bias	W3	0,2			
					B	0,1			

Tabel di atas menampilkan perhitungan iterasi pertama pada algoritma ANN untuk klasifikasi status gizi, meliputi normalisasi data, perhitungan output neuron dengan bobot awal, dan penerapan fungsi aktivasi biner (step function) untuk menentukan klasifikasi. Berikut analisis Iterasi-1:

1. *Input & Normalisasi:*

Variabel input yang digunakan adalah berat badan (BB), tinggi badan (TB), dan usia (UMUR), yang dinormalisasi menjadi X1, X2, dan X3. Ini penting agar semua fitur berada dalam skala yang sama untuk pemrosesan di jaringan saraf.

2. *Parameter Awal:*

Bobot awal (W1, W2, W3) dan bias (b) diinisialisasi masing-masing:

$$W1 = 0.3, W2 = 0.5, W3 = 0.1, b = 0.1$$

3. *Output Neuron & Klasifikasi:*

Output neuron dihitung berdasarkan rumus:

$$Output = (X1 * W1) + (X2 * W2) + (X3 * W3) + b \quad (3.9)$$

4. Fungsi aktivasi yang digunakan adalah *step function* (fungsi ambang), yang mengubah *output* menjadi 0 atau 1.
5. Hasil klasifikasi awal semuanya bernilai 0, menunjukkan bahwa pada iterasi ini model belum mampu membedakan antara status gizi dengan baik.
6. Evaluasi Sementara:

Kemungkinan model mengalami *underfitting* di iterasi awal ini karena bobot belum cukup terlatih. Untuk itu maka perlu dilakukan iterasi berikutnya, dapat dilihat pada tabel 3.9.

Tabel 3. 9 *Looping* Iterasi ke-2

looping iterasi	X1	X2	X3	PREDIKSI	Fungsi Aktivasi (Step Function)	PREDIKSI AKTUAL		
ke -2	0,5313	0,90625	0,9	0,5730234	1	0		
	0,1719	0	0	0,5596191	1	1		
	0,4375	0,5	0,6	1,3224453	1	0		
	0,5156	1	0,95	0,7427324	1	1		
	0	0,25	0,55	0,9253438	1	0		
	0,5469	0,8125	0,75	1,0048145	1	0	BOBOT BARU	
	0,375	0,5625	0,65	1,1943516	1	0	W1	0,3
	1	0,875	1	1,3900781	1	0	W2	0,2
	0,4375	0,75	0,75	0,5162891	1	0	W3	0,1
	0,2031	0,125	0,3	0,3796309	0	0	B	0,1
	0,5	0,25	0,2	0,2477813	0	0		
	0	0	0,15	0,7665	1	1		
	0,5	0,75	0,85	0,6114688	1	0		
	0,25	0,1875	0,5	0,3684766	0	0		
	0,25	0,1875	0,5	0,4309766	0	0		
	0,375	0,25	0,35	0,4084219	0	0		
	0,0469	0,1875	0,2	0,2150176	0	0		
	0,125	0,1375	0,2	0,0941484	0	0		

Tabel di atas menunjukkan proses iterasi kedua pada pelatihan ANN untuk klasifikasi status gizi, di mana bobot ($W1=0,3$; $W2=0,2$; $W3=0,1$) dan bias ($b=0,1$) diperbarui berdasarkan error iterasi sebelumnya. Output neuron dihitung ulang dan diproses menggunakan *step function*, menghasilkan variasi klasifikasi pertama dibanding iterasi awal yang seluruhnya 0. Meskipun belum sepenuhnya akurat, perubahan ini menunjukkan model mulai belajar mengenali pola data dan peningkatan performa klasifikasi. Iterasi ini menjadi tahap penting dalam penyesuaian awal agar ANN mampu menghasilkan klasifikasi status gizi yang lebih mendekati kondisi aktual.

Tabel 3. 10 *Looping* Iterasi ke-3

looping iterasi	X1	X2	X3	PREDIKSI	Fungsi Aktivasi (Step Function)	PREDIKSI AKTUAL		
ke -3	0,5313	0,90625	0,9	0,4250774	0	0		
	0,1719	0	0	0,0543869	0	1		
	0,4375	0,5	0,6	0,2522186	0	0		
	0,5156	1	0,95	0,3761121	0	1		
	0	0,25	0,55	0,0714965	0	0	W1	0,3
	0,5469	0,8125	0,75	0,3447941	0	0	W2	0,2
	0,375	0,5625	0,65	0,2452032	0	0	W3	0,1
	1	0,875	1	0,5126252	1	0	B	0,1
	0,4375	0,75	0,75	0,300344	0	0		
	0,2031	0,125	0,3	0,1014844	0	0		
	0,5	0,25	0,2	0,2092633	0	0		
	0	0	0,15	0,0087642	0	1		
	0,5	0,75	0,85	0,3259638	0	0		
	0,25	0,1875	0,5	0,1378431	0	0		
	0,25	0,1875	0,5	0,1378431	0	0		
	0,375	0,25	0,35	0,1784734	0	0		
	0,0469	0,1875	0,2	0,0560393	0	0		
	0,125	0,1375	0,2	0,3667393	0	0		

Pada iterasi ke tiga dengan bobot yang sudah diperbaharui, yaitu $W1=0.3$, $W2=0.2$, $W3=0.1$, dan bias 0.1 . menghasilkan klasifikasi yang berbeda dengan iterasi sebelumnya, hal ini menunjukkan bahwa model masih harus belajar lagi agar bisa mendapatkan hasil yang lebih optimal.

Tabel 3. 11 *Looping* Iterasi ke -4

looping iterasi	X1	X2	X3	PREDIKSI	Fungsi Aktivasi (Step Function)	PREDIKSI AKTUAL		
ke -4	0,5313	0,90625	0,9	0,3255843	0	0		
	0,1719	0	0	0,0505056	0	1		
	0,4375	0,5	0,6	0,1880208	0	0		
	0,5156	1	0,95	0,270439	0	1		
	0	0,25	0,55	0,0297305	0	0	W1	0,3
	0,5469	0,8125	0,75	0,2573239	0	0	W2	0,1
	0,375	0,5625	0,65	0,1770878	0	0	W3	0
	1	0,875	1	0,3979077	0	0	B	0,1
	0,4375	0,75	0,75	0,2177514	0	0		
	0,2031	0,125	0,3	0,0745537	0	0		
	0,5	0,25	0,2	0,176656	0	0		
	0	0	0,15	0	0	1		
	0,5	0,75	0,85	0,236117	0	0		
	0,25	0,1875	0,5	0,0957606	0	0		
	0,25	0,1875	0,5	0,0957606	0	0		
	0,375	0,25	0,35	0,1399246	0	0		
	0,0469	0,1875	0,2	0,0360722	0	0		
	0,125	0,1375	0,2	0,3296374	0	0		

Pada iterasi ke-empat, model diberikan bobot yang berbeda yaitu $W1=0.3$, $W2=0.1$, dan $W3=0.0$, dan Bias $=0.1$. hasil klasifikasi pada iterasi ke-empat ini juga berbeda dengan iterasi sebelumnya. Yang dimana dapat disimpulkan bahwa dengan bobot yang berbeda-beda setiap iterasi akan memberikan pembelajaran baru terhadap model yang dibuat. Karena hasil masih beragam, maka perlu dilakukan untuk iterasi selanjutnya.

Tabel 3. 12 *Looping* Iterasi ke-5

looping iterasi	X1	X2	X3	PREDIKSI	Fungsi Aktivasi (Step Function)	PREDIKSI AKTUAL		
ke -5	0,5313	0,90625	0,9	0,2632832	0	0		
	0,1719	0	0	0,0475328	0	1		
	0,4375	0,5	0,6	0,1481189	0	0		
	0,5156	1	0,95	0,2041769	0	1		
	0	0,25	0,55	0,0062376	0	0	W1	0,3
	0,5469	0,8125	0,75	0,2019142	0	0	W2	0,1
	0,375	0,5625	0,65	0,1349577	0	0	W3	0,0
	1	0,875	1	0,3254907	0	0	B	0,1
	0,4375	0,75	0,75	0,1660776	0	0		
	0,2031	0,125	0,3	0,0585613	0	0		
	0,5	0,25	0,2	0,1547706	0	0		
	0	0	0,15	-0,004395	0	1		
	0,5	0,75	0,85	0,1804319	0	0		
	0,25	0,1875	0,5	0,0712528	0	0		
	0,25	0,1875	0,5	0,0712528	0	0		
	0,375	0,25	0,35	0,1158059	0	0		
	0,0469	0,1875	0,2	0,0238685	0	0		
	0,125	0,1375	0,2	0,0410035	0	0		

Pada iterasi kelima, dengan bobot $W1=0,3$; $W2=0,1$; $W3=0$; dan bias 0,1, nilai klasifikasi hampir sama dengan iterasi sebelumnya, sehingga proses iterasi dihentikan. Hal ini menandakan bahwa model telah cukup belajar (converged). Kesimpulan sementara menunjukkan bahwa bobot sangat berpengaruh terhadap hasil klasifikasi, dan pada simulasi ini diperlukan lima iterasi untuk mencapai akurasi yang lebih optimal.

3.4.2 Pengukuran Status Gizi dengan *Random Forest*

Random Forest sesungguhnya terdiri dari ratusan hingga ribuan pohon keputusan, di mana setiap pohon dilatih menggunakan subset data acak dengan pengembalian (bootstrapping). Selain itu, setiap node pohon hanya mempertimbangkan subset fitur acak untuk menentukan pemisahan terbaik, sehingga meningkatkan diversitas pohon dan akurasi keseluruhan model. Rincian proses ini ditunjukkan pada Tabel 3.13.

Tabel 3. 13 Pohon ke-1 berdasarkan IMT

ID	NAMA	BB (kg)	TB (cm)	TB (m)	Hitung IMT	aturan keputusan sederhana berdasarkan IMT
1	M ZAFRAN	15,6	106,5	1,07	13,75388	Gizi Kurang
1	RAFAEL JAM	13,3	92	0,92	15,71361	Gizi Baik
1	DINAR NADIA	15	100	1,00	15,00000	Gizi Baik
1	ADITYA RIS	15,5	108	1,08	13,28875	Gizi Kurang
1	ZAURA MIO	12,2	96	0,96	13,23785	Gizi Kurang
1	AULAN BAK	15,7	105	1,05	14,24036	Gizi Baik
2	SILMI AFIF	14,6	101	1,01	14,31232	Gizi Baik
2	M NUR HA	18,6	106	1,06	16,55393	Gizi Baik
2	TATU DIIGA	15	104	1,04	13,86834	Gizi Kurang
2	M ALFARIS	13,5	94	0,94	15,27841	Gizi Baik
2	JESIFA KHA	15,4	96	0,96	16,71007	Gizi Baik
3	DAVINA NU	12,2	92	0,92	14,41399	Gizi Baik
3	AZIZAH SA	15,4	104	1,04	14,23817	Gizi Baik
3	AISYAH FIS	13,8	95	0,95	15,29086	Gizi Baik
3	TANZILA N	13,8	95	0,95	15,29086	Gizi Baik
3	M SAKBAN	14,6	96	0,96	15,84201	Gizi Baik
3	M OZIL JAR	12,5	95	0,95	13,85042	Gizi Kurang
3	RAFLIAN A	13	94,2	0,94	14,65013	Gizi Baik

Pada gambar tersebut, pohon pertama dibangun dengan membuat keputusan berdasarkan nilai Indeks Massa Tubuh (IMT) atau Body Mass Index (BMI). Mengacu pada standar WHO, balita dengan IMT di bawah 14 dikategorikan sebagai “gizi kurang”.

Tabel 3. 14 Pohon ke-2 berdasarkan Umur/Usia

ID	NAMA	UMUR (bln)	berdasarkan Usia
1	M ZAFRAN	57	Gizi Kurang
1	RAFAEL JA	39	Gizi Baik
1	DINAR NA	51	Gizi Kurang
1	ADITYA RIS	58	Gizi Kurang
1	ZAURA MIO	50	Gizi Kurang
1	AULAN BA	54	Gizi Kurang
2	SILMI AFIF	52	Gizi Kurang
2	M NUR HA	59	Gizi Kurang
2	TATU DIIGA	54	Gizi Kurang
2	M ALFARIS	45	Gizi Baik
2	JESIFA KHA	43	Gizi Baik
3	DAVINA NU	42	Gizi Baik
3	AZIZAH SA	56	Gizi Kurang
3	AISYAH FIS	49	Gizi Baik
3	TANZILA N	49	Gizi Baik
3	M SAKBAN	46	Gizi Baik
3	M OZIL JAR	43	Gizi Baik
3	RAFLIAN A	43	Gizi Baik

Selanjutnya, pohon keputusan dibangun berdasarkan umur atau usia balita. Mengacu pada standar WHO, balita dengan usia lebih dari 47 bulan dikategorikan sebagai “gizi kurang”.

Tabel 3. 15 Pohon ke-3 berdasarkan Berat Badan (BB)

NAMA	BB	Berdasarkan TB & BB
M ZAFRAN AL G	15,6	Gizi Baik
RAFAEL JAKFAR	13,3	Gizi Kurang
DINAR NADIYA W	15	Gizi Kurang
ADITYA RISQI ANANTA	15,5	Gizi Baik
ZAURA MIQAILA A	12,2	Gizi Kurang
AULAN BASIRA F	15,7	Gizi Baik
SILMI AFIFAH	14,6	Gizi Kurang
M NUR HAFIZAH	18,6	Gizi Baik
TATU DIIGAHAYU	15	Gizi Kurang
M ALFARISQI	13,5	Gizi Kurang
JESIFA KHAIRUNNISA	15,4	Gizi Baik
DAVINA NUR LATIFA	12,2	Gizi Kurang
AZIZAH SALSABILA	15,4	Gizi Baik
AISYAH FISSILMI	13,8	Gizi Kurang
TANZILA NUR F	13,8	Gizi Kurang
M SAKBANI R	14,6	Gizi Kurang
M OZIL JARIS	12,5	Gizi Kurang
RAFLIAN ARKA AYUBI	13	Gizi Kurang

Terakhir, pohon ketiga dibuat berdasarkan berat badan. Mengacu pada standar WHO, balita usia 1–5 tahun dengan berat badan kurang dari 15 kg dikategorikan sebagai “gizi kurang”.

Tabel 3. 16 Ringkasan Klasifikasi

ID	NAMA	Pohon 1	Pohon 2	Pohon 3	Prediksi Akhir (Mayoritas)	PREDIKSI I	AKTUAL
1	M ZAFRAN AL G	Gizi Kurang	Gizi Kurang	Gizi Baik	Gizi Kurang	0	0
1	RAFAEL JAKFAR	Gizi Baik	Gizi Baik	Gizi Kurang	Gizi Baik	1	1
1	DINAR NADIYA W	Gizi Baik	Gizi Kurang	Gizi Kurang	Gizi Kurang	0	0
1	ADITYA RISQI ANANTA	Gizi Kurang	Gizi Kurang	Gizi Baik	Gizi Kurang	0	1
1	ZAURA MIQAILA A	Gizi Kurang	Gizi Kurang	Gizi Kurang	Gizi Kurang	0	0
1	AULAN BASIRA F	Gizi Baik	Gizi Kurang	Gizi Baik	Gizi Baik	1	0
2	SILMI AFIFAH	Gizi Baik	Gizi Kurang	Gizi Kurang	Gizi Kurang	0	0
2	M NUR HAFIZAH	Gizi Baik	Gizi Kurang	Gizi Baik	Gizi Baik	1	0
2	TATU DIIGAHAYU	Gizi Kurang	Gizi Kurang	Gizi Kurang	Gizi Kurang	0	0
2	M ALFARISQI	Gizi Baik	Gizi Baik	Gizi Kurang	Gizi Baik	1	0
2	JESIFA KHAIRUNNISA	Gizi Baik	Gizi Baik	Gizi Baik	Gizi Baik	1	0
3	DAVINA NUR LATIFA	Gizi Baik	Gizi Baik	Gizi Kurang	Gizi Baik	1	1
3	AZIZAH SALSABILA	Gizi Baik	Gizi Kurang	Gizi Baik	Gizi Baik	1	0
3	AISYAH FISSILMI	Gizi Baik	Gizi Baik	Gizi Kurang	Gizi Baik	1	0
3	TANZILA NUR F	Gizi Baik	Gizi Baik	Gizi Kurang	Gizi Baik	1	0
3	M SAKBANI R	Gizi Baik	Gizi Baik	Gizi Kurang	Gizi Baik	1	0
3	M OZIL JARIS	Gizi Kurang	Gizi Baik	Gizi Kurang	Gizi Kurang	0	0
3	RAFLIAN ARKA AYUBI	Gizi Baik	Gizi Baik	Gizi Kurang	Gizi Baik	1	0

Kesimpulan sementara, Random Forest menggunakan prinsip mayoritas untuk menggabungkan klasifikasi dari ketiga pohon sehingga menghasilkan klasifikasi akhir yang lebih robust dan akurat dibandingkan satu pohon tunggal. Pada gambar di atas, output yang dihasilkan masih kurang akurat karena beberapa faktor:

- a) Data mungkin masih kurang karena ini hanya simulasi menggunakan sample data seadanya.
- b) Jumlah "Pohon" Sedikit: Tiga pohon tidak cukup untuk mendapatkan manfaat penuh dari *ensemble learning*.
- c) Belum ada *Bootstrapping* yang sebenarnya, pada simulasi ini tidak secara eksplisit membuat subset data acak dengan penggantian untuk setiap pohon. Semua pohon pada dasarnya melihat semua data (meskipun dengan fokus pada fitur yang berbeda).
- d) Pemilihan Fitur Statis: Kita secara manual memilih fitur untuk setiap "pohon" di awal. Dalam *Random Forest*, pemilihan fitur acak terjadi di setiap node saat pohon dibangun.

3.4.3 Pemetaan Berbasis Spasial

Tabel 3. 17 Zonasi Resiko Gizi Balita

ID	JUMLAH BALITA	GIZI KURANG (X)	GIZI BAIK (Y)	Persentase Risiko (%)	Zona
Kec. Bluto	40	22	18	55,0	Merah
Kec. Raas	40	28	12	70,0	Merah
Kec. Kalianget	40	8	32	20,0	Kuning
Kec. Kota	40	3	37	7,5	Hijau
Kec. Ambunten	40	25	15	62,5	Merah

1. Kecamatan Bluto (55,0%) – Zona Merah

Lebih dari setengah balita masuk kategori gizi kurang. Kondisi ini menunjukkan risiko tinggi terjadinya masalah kesehatan jangka panjang, sehingga perlu intervensi serius dari tenaga kesehatan.

2. Kecamatan Raas (70,0%) – Zona Merah

Persentase tertinggi di antara kecamatan lain. Angka ini mengindikasikan kondisi darurat gizi yang sangat mengkhawatirkan. Intervensi segera dan program pemulihan gizi harus menjadi prioritas utama.

3. Kecamatan Kalianget (20,0%) – Zona Kuning

Masuk kategori waspada. Artinya, meskipun sebagian besar balita dalam kondisi gizi baik, masih ada 20% yang perlu perhatian. Pencegahan lebih dini bisa mencegah wilayah ini bergeser ke zona merah.

4. Kecamatan Kota (7,5%) – Zona Hijau

Kondisi relatif aman karena prevalensi gizi kurang masih di bawah 10%. Meski demikian, tetap diperlukan monitoring berkala agar tidak ada peningkatan kasus.

5. Kecamatan Ambunten (62,5%) – Zona Merah

Tingkat risiko tinggi, lebih dari separuh balita bermasalah gizi. Sama seperti Bluto dan Raas, perlu adanya intervensi intensif.

BAB IV

KLASIFIKASI STATUS GIZI DENGAN ANN

4.1. Arsitektur ANN dan Parameter Pelatihan

Arsitektur jaringan saraf yang digunakan dalam penelitian ini dirancang untuk mengolah 11 fitur input yang merepresentasikan variabel penentu status gizi. Model dibangun menggunakan struktur multi-layer perceptron (MLP). Pada lapisan input, digunakan 256 neuron dengan fungsi aktivasi ReLU yang mampu mempercepat konvergensi dan mengatasi vanishing gradient. Proses pelatihan distabilkan dengan Batch Normalization dan dicegah dari overfitting menggunakan Dropout sebesar 0,3. Selanjutnya, terdapat tiga lapisan tersembunyi dengan jumlah neuron 128, 64, dan 32, masing-masing dilengkapi ReLU, Batch Normalization, dan Dropout 0,24. Penurunan jumlah neuron pada tiap lapisan ini bertujuan menyederhanakan representasi data secara bertahap sehingga jaringan dapat belajar pola yang lebih relevan tanpa kehilangan informasi penting.

Table 4. 1 Tabel Arsitektur Jaringan Neural Network

Lapisan	Jumlah Neuron	Fungsi Aktivasi	Batch Normalization	Dropout Rate	Keterangan
Input Layer	256	ReLU	Ya	0.30	Memproses 11 fitur input dan menstabilkan distribusi data awal
Hidden Layer 1	128	ReLU	Ya	0.24	Ekstraksi pola kompleks dari fitur awal
Hidden Layer 2	64	ReLU	Ya	0.24	Reduksi kompleksitas representasi data

Hidden Layer 3	32	ReLU	Ya	0.24	Memperkuat pembelajaran fitur yang lebih spesifik
Output Layer*	2	Softmax	tidak	-	Menghasilkan klasifikasi status gizi

Secara keseluruhan, arsitektur ini menunjukkan bahwa model ANN yang dibangun mengutamakan keseimbangan antara kompleksitas dan generalisasi. Dengan adanya kombinasi ReLU, Batch Normalization, dan Dropout, jaringan diharapkan dapat mencapai performa yang stabil, akurat, dan mampu menghindari masalah overfitting meskipun data yang digunakan memiliki variasi yang cukup tinggi.

4.2. Konfigurasi Pelatihan Jaringan Neural (Training Configuration)

Tahap konfigurasi pelatihan menetapkan parameter dan teknik yang digunakan dalam proses pembelajaran ANN, termasuk pemilihan optimizer, fungsi kehilangan, metrik evaluasi, dan teknik regularisasi untuk mencegah overfitting. Pengaturan ini memastikan model belajar secara optimal, stabil, dan mampu menghasilkan akurasi tinggi dalam klasifikasi status gizi.

4.3.1. Optimizer

Dalam penelitian ini digunakan Adam Optimizer (Adaptive Moment Estimation) karena memiliki kemampuan adaptif dalam menyesuaikan learning rate dan mempercepat proses konvergensi. Adam menggabungkan konsep Momentum dan RMSProp, sehingga lebih efisien untuk data berdimensi tinggi dan bersifat non-linear.

Table 4. 2 Konfigurasi Optimizer Model ANN

Parameter	Nilai / Jenis	Keterangan
Algoritma Optimasi	Adam Optimizer	Menggabungkan metode Momentum dan RMSProp
Learning Rate	0.001	Mengatur kecepatan pembaruan bobot
Beta 1	0.9	Faktor peluruhan momentum pertama
Beta 2	0.999	Faktor peluruhan momentum kedua
Epsilon	1e-07	Menghindari pembagian dengan nol

4.3.2. Parameter Pelatihan

Parameter pelatihan menentukan bagaimana model ANN belajar dari data. Pengaturan ini mencakup fungsi kehilangan, metrik evaluasi, ukuran batch, jumlah epoch, serta pembagian data validasi.

Table 4. 3 Parameter Pelatihan Model ANN

Parameter	Nilai / Jenis	Keterangan
Loss Function	Sparse Categorical Crossentropy	Cocok untuk klasifikasi multi-kelas dengan label integer
Metrics	Accuracy	Mengukur tingkat keberhasilan klasifikasi
Batch Size	32	Jumlah data per iterasi pembelajaran
Epochs	200	Jumlah siklus pembelajaran penuh
Validation Split	0.2 (20%)	Proporsi data validasi terhadap data latih

4.3.3. Regularisasi

Regularisasi digunakan untuk mencegah overfitting agar model tidak hanya menghafal data latih, tetapi juga mampu melakukan generalisasi terhadap data baru. Beberapa teknik yang digunakan adalah Dropout, Batch Normalization, dan Early Stopping.

Table 4. 4 Teknik Regularisasi Model ANN

Teknik Regularisasi	Parameter / Nilai	Keterangan
Dropout	Input Layer: 0.3 Hidden Layers: 0.24	Menghapus neuron secara acak untuk mencegah overfitting
Batch Normalization	Diterapkan di setiap layer	Menormalkan aktivasi dan mempercepat konvergensi
Early Stopping	Patience = 20, Monitor = 'val_loss'	Menghentikan pelatihan saat tidak ada peningkatan performa

4.3. Hyperparameter Tuning Grid

Tahap hyperparameter tuning dilakukan untuk menemukan kombinasi parameter terbaik bagi model ANN. Proses ini menguji variasi nilai seperti learning rate, dropout rate, arsitektur hidden layer, dan fungsi aktivasi. Pendekatan grid search digunakan untuk mengevaluasi setiap kombinasi secara menyeluruh, sehingga diperoleh konfigurasi dengan akurasi tertinggi dan error terendah..

4.4.1. Learning Rate Options

Learning rate menentukan besarnya langkah pembaruan bobot pada setiap iterasi. Nilai yang terlalu besar membuat model tidak stabil, sedangkan nilai terlalu kecil memperlambat proses belajar. Karena itu, penelitian ini menguji tiga variasi learning rate untuk menemukan nilai yang paling optimal.

Table 4. 5 Variasi Learning Rate yang diuji

Nilai Learning Rate	Keterangan
0.001	Nilai umum untuk pelatihan model standar
0.0005	Nilai sedang, memberikan keseimbangan antara stabilitas dan kecepatan konvergensi
0.0001	Nilai kecil, digunakan untuk pelatihan yang lebih

	hati-hati agar tidak melewati minimum lokal
--	---

4.4.2. *Dropout Rate Options*

Dropout rate digunakan untuk mencegah overfitting dengan cara mengabaikan sebagian neuron secara acak selama pelatihan. Beberapa nilai dropout diuji untuk mencari keseimbangan antara kapasitas pembelajaran dan kemampuan generalisasi.

Table 4. 6 Variasi Dropout Rate yang diuji

Nilai Dropout Rate	Keterangan
0.2	Regularisasi ringan, menjaga sebagian besar neuron aktif
0.3	Regularisasi sedang, mencegah overfitting tanpa menurunkan kapasitas belajar
0.4	Regularisasi kuat, cocok untuk data kompleks namun berisiko memperlambat pembelajaran

4.4.3. *Hidden Layer Architectures*

Struktur hidden layer memengaruhi kompleksitas model dan kemampuannya mengekstraksi fitur. Penelitian ini menguji tiga arsitektur dengan variasi jumlah lapisan dan neuron untuk memperoleh keseimbangan terbaik antara akurasi dan efisiensi

Table 4. 7 Arsitektur Hidden layer yang diuji

Struktur Lapisan Tersembunyi	Keterangan
[128, 64]	Struktur sederhana dengan dua lapisan tersembunyi
[256, 128, 64]	Struktur kompleks dengan kapasitas pembelajaran tinggi
[128, 64, 32]	Struktur menengah dengan keseimbangan antara kompleksitas dan kecepatan pelatihan

4.4.4. Activation Functions

Fungsi aktivasi digunakan untuk memberikan non-linearitas pada jaringan, sehingga model dapat belajar pola yang lebih kompleks. Dua jenis fungsi aktivasi diuji untuk mengetahui pengaruhnya terhadap performa jaringan.

Table 4. 8 Fungsi Aktivasi yang diuji

Fungsi Aktivasi	Keterangan
ReLU	Cepat dan stabil, umum digunakan untuk data non-linear
Tanh	Memberikan keluaran antara -1 hingga 1, efektif untuk data dengan rentang kecil

4.4. Skenario Pelatihan Model ANN

Skenario pelatihan dilakukan untuk memperoleh performa ANN yang optimal dalam mengklasifikasi status gizi anak. Tiga skenario pembagian data diterapkan, yakni 80:20, 60:40, dan 70:30, untuk menilai konsistensi dan kestabilan model terhadap variasi proporsi data latih dan uji. Pendekatan ini membantu melihat pengaruh perubahan proporsi data terhadap akurasi dan kemampuan generalisasi model.

4.5.1. Skenario Awal (Default)

Model Artificial Neural Network (ANN) dikembangkan dengan konfigurasi arsitektur yang akan di tampilkan pada tabel berikut:

Gambar 4. 1 Hyperparamter ANN

```
param_grid = {
    'learning_rate': [0.001, 0.0005, 0.0001],
    'dropout_rate': [0.2, 0.3, 0.4],
    'hidden_layers': [[128, 64], [256, 128, 64], [128, 64, 32]],
    'activation': ['relu', 'tanh']
}
```

Pada gambar di atas merupakan parameter-parameter yang akan dijadikan sebagai acuan yang akan d kombinasi untuk dicari kombinasi parameter yang terbaik.

Table 4. 9 parameter Terbaik ANN Skenario Awal (80-20)

Best parameters:	
learning_rate	0.0005
dropout_rate	0.3
hidden_layers	[128, 64, 32]
activation	relu

Dari tabel di atas merupakan hasil parameter yang terbaik, yang nantinya kan dijadikan sebagai parameter untuk menentukan status gizi.

Table 4. 10 Hasil pelatihan model ANN per-epoch

Epoch	Accuracy (Train)	Loss (Train)	Val_Accuracy	Val_Loss	Learning Rate
1/200	0.9636	0.3186	0.9420	0.1268	0.0005
2/200	0.9333	0.2375	0.9324	0.1679	0.0005
3/200	0.9297	0.2184	0.9517	0.1396	0.0005
4/200	0.9212	0.2380	0.9517	0.1317	0.0005
5/200	0.9333	0.1836	0.9469	0.1530	0.0005
6/200	0.9200	0.2063	0.9614	0.1377	0.0005
7/200	0.9333	0.1849	0.9517	0.1365	0.0005
8/200	0.9248	0.1948	0.9517	0.1489	0.0005
9/200	0.9345	0.1981	0.9324	0.1738	0.0005
10/200	0.9200	0.1980	0.9420	0.1723	0.0005
11/200	0.9382	0.1979	0.9372	0.1576	0.0005
12/200	0.9261	0.1938	0.9469	0.1564	0.00035
13/200	0.9345	0.1677	0.9469	0.1449	0.00035
14/200	0.9394	0.1587	0.9517	0.1370	0.00035
15/200	0.9394	0.1602	0.9517	0.1339	0.00035
16/200	0.9455	0.1755	0.9469	0.1399	0.00035
17/200	0.9370	0.1710	0.9517	0.1475	0.00035
18/200	0.9345	0.1474	0.9469	0.1510	0.00035
19/200	0.9442	0.1606	0.9469	0.1517	0.00035
20/200	0.9358	0.1792	0.9469	0.1643	0.00035
21/200	0.9394	0.1508	0.9517	0.1436	0.00035

Hasil skenario awal (80–20), model ANN mencapai performa optimal pada epoch ke-21 dengan akurasi pelatihan 93,94% dan akurasi

validasi 95,17%. Nilai loss turun hingga 0,1508 (latih) dan 0,1436 (validasi), menunjukkan proses belajar yang stabil. Selisih akurasi yang kecil menandakan model tidak mengalami overfitting dan mampu melakukan generalisasi dengan baik.

Table 4. 11 Hasil Klasifikasi ANN Skenario Awal (80-20)

Kecamatan	Total Anak	Gizi Baik (Real)	Gizi Kurang (Real)	Prediksi Gizi Baik (ANN)	Prediksi Gizi Kurang ANN
AMBUNTEN	40	36	4	36	4
ARIASA	15	15	0	14	1
BATANG-	30	30	0	29	1
BATUAN	54	54	0	51	3
BATUPUTIH	50	35	15	34	16
BLUTO	40	39	1	39	1
DASUK	55	44	11	44	11
DUNGKEK	70	61	9	59	11
GANDING	70	57	13	50	20
GAPURA	40	38	2	37	3
GAYAM	15	11	4	11	4
GILIGINTING	50	45	5	45	5
GULUK-GULUK	49	41	8	38	11
KALIANGET	60	58	2	54	6
KANGAYAN	35	29	6	29	6
KANGEAN	50	46	4	41	9
Kota Sumenep	90	90	0	90	0
LENTENG	68	59	9	57	11
MANDING	50	48	2	46	4
MASALEMBU	40	37	3	37	3
NONGGUNON	35	26	9	26	9
PASONGSONG	50	47	3	45	5
PRAGAAN	20	18	2	16	4
RAAS	35	31	4	29	6
RUBARU	45	26	19	23	22
SAPEKEN	20	17	3	17	3
SARONGGI	45	42	3	40	5
TALANGO	70	68	2	68	2

Hasil tabel menunjukkan bahwa model ANN mampu mengklasifikasi status gizi anak dengan tingkat kesesuaian tinggi pada hampir seluruh kecamatan. Sebagian besar prediksi sesuai dengan kondisi aktual, dengan hanya sedikit perbedaan di beberapa wilayah. Secara keseluruhan, performa model sangat baik dengan akurasi 95,8%, presisi 96,4%, recall 95,8%, dan

F1-score 95,9%, yang mengindikasikan ketepatan dan keseimbangan model dalam mengenali status gizi anak.

4.5.2. Skenario 1 (60-40)

Model Artificial Neural Network (ANN) Skenario-1 dikembangkan dengan hasil best parameter yang akan di tampilkan pada tabel berikut:

Table 4. 12 Parameter Terbaik ANN Skenario-1 (60-40)

Parameter	Nilai / Konfigurasi
Learning Rate	0.0005
Dropout Rate	0.4
Hidden Layers	[128, 64, 32]
Activation Function	<i>tanh</i>

Dari tabel di atas merupakan hasil parameter yang terbaik, dari split data 60-40 yang nantinya akan dijadikan sebagai parameter untuk menentukan status gizi.

Table 4. 13 Hasil pelatihan model ANN per-epoch Skenario-1

Epoch	Accuracy (Train)	Loss (Train)	Val_Accuracy	Val_Loss	Leaming Rate
1	0.9063	0.4697	0.8968	0.2218	0.0005
5	0.8595	0.3522	0.9097	0.2147	0.0005
10	0.8530	0.3391	0.8968	0.2588	0.0005
15	0.8643	0.3152	0.9161	0.1984	0.0005
20	0.8805	0.2575	0.9161	0.1913	0.0005
25	0.8837	0.2719	0.9290	0.1897	0.0005
30	0.8853	0.3165	0.9290	0.1799	0.0005
35	0.8982	0.2281	0.9419	0.1612	0.0005
40	0.9031	0.2396	0.9419	0.1753	0.00035
45	0.8837	0.2603	0.9290	0.1931	0.00035
49	0.8966	0.2198	0.9355	0.1752	0.00035

Model ANN menunjukkan peningkatan akurasi yang stabil hingga mencapai val_accuracy tertinggi 94,19% pada epoch ke-35, menandakan kemampuan generalisasi yang baik terhadap data validasi.

Table 4. 14 Hasil Klasifikasi ANN Skenario-1 (60-30)

Kecamatan	Total	Gizi Baik	Gizi Kurang	Prediksi ANN	
				Gizi Kurang ANN	Gizi Baik
AMBUNTEN	40	36	4	5	35
ARJASA	15	15	0	0	15
BATANG-BATA	30	30	0	0	30
BATUAN	54	54	0	4	50
BATUPUTIH	50	35	15	17	33
BLUTO	40	39	1	1	39
DASUK	55	44	11	12	43
DUNGKEK	70	61	9	14	56
GANDING	70	57	13	23	47
GAPURA	40	38	2	3	37
GAYAM	15	11	4	4	11
GILIGINTING	50	45	5	5	45
GULUK-GULUK	49	41	8	7	42
KALIANGET	60	58	2	10	50
KANGAYAN	35	29	6	17	18
KANGEAN	50	46	4	9	41
Kota Sumenep	90	90	0	2	88
LENTENG	68	59	9	10	58
MANDING	50	48	2	5	45
MASALEMBU	40	37	3	4	36
NONGGUNONG	35	26	9	11	24
PASONGSONGA	50	47	3	3	47
PRAGAAN	20	18	2	6	14
RUBARU	45	26	19	24	21
Ra'as	35	31	4	7	28
SAPEKEN	20	17	3	4	16
SARONGGI	45	42	3	5	40
TALANGO	70	68	2	3	67

Hasil tabel menunjukkan bahwa model ANN memiliki kemampuan klasifikasi yang tinggi terhadap data aktual di sebagian besar kecamatan. Mayoritas hasil prediksi sesuai dengan kondisi sebenarnya, dengan hanya sedikit perbedaan pada beberapa wilayah. Secara umum, model menunjukkan performa sangat baik dengan akurasi 94,8%, presisi 97,99%, recall 98,87%, dan F1-score 97,91%. Hal ini menegaskan bahwa ANN mampu mengidentifikasi status gizi anak dengan tingkat ketelitian dan konsistensi yang sangat tinggi.

4.5.3. Skenario 2 (70-30)

Model Artificial Neural Network (ANN) Skenario-2 dikembangkan dengan hasil best parameter yang akan di tampilkan pada tabel berikut:

Table 4. 15 parameter Terbaik ANN Skenario-2 (70-30)

Parameter	Nilai
Leaming Rate	0.001
Dropout Rate	0.2
Hidden Layers	[128, 64]
Activation	tanh
Optimizer	Adam (default)

Dari tabel di atas merupakan hasil parameter yang terbaik, dari split data 70-30 yang nantinya akan dijadikan sebagai parameter untuk menentukan status gizi.

Table 4. 16 Hasil pelatihan model ANN per-epoch Skenario-2

Epoch	Accuracy	Loss	Val_Accuracy	Val_Loss	Learning Rate
1	0.9418	0.2029	0.9282	0.1874	0.0010
2	0.9238	0.2034	0.9282	0.1646	0.0010
3	0.9280	0.1930	0.9669	0.1153	0.0010
10	0.8961	0.1971	0.9503	0.1244	0.0010
11	0.9280	0.1762	0.9558	0.1419	0.0010
12	0.9363	0.1607	0.9613	0.1318	0.0010
16	0.9266	0.1622	0.9448	0.1426	0.0007
17	0.9266	0.1766	0.9613	0.1292	0.0007
18	0.9294	0.1416	0.9724	0.1286	0.0007
19	0.9321	0.1330	0.9613	0.1288	0.0007
20	0.9335	0.1480	0.9558	0.1482	0.0007
21	0.9183	0.1518	0.9669	0.1341	0.0007
22	0.9211	0.1551	0.9503	0.1687	0.0007
23	0.9321	0.1367	0.9613	0.1340	0.0007

Model ANN menunjukkan peningkatan akurasi yang stabil hingga mencapai val_accuracy tertinggi 97,29% pada epoch ke-18, menandakan kemampuan generalisasi yang baik terhadap data validasi.

Table 4. 17 Hasil Klasifikasi ANN Skenario-2

Kecamatan	Total	Gizi Baik	Gizi Kurang	Hasil Prediksi ANN	
				Gizi Baik ANN	Gizi Kurang ANN
AMBUNTEN	40	36	4	35	5
ARJASA	15	15	0	14	1
BATANG-BATAN	30	30	0	29	1
BATUAN	54	54	0	50	4
BATUPUTIH	50	35	15	32	18
BLUTO	40	39	1	39	1
DASUK	55	44	11	43	12
DUNGKEK	70	61	9	57	13
GANDING	70	57	13	51	19
GAPURA	40	38	2	37	3
GAYAM	15	11	4	11	4
GILIGINTING	50	45	5	45	5
GULUK-GULUK	49	41	8	43	6
KALIANGET	60	58	2	56	4
KANGAYAN	35	29	6	29	6
KANGEAN	50	46	4	42	8
Kota Sumenep	90	90	0	87	3
LENTENG	68	59	9	58	10
MANDING	50	48	2	47	3
MASALEMBU	40	37	3	37	3
NONGGUNONG	35	26	9	25	10
PASONGSONGAN	50	47	3	44	6
PRAGAAN	20	18	2	16	4
RUBARU	45	26	19	25	20
Ra'as	35	31	4	29	6
SAPEKEN	20	17	3	16	4
SARONGGI	45	42	3	38	7
TALANGO	70	68	2	67	3

Model ANN menunjukkan kemampuan klasifikasi yang tinggi dengan mayoritas hasil prediksi sesuai data aktual. Pada skenario ini, model mencapai akurasi 95,88%, presisi 96,64%, recall 95,88%, dan F1-score 90,09%. Nilai akurasi tersebut lebih baik dibandingkan skenario sebelumnya, menandakan peningkatan performa model dalam mengklasifikasi status gizi.

4.5.4. Perbandingan Hasil Setiap Skenario

Setelah menyelesaikan tahap pengujian pada Skenario awal, skenario 1 dan skenario 2, hasil kedua skenario dapat dilihat pada tabel berikut:

Table 4. 18 Perbandingan performa setiap skenario

Skenario	Akurasi ANN
Skenario Awal (Default)	95.8 %
Skenario -1	94.58 %
Skenario -2	95.88 %

Hasil perbandingan menunjukkan adanya peningkatan kinerja model pada Skenario 2 yaitu akurasi sebesar 95.88 %

BAB V

KLASIFIKASI STATUS GIZI DENGAN RANDOM FOREST

5.1. Arsitektur Random Forest

Model Random Forest (RF) adalah algoritma ensemble learning yang terdiri dari banyak pohon keputusan (Decision Trees) yang dilatih secara terpisah. Setiap pohon menghasilkan prediksi, kemudian keputusan akhir ditentukan melalui voting mayoritas. Dengan memanfaatkan variasi sampel dan fitur pada setiap pohon, Random Forest menjadi lebih stabil dan tahan terhadap overfitting dibandingkan satu pohon tunggal, sehingga menghasilkan performa klasifikasi yang lebih akurat.

5.1.1. Base Estimator

Setiap pohon dalam Random Forest berfungsi sebagai base estimator yang mempartisi data berdasarkan fitur untuk memisahkan kelas target. Kriteria pemisahan yang digunakan adalah Gini Impurity, di mana nilai yang lebih kecil menunjukkan kelompok data yang lebih homogen dan terstruktur dengan baik.

Tabel 5. 1 Parameter Base Estimator pada Model Random Forest

Komponen	Nilai / Keterangan
Algoritma Dasar	Decision Tree
Kriteria Pemisahan (Criterion)	Gini Impurity (default)
Strategi Pemisahan (Split Strategy)	Pemilihan <i>best split</i> pada setiap node untuk memaksimalkan kehomogenan kelas

5.1.2. Ensemble Parameters

Arsitektur Random Forest ditentukan oleh jumlah pohon (50–500) untuk memperoleh akurasi optimal dan efisiensi komputasi. Parameter

bootstrap mengatur pengambilan sampel acak dengan pengembalian, sedangkan random state memastikan hasil yang konsisten dan dapat direproduksi. Tabel 5. 2 Parameter Ensemble Random Forest

Parameter	Nilai / Rentang Uji	Keterangan
Number of Estimators	50 – 500 pohon	Jumlah pohon keputusan dalam satu model Random Forest
Bootstrap	True / False	Menentukan apakah data diambil dengan pengembalian
Random State	42	Menjamin hasil eksperimen dapat direplikasi secara konsisten

5.2. Hyperparameter Random Forest

Model Random Forest dioptimalkan dengan pengaturan hyperparameter seperti struktur pohon, bobot kelas, dan konfigurasi pelatihan. Evaluasi dilakukan menggunakan random search dan 10-fold cross-validation untuk menemukan kombinasi yang memberikan akurasi, presisi, dan generalisasi terbaik.

5.2.1. Tree Structure Parameters

Parameter struktur pohon menentukan kompleksitas setiap base estimator dalam Random Forest. Variasi parameter ini mempengaruhi kemampuan model dalam menangkap pola data tanpa menimbulkan overfitting.

Tabel 5. 3 Parameter Struktur Pohon yang Diuji

Parameter	Nilai / Rentang Uji	Keterangan
Max Depth	5, 10, 15, 20, 25, 30, None	Kedalaman maksimum setiap pohon; None berarti tidak dibatasi
Min Samples	2, 5, 10, 15, 20	Jumlah minimal sampel untuk membagi

Split		node
Min Samples Leaf	1, 2, 4, 6, 8	Jumlah minimal sampel pada setiap daun
Max Features	'sqrt', 'log2', 0.5, 0.7, 0.9, None	Jumlah fitur yang dipertimbangkan saat mencari <i>best split</i>

5.2.2. Class Weight Options

Pada data yang tidak seimbang, pengaturan class weight sangat penting untuk mencegah bias terhadap kelas mayoritas. Penelitian ini menguji beberapa opsi bobot kelas.

Tabel 5. 4 Opsi Bobot Kelas Random Forest

Opsi	Keterangan
Balanced	Menyeimbangkan secara otomatis proporsi kelas
Balanced Subsample	Menyeimbangkan setiap bootstrap sample
Custom Weight	Memberikan bobot sesuai distribusi kelas yang diinginkan

5.2.3. Training Parameters

Random Forest dilatih menggunakan random search untuk eksplorasi hyperparameter, dengan 10-fold cross-validation agar evaluasi stabil, serta dievaluasi menggunakan berbagai metrik untuk menilai performa secara menyeluruh.

Tabel 5. 5 Parameter Pelatihan Random Forest

Parameter	Nilai / Jenis	Keterangan
Cross Validation	10-fold CV	Membagi data menjadi 10 subset untuk evaluasi yang robust
Scoring	accuracy, precision_weighted, recall_weighted, f1_weighted	Metrik evaluasi model multi-kelas
Random Search	100 iterasi	Jumlah percobaan kombinasi parameter

Parallel Processing	n_jobs = -1	Menggunakan semua core CPU untuk mempercepat proses
---------------------	-------------	---

5.3. Hyperparameter Tuning Grid

Hyperparameter tuning Random Forest dilakukan untuk menemukan kombinasi parameter terbaik dengan grid search atau random search, sehingga model mampu belajar pola data secara optimal, menghindari overfitting, dan menghasilkan klasifikasi status gizi yang akurat.

5.3.1. Number of Estimators

Jumlah pohon dalam hutan keputusan mempengaruhi stabilitas dan akurasi model. Semakin banyak pohon, model cenderung lebih stabil namun memerlukan waktu komputasi lebih lama.

Tabel 5. 6 Jumlah Estimator yang Diuji

Number of Estimators	Keterangan
50	Hutan kecil, cepat namun kurang stabil
100	Ukuran menengah, akurasi meningkat
200	Keseimbangan antara performa dan waktu komputasi
300	Performa stabil, waktu komputasi lebih lama
500	Performa maksimal, biaya komputasi tinggi

5.3.2. Max Depth Options

Kedalaman maksimum pohon mempengaruhi kompleksitas model. Kedalaman terbatas dapat mencegah overfitting, sedangkan kedalaman tak terbatas memungkinkan model menangkap pola lebih kompleks.

Tabel 5. 7 Max Depth yang Diuji

Max Depth	Keterangan
5	Kedalaman sangat dangkal, risiko underfitting tinggi
10	Dangkal, masih sederhana

15	Cukup kompleks, performa meningkat
20	Kompleks, risiko overfitting sedikit lebih tinggi
25	Kompleks, potensi overfitting meningkat
30	Sangat kompleks, overfitting mungkin terjadi
None	Tidak dibatasi, pohon dapat berkembang penuh

5.3.3. Min Samples Split Options

Jumlah minimum sampel untuk membagi node memengaruhi ukuran dan kompleksitas pohon. Nilai kecil memungkinkan pohon lebih detail namun bisa overfit.

Tabel 5. 8 Min Samples Split yang Diuji

Min Samples Split	Keterangan
2	Node bisa terbagi dengan sedikit data, risiko overfitting tinggi
5	Split moderat, lebih stabil
10	Node hanya terbagi jika ada data cukup
15	Node besar, model lebih sederhana
20	Node besar, model sangat sederhana

5.3.4. Min Samples Leaf Options

Jumlah minimum sampel pada daun memengaruhi ukuran akhir daun. Leaf besar mencegah overfitting, leaf kecil membuat pohon lebih detail.

Tabel 5. 9 Min Samples Leaf yang Diuji

Min Samples Leaf	Keterangan
1	Daun paling kecil, detail maksimal, risiko overfitting tinggi
2	Sedikit regularisasi
4	Leaf sedang, keseimbangan antara detail dan generalisasi
6	Leaf besar, lebih generalisasi
8	Leaf sangat besar, model sederhana

5.3.5. Max Features Options

Jumlah fitur yang dipertimbangkan saat mencari best split memengaruhi diversitas pohon.

Tabel 5. 10 Max Features yang Diuji

Opsi	Keterangan
'sqrt'	Menggunakan akar kuadrat dari jumlah fitur
'log2'	Menggunakan logaritma basis 2 dari jumlah fitur
0.5	Menggunakan 50% fitur acak
0.7	Menggunakan 70% fitur acak
0.9	Menggunakan 90% fitur acak
None	Menggunakan semua fitur

5.3.6. Bootstrap Options

Pengambilan sampel dengan atau tanpa pengembalian memengaruhi keragaman pohon dalam hutan.

Tabel 5. 11 Opsi Bootstrap

Opsi	Keterangan
True	Mengambil sampel dengan pengembalian, meningkatkan diversitas pohon
False	Tanpa pengembalian, setiap pohon dibangun dari sampel unik

5.4. Random Forest Optimal Configuration

Hasil hyperparameter tuning Random Forest menunjukkan konfigurasi optimal yang memberikan keseimbangan akurasi, stabilitas, dan efisiensi, yaitu: jumlah estimators 200–500, max depth 15–25, min samples split 5–10, min samples leaf 2–4, max features 'sqrt' atau 0,7, dan bootstrap True. Kombinasi ini memungkinkan model mengklasifikasi status gizi secara akurat tanpa overfitting.

Tabel 5. 12 Konfigurasi Optimal Model Random Forest

Parameter	Nilai Optimal	Keterangan
n_estimators	200–500	Jumlah pohon keputusan yang memberikan keseimbangan antara akurasi dan efisiensi komputasi
max_depth	15–25	Kedalaman pohon yang optimal untuk menangkap pola tanpa overfitting
min_samples_split	5–10	Minimum sampel untuk memisahkan node, mencegah pembentukan pohon yang terlalu detail
min_samples_leaf	2–4	Minimum sampel di setiap daun untuk menjaga generalisasi
max_features	'sqrt' atau 0.7	Proporsi fitur yang digunakan dalam setiap split, meningkatkan diversitas pohon
bootstrap	True	Mengaktifkan pengambilan sampel dengan pengembalian untuk menambah variasi pohon

5.4.1. Jumlah Estimator (200–500 Pohon)

Jumlah pohon berperan penting dalam menentukan stabilitas hasil klasifikasi.

- Jika terlalu sedikit (<100), model cenderung memiliki high variance karena setiap pohon sangat bergantung pada data latih tertentu.
- Jika terlalu banyak (>1000), peningkatan akurasi biasanya tidak signifikan lagi, namun waktu komputasi meningkat tajam.
- Oleh karena itu, jumlah **antara 200 hingga 500 pohon** memberikan hasil paling seimbang antara performa dan efisiensi.

5.4.2. Max Depth (15–25 Level)

Kedalaman pohon menentukan kompleksitas model dalam mempelajari hubungan antar fitur.

- Pohon yang terlalu dalam (>30) dapat menyesuaikan diri terlalu baik pada data latih dan menimbulkan overfitting.
- Sebaliknya, kedalaman yang terlalu dangkal (<10) dapat membuat model tidak cukup menangkap pola kompleks (underfitting).

- Rentang **15–25** terbukti mampu menangkap interaksi antar fitur yang relevan, terutama untuk dataset dengan 11 fitur input, tanpa kehilangan interpretabilitas model.

5.4.3. *Min Samples Split (5–10)*

Parameter ini mengontrol jumlah minimum sampel yang diperlukan agar suatu node dapat dibagi.

- Nilai kecil memungkinkan banyak pemisahan, namun meningkatkan risiko overfitting.
- Nilai besar membuat model lebih konservatif dan stabil.
- Rentang **5–10** memberikan keseimbangan terbaik antara stabilitas, kompleksitas, dan kecepatan pembelajaran.

5.4.4. *Min Samples Leaf (2–4)*

Menentukan jumlah minimum sampel dalam setiap daun (terminal node).

- Daun dengan terlalu sedikit sampel dapat menyebabkan model belajar dari noise atau data yang tidak representatif.
- Nilai 2–4 menjaga keseimbangan antara detail dan kemampuan generalisasi model, sekaligus memberikan statistical significance yang cukup pada setiap node akhir.

5.4.5. *Max Features ('sqrt' atau 0.7)*

Parameter ini mengontrol jumlah fitur yang dipertimbangkan pada setiap pemisahan node.

- Penggunaan 'sqrt' (sekitar 3–4 dari 11 fitur) dan 0.7 (sekitar 7–8 fitur) mendorong keragaman antar pohon dan mengurangi bias.
- Pendekatan ini mempercepat proses pelatihan, menghindari overfitting, dan tetap menjaga kemampuan model dalam menangkap pola penting antar fitur.

5.4.6. *Bootstrap = True*

Dengan bootstrap sampling, setiap pohon dilatih menggunakan subset data yang diambil secara acak dengan pengembalian.

- Strategi ini meningkatkan keragaman pohon, mengurangi variansi, dan memperkuat kemampuan generalisasi model.
- Selain itu, bootstrap memberikan estimasi performa yang lebih robust karena setiap pohon dibangun dari data yang sedikit berbeda.

5.5. Skenario Pelatihan Model

Skenario pelatihan Random Forest dilakukan untuk memperoleh performa optimal dalam klasifikasi status gizi anak dengan menguji tiga pembagian data: skenario awal 80% training–20% testing, skenario 1 dengan 60%–40%, dan skenario 2 dengan 70%–30%. Pendekatan ini bertujuan menilai konsistensi dan kemampuan generalisasi model terhadap variasi proporsi data..

Tabel 5. 13 Skenario Pelatihan Model Random Forest

Skenario	Data Training	Data Uji (Testing)
Skenario Awal	80 %	20 %
Skenario 1	60%	40%
Skenario 2	70%	30%

5.5.1. Skenario Awal (Default) 80-20

Dalam skenario ini, data status gizi balita dibagi 80% untuk pelatihan dan 20% untuk pengujian, agar model Random Forest dapat belajar dari data historis sekaligus diuji pada data baru. Hasil klasifikasi kemudian dibandingkan dengan data aktual untuk menghitung akurasi dan tingkat kesalahan, sebagai dasar evaluasi performa model dalam mengklasifikasi status gizi di tiap kecamatan.

Tabel 5. 14 Hasil Evaluasi Cross-Validation Random Forest

Metrik	Nilai Rata-rata	Standar Deviasi ($\pm$)
Accuracy	0.9855	0.0320
Precision	0.9861	0.0314
Recall (Sensitivity)	0.9855	0.0320
F1-Score	0.9855	0.0320

Hasil Cross Validation menunjukkan bahwa model Random Forest memiliki performa sangat baik dan stabil dengan nilai accuracy 0.9855, precision 0.9861, recall 0.9855, dan F1-score 0.9855. Nilai standar deviasi yang kecil menandakan konsistensi model dalam melakukan klasifikasi pada setiap lipatan validasi.

Tabel 5. 15 Tabel Hasil Klasifikasi ANN per Kecamatan

Kecamatan	Total Anak	Gizi Baik (Real)	Gizi Kurang (Real)	Prediksi Gizi Baik (RF)	Prediksi Gizi Kurang RF
AMBUNTEN	40	36	4	36	4
ARIASA	15	15	0	15	0
BATANG-BATANG	30	30	0	30	0
BATUAN	54	54	0	54	0
BATU PUTIH	50	35	15	35	15
BLUTO	40	39	1	39	1
DASUK	55	44	11	44	11
DUNGKEK	70	61	9	61	9
GANDING	70	57	13	58	12
GAPURA	40	38	2	38	2
GAYAM	15	11	4	11	4
GILIGINTING	50	45	5	46	4
GULUK-GULUK	49	41	8	41	8
KALIANGET	60	58	2	58	2
KANGAYAN	35	29	6	29	6
KANGEAN	50	46	4	45	5
Kota Sumenep	90	90	0	90	0
LENTENG	68	59	9	59	9
MANDING	50	48	2	49	1
MASALEMBU	40	37	3	37	3
NONGGUNONG	35	26	9	26	9
PASONGSONGAN	50	47	3	47	3
PRAGAAN	20	18	2	17	3
RAAS	35	31	4	30	5
RUBARU	45	26	19	26	19
SAPEKEN	20	17	3	17	3
SARONGGI	45	42	3	42	3
TALANGO	70	68	2	68	2

Hasil menunjukkan bahwa model Random Forest mampu melakukan klasifikasi status gizi anak dengan akurat dan konsisten. Meskipun terdapat beberapa kesalahan di beberapa kecamatan, performa model tetap tinggi

dengan akurasi 94,58%, presisi 96,5%, recall 94,58%, dan F1-score 94,98%, menegaskan keefektifan dan keandalannya dalam klasifikasi gizi di berbagai kecamatan.

5.5.2. Skenario 1 (60-40)

Pada skenario Pertama, data dibagi menjadi 60% untuk pelatihan dan 40% untuk pengujian untuk mengevaluasi pengaruh peningkatan proporsi data pelatihan terhadap performa model Random Forest dalam klasifikasi status gizi balita. Hasil pengaturan hyperparameter terbaik untuk skenario ini disajikan pada Tabel berikut:

Tabel 5. 16 Parameter Terbaik Random Forest Skenario - 1

Parameter	Nilai / Konfigurasi
n_estimators	100
min_samples_split	5
min_samples_leaf	2
max_features	0.5
max_depth	5
bootstrap	FALSE
class_weight	{0: 0.5627, 1: 4.4855}

Selanjutnya, melakukan Training Performa Cross-Validation sesuai dengan parameter terpilih.

Tabel 5. 17 Hasil Evaluasi Cross-Validation Random Forest

Metrik	Nilai Rata-rata	Standar Deviasi	Interpretasi
Accuracy	0.9806	±0.0281	Model memiliki ketepatan prediksi sangat tinggi dan konsisten antar fold.
Precision	0.9820	±0.0284	Proporsi prediksi positif yang benar sangat tinggi.
Recall	0.9806	±0.0281	Model mampu menangkap sebagian besar kasus positif (gizi kurang).
F1-Score	0.9808	±0.0286	Keseimbangan antara presisi dan recall sangat baik.

Model Random Forest dengan parameter optimal mampu mencapai performa sangat tinggi (akurasi 98.06%) dengan stabilitas yang baik dan sensitivitas yang seimbang terhadap kedua kelas status gizi. Pendekatan class weighting berhasil mengatasi ketimpangan distribusi data tanpa menurunkan akurasi model secara signifikan.

Tabel 5. 18 Tabel Hasil Klasifikasi Random Forest per Kecamatan

Kecamatan	Total	Gizi Baik	Gizi Kurang	Prediksi RF	
				Gizi Kurang RF	Gizi Baik RF
AMBUNTEN	40	36	4	4	36
ARJASA	15	15	0	0	15
BATANG-BATA	30	30	0	0	30
BATUAN	54	54	0	1	53
BATUPUTIH	50	35	15	15	35
BLUTO	40	39	1	1	39
DASUK	55	44	11	11	44
DUNGKEK	70	61	9	9	61
GANDING	70	57	13	11	59
GAPURA	40	38	2	2	38
GAYAM	15	11	4	4	11
GILIGINTING	50	45	5	4	46
GULUK-GULUK	49	41	8	20	29
KALIANGET	60	58	2	4	56
KANGAYAN	35	29	6	6	29
KANGEAN	50	46	4	9	41
Kota Sumenep	90	90	0	0	90
LENTENG	68	59	9	9	59
MANDING	50	48	2	2	48
MASALEMBU	40	37	3	3	37
NONGGUNONG	35	26	9	9	26
PASONGSONG	50	47	3	3	47
PRAGAAN	20	18	2	5	15
RUBARU	45	26	19	19	26
Ra'as	35	31	4	7	28
SAPEKEN	20	17	3	3	17
SARONGGI	45	42	3	3	42
TALANGO	70	68	2	2	68

Hasil tabel menunjukkan bahwa model Random Forest mampu melakukan klasifikasi status gizi anak dengan sangat akurat dan konsisten, mencapai akurasi 97,87%, presisi 97,99%, recall 97,87%, dan F1-score

97,91%. Meski terdapat beberapa kesalahan klasifikasi di beberapa kecamatan, model ini tetap terbukti efektif dan andal dalam mengklasifikasikan status gizi secara keseluruhan.

5.5.3. Skenario 2 (70-30)

Pada skenario kedua, data dibagi menjadi 70% untuk pelatihan dan 30% untuk pengujian untuk mengevaluasi pengaruh peningkatan proporsi data latih terhadap kinerja model Random Forest dalam mengklasifikasi status gizi balita, serta membandingkannya dengan skenario sebelumnya. Setelah dilakukan hyperparameter tuning pada skenario pertama, diperoleh konfigurasi parameter optimal yang disajikan pada tabel berikut.

Tabel 5. 19 Parameter Terbaik Random Forest Skenario - 1

Parameter	Nilai Optimal
n_estimators	300
min_samples_split	5
min_samples_leaf	2
max_features	<i>log2</i>
max_depth	<i>None</i>
class_weight	<i>balanced_subsample</i>
bootstrap	<i>FALSE</i>

Selanjutnya, melakukan Training Performa Cross-Validation sesuai dengan parameter terpilih.

Tabel 5. 20 Hasil Evaluasi Cross-Validation Random Forest Skenario – 2

Metrik Evaluasi	Rata-rata (Mean)	Deviasi Standar ($\pm$)
Accuracy	0.9876	0.0230
Precision	0.9891	0.0199
Recall	0.9876	0.0230
F1-Score	0.9877	0.0228

Nilai metrik yang tinggi dan konsisten menunjukkan bahwa model memiliki kinerja yang sangat baik dan stabil pada berbagai subset data pelatihan.

Tabel 5. 21 Tabel Hasil Klasifikasi Random Forest per Kecamatan

Kecamatan	Total	Gizi Baik	Gizi Kurang	Gizi Baik RF	Gizi Kurang RF
AMBUNTEN	40	36	4	4	4
ARJASA	15	15	0	0	0
BATANG-BATAN	30	30	0	0	0
BATUAN	54	54	0	1	1
BATUPUTIH	50	35	15	15	15
BLUTO	40	39	1	1	1
DASUK	55	44	11	11	11
DUNGKEK	70	61	9	9	9
GANDING	70	57	13	12	12
GAPURA	40	38	2	2	2
GAYAM	15	11	4	4	4
GILIGINTING	50	45	5	5	5
GULUK-GULUK	49	41	8	10	10
KALIANGET	60	58	2	3	3
KANGAYAN	35	29	6	5	5
KANGEAN	50	46	4	5	5
Kota Sumenep	90	90	0	0	0
LENTENG	68	59	9	9	9
MANDING	50	48	2	1	1
MASALEMBU	40	37	3	3	3
NONGGUNONG	35	26	9	9	9
PASONGSONGAN	50	47	3	3	3
PRAGAAN	20	18	2	3	3
RUBARU	45	26	19	19	19
Ra'as	35	31	4	6	6
SAPEKEN	20	17	3	3	3
SARONGGI	45	42	3	4	4
TALANGO	70	68	2	2	2

Hasil tabel menunjukkan bahwa pada skenario ini, model Random Forest mencapai performa paling optimal dibanding skenario lainnya, dengan akurasi 98,71%, presisi 98,7%, recall 98,71%, dan F1-score 98,7%. Hasil ini menegaskan bahwa model bekerja efektif dan andal dalam mengklasifikasi

status gizi anak di berbagai kecamatan, meskipun masih terdapat sedikit kesalahan pada beberapa wilayah.

5.5.4. Perbandingan Hasil Kinerja

Evaluasi performa model ANN dan Random Forest dilakukan melalui tiga skenario pembagian data—80:20, 60:40, dan 70:30 antara data pelatihan dan pengujian—untuk menilai pengaruh variasi proporsi data terhadap kemampuan model dalam mengklasifikasi status gizi anak.

Tabel 5. 22 Perbandingan Hasil Kinerja ANN dan Random Forest

% AKURASI MODEL/ SKENARIO	ANN	RANDOM FOREST
SKENARIO AWAL (80-20)	95.8 %	97.7 %
SKENARIO - 1 (60-40)	94.58 %	97.87 %
SKENARIO - 2 (70-30)	95.88 %	98.71 %

Secara keseluruhan, model Random Forest menunjukkan performa lebih tinggi dan stabil dibandingkan ANN di semua skenario, dengan hasil terbaik dicapai pada skenario 2 (70:30) sebesar 98,71%.

BAB VI

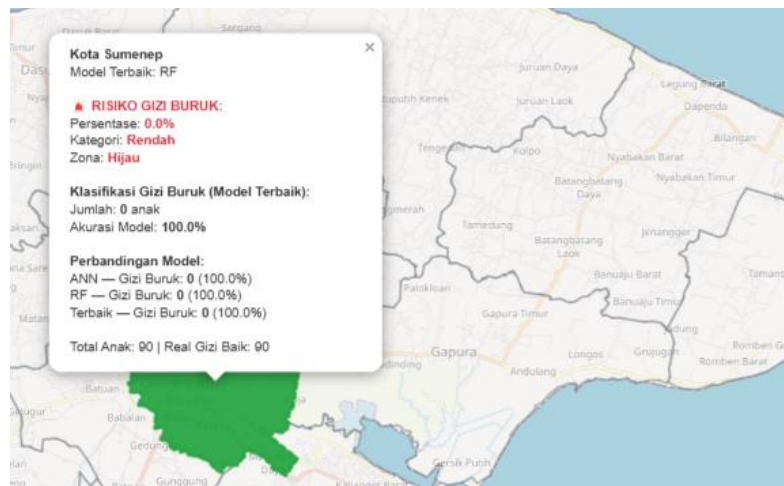
ANALISIS DATA SPASIAL

6.1. Analisis Status Gizi Per-Kecamatan

Analisis ini membandingkan klasifikasi status gizi anak oleh model ANN dan Random Forest terhadap data aktual di setiap kecamatan Kabupaten Sumenep, dengan tujuan menilai akurasi, pola sebaran risiko gizi, dan kemampuan model dalam merepresentasikan kondisi gizi nyata.

6.1.1. Kecamatan Kota Sumenep

Berdasarkan hasil pemetaan, Kecamatan Kota Sumenep menunjukkan zona hijau dengan persentase risiko gizi kurang sebesar 0% yang dikategorikan dalam tingkat risiko rendah.



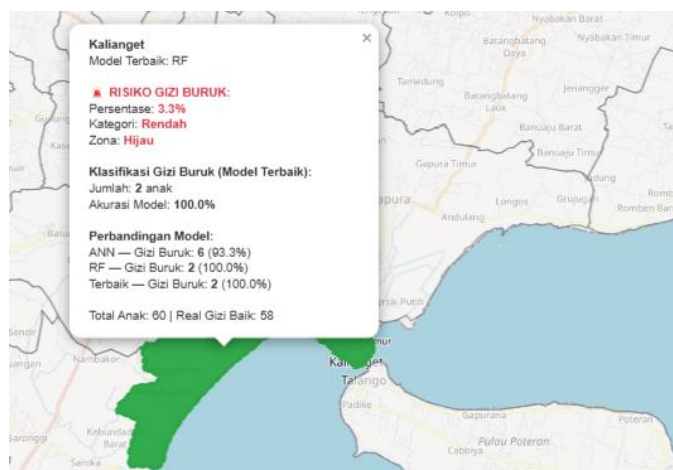
Gambar 6. 1 Peta hasil klasifikasi status gizi kec Sumenep

Hasil klasifikasi ANN dan Random Forest di Kecamatan Kota Sumenep menunjukkan akurasi 100%, tanpa adanya kasus gizi kurang, sesuai dengan data aktual yang mencatat 90 anak tergolong gizi baik. Kinerja optimal ini menandakan pola gizi anak di wilayah tersebut relatif homogen dan mudah dipelajari oleh model,

sementara kondisi sosial ekonomi dan akses fasilitas kesehatan yang lebih baik turut mendukung akurasi tinggi.

6.1.2. Kecamatan Kalianget

Hasil klasifikasi di Kecamatan Kalianget menunjukkan risiko gizi kurang rendah, yakni 3,3% atau sekitar 2 dari 60 anak. Model Random Forest dipilih sebagai yang terbaik karena memberikan akurasi 100%, sesuai dengan kondisi aktual yang mayoritas berstatus gizi baik.



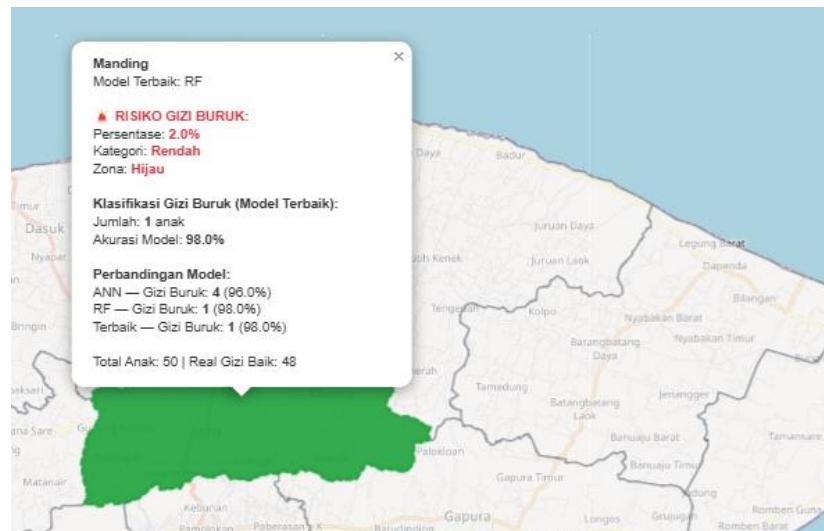
Gambar 6. 2 Peta hasil klasifikasi status gizi kec. Kalianget

Perbandingan dengan model ANN menunjukkan ANN cenderung overestimasi dengan mengklasifikasikan 6 anak berisiko gizi kurang, sementara Random Forest mampu mencocokkan klasifikasi dengan distribusi data sebenarnya. Hal ini menegaskan keunggulan RF dalam menangani ketidakseimbangan kelas dan memberikan hasil lebih stabil pada data dengan risiko gizi rendah.

6.1.3. Kecamatan Manding

Berdasarkan hasil klasifikasi, Kecamatan Manding memiliki risiko gizi kurang 2,0% (kategori rendah, zona hijau). Model terbaik, Random Forest (RF),

mencapai akurasi 98,0%, dengan hanya 1 anak dari 50 diklasifikasi berisiko gizi kurang, sesuai dengan data riil yang menunjukkan 48 anak bergizi baik.



Gambar 6. 3 Peta hasil klasifikasi status gizi kec. Manding

Dibandingkan dengan metode ANN, yang memklasifikasi 4 anak berisiko gizi kurang dengan akurasi 96.0%, model RF memberikan hasil yang lebih akurat dan stabil.

6.1.4. Kecamatan Talango

Berdasarkan hasil klasifikasi, Kecamatan Talango memiliki risiko gizi kurang 2,9% (kategori rendah, zona hijau). Model terbaik, Random Forest (RF), mencapai akurasi 100%, dengan 2 dari 70 anak diklasifikasi berisiko gizi kurang dan 68 anak tergolong bergizi baik, sesuai data riil.



Gambar 6. 4 Peta hasil klasifikasi status gizi kec. Talango

Baik model ANN maupun RF memberikan hasil klasifikasi yang identik dengan tingkat akurasi sempurna. Namun, RF dinilai lebih optimal karena memiliki stabilitas yang lebih tinggi dalam mengatasi variasi data dan noise.

6.1.5. Kecamatan Bluto

Dari total 40 anak, data real menunjukkan 39 anak berstatus gizi baik dan 1 anak gizi kurang, menunjukkan distribusi yang sangat imbalanced.

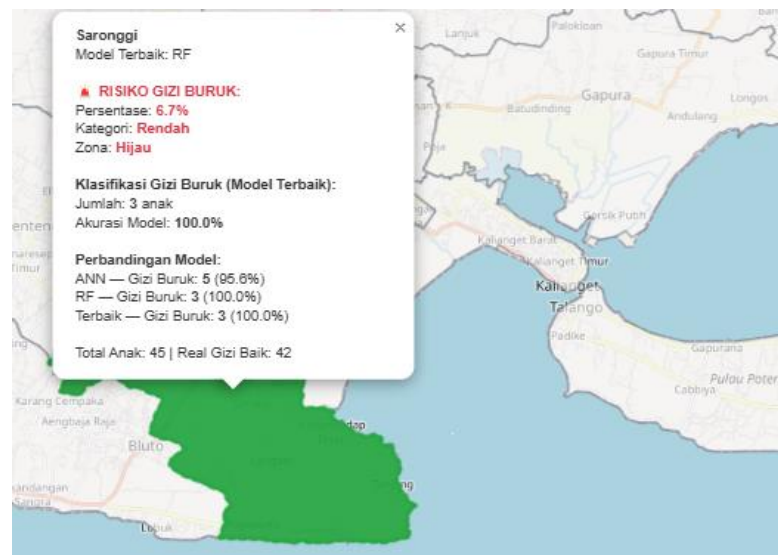


Gambar 6. 5 Peta hasil klasifikasi status gizi kec. Bluto

Klasifikasi menggunakan ANN dan Random Forest (RF) sama-sama berhasil mengidentifikasi 1 anak gizi kurang, sehingga menghasilkan akurasi 100%.

6.1.6. Kecamatan Saronggi

Dari total 45 anak di Saronggi, data real menunjukkan 42 anak berstatus gizi baik dan 3 anak mengalami gizi kurang, sehingga distribusi data tergolong imbalanced dengan kasus gizi kurang sekitar 6,7%.

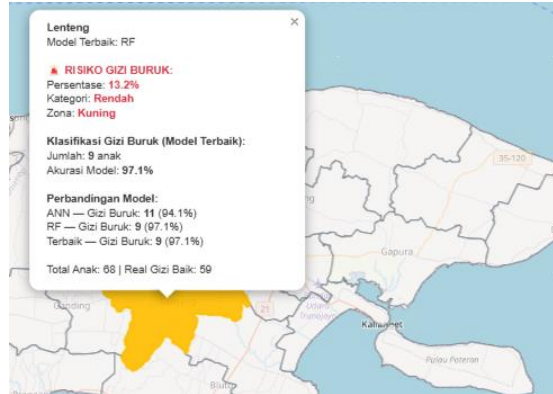


Gambar 6. 6 Peta hasil klasifikasi status gizi kec. Saronggi

Klasifikasi menggunakan ANN menunjukkan 5 anak gizi kurang, sedangkan Random Forest (RF) memklasifikasi 3 anak gizi kurang sesuai data real. Akurasi klasifikasi RF mencapai 100%, sedangkan ANN sedikit lebih rendah yaitu 95,6%.

6.1.7. Kecamatan Lenteng

Di Kecamatan Lenteng, dari 68 anak, 59 anak bergizi baik dan 9 anak mengalami gizi kurang (13,2%). Model ANN mengklasifikasi 11 anak gizi kurang, sementara Random Forest (RF) menghasilkan klasifikasi yang sesuai data riil, yakni 9 anak gizi kurang.

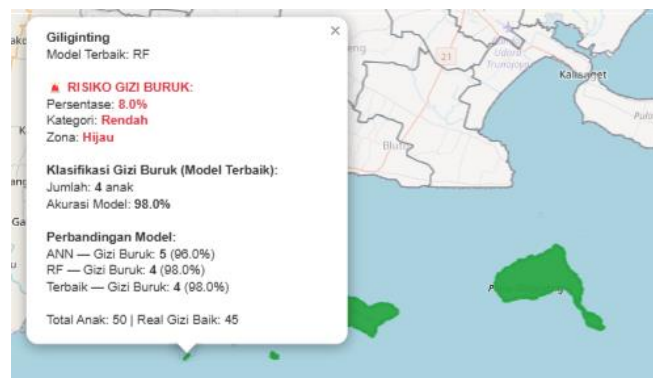


Gambar 6. 7 Peta hasil klasifikasi status gizi kec. Lenteng

Akurasi RF mencapai 97,1%, sedikit lebih tinggi dibanding ANN yang memiliki akurasi 94,1%.

6.1.8. Kecamatan Giligenting

Dari total 50 anak di Giligenting, data real menunjukkan 45 anak berstatus gizi baik dan 5 anak gizi buruk, sehingga kasus gizi buruk sekitar 8,0%.

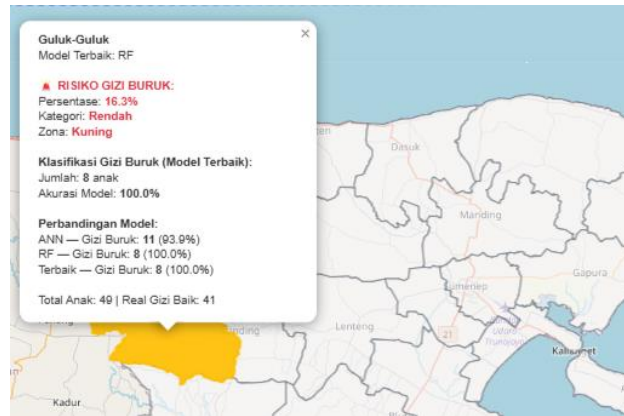


Gambar 6. 8 Peta hasil klasifikasi status gizi kec. Giligenting

Klasifikasi menggunakan ANN menunjukkan 5 anak gizi buruk, sedangkan Random Forest (RF) memklasifikasi 4 anak gizi buruk, menghasilkan akurasi 98,0% untuk RF dan 96,0% untuk ANN.

6.1.9. Kecamatan Guluk-guluk

Dari total 49 anak di Guluk-Guluk, data real menunjukkan 41 anak berstatus gizi baik dan 8 anak mengalami gizi buruk, sehingga kasus gizi buruk mencapai 16,3% dan tergolong cukup signifikan.

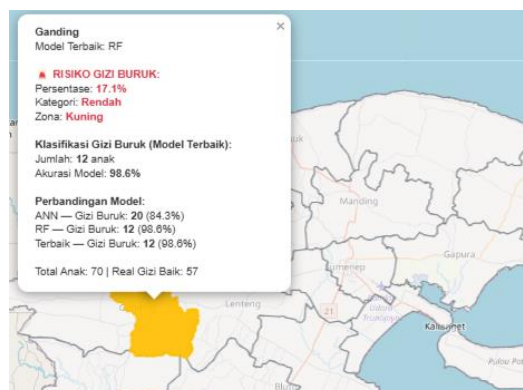


Gambar 6. 9 Peta hasil klasifikasi status gizi kec. Guluk-guluk

Klasifikasi menggunakan ANN menunjukkan 11 anak gizi buruk, sementara Random Forest (RF) memklasifikasi 8 anak gizi buruk sesuai data real, menghasilkan akurasi 100% untuk RF dan 93,9% untuk ANN.

6.1.10. Kecamatan Ganding

Dari total 70 anak di Ganding, data real menunjukkan 57 anak berstatus gizi baik dan 13 anak mengalami gizi buruk, sehingga kasus gizi buruk mencapai 17,1%.



Gambar 6. 10 Peta hasil klasifikasi status gizi kec. Ganding

Klasifikasi menggunakan ANN menunjukkan 20 anak gizi buruk, sementara Random Forest (RF) memklasifikasi 12 anak gizi buruk, menghasilkan akurasi 98,6% untuk RF dan 84,3% untuk ANN.

6.1.11. Kecamatan Pragaan

Dari total 20 anak di wilayah Pragaan, data aktual menunjukkan bahwa 18 anak berada dalam kategori gizi baik dan 2 anak mengalami gizi buruk, dengan proporsi gizi buruk sebesar 15,0%.

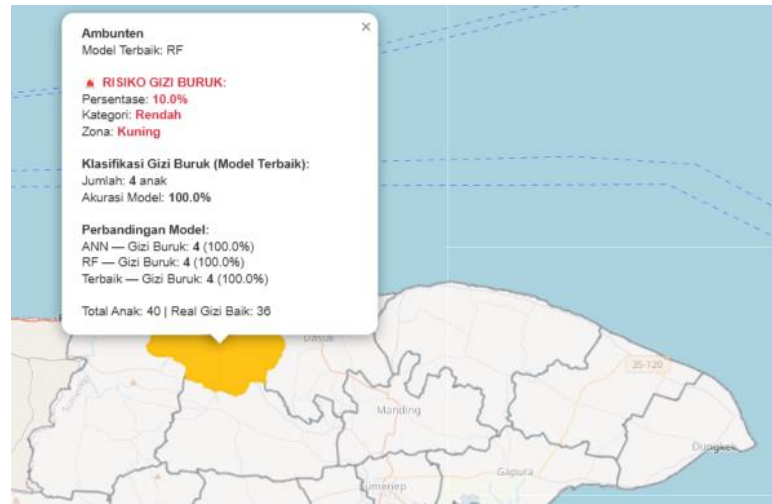


Gambar 6. 11 Peta hasil klasifikasi status gizi kec. Pragaan

Hasil klasifikasi menunjukkan bahwa metode ANN mengidentifikasi 4 anak gizi buruk, sementara Random Forest (RF) memklasifikasi 3 anak gizi buruk. Model RF memperoleh tingkat akurasi 95,0%, lebih tinggi dibanding ANN yang mencapai 90,0%.

6.1.12. Kecamatan Ambunten

Dari total 40 anak di wilayah Ambunten, data aktual menunjukkan 36 anak berstatus gizi baik dan 4 anak mengalami gizi buruk, dengan persentase kasus gizi buruk sebesar 10,0%.



Gambar 6. 12 Peta hasil klasifikasi status gizi kec. Ambunten

Hasil klasifikasi menggunakan metode ANN dan Random Forest (RF) menunjukkan performa yang sama, yaitu mampu mengidentifikasi 4 anak gizi buruk dengan tingkat akurasi 100%.

6.1.13. Kecamatan Pasongsongan

Dari total 50 anak di wilayah Pasongsongan, data aktual menunjukkan 47 anak berstatus gizi baik dan 3 anak mengalami gizi buruk, dengan persentase kasus gizi buruk sebesar 6,0%.

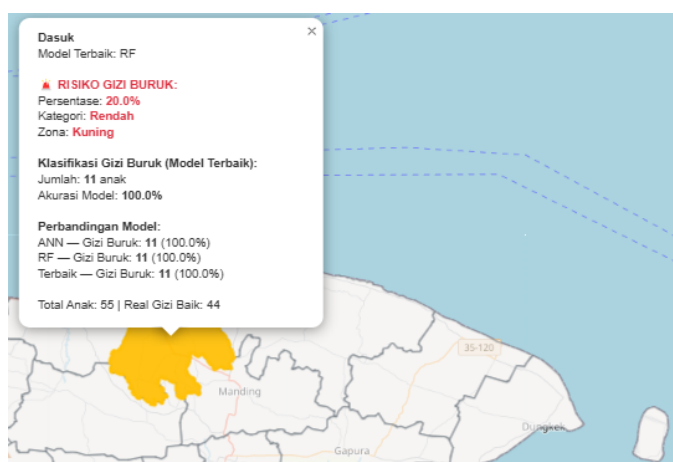


Gambar 6. 13 Peta hasil klasifikasi status gizi kec. Pasongsongan

Hasil klasifikasi menunjukkan bahwa metode ANN memperkirakan 5 anak gizi buruk dengan akurasi 96,0%, sedangkan Random Forest (RF) berhasil memklasifikasi 3 anak gizi buruk sesuai dengan data real dan mencapai akurasi sempurna sebesar 100%.

6.1.14 Kecamatan Dasuk

Dari total 55 anak di wilayah Dasuk, data aktual menunjukkan 44 anak berstatus gizi baik dan 11 anak mengalami gizi buruk, dengan persentase kasus gizi buruk sebesar 20,0%.

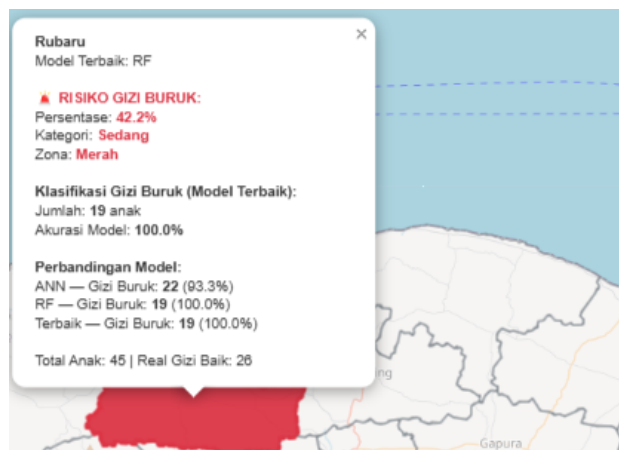


Gambar 6. 14 Peta hasil klasifikasi status gizi kec. Dasuk

Hasil klasifikasi menunjukkan bahwa baik metode Artificial Neural Network (ANN) maupun Random Forest (RF) sama-sama mampu mengidentifikasi 11 anak gizi buruk secara tepat, dengan tingkat akurasi mencapai 100%.

6.1.15 Kecamatan Rubaru

Dari hasil analisis di wilayah Rubaru, diperoleh total anak dengan risiko gizi buruk sebesar 42,2%, menunjukkan kategori **sedang** dengan zona berwarna merah.

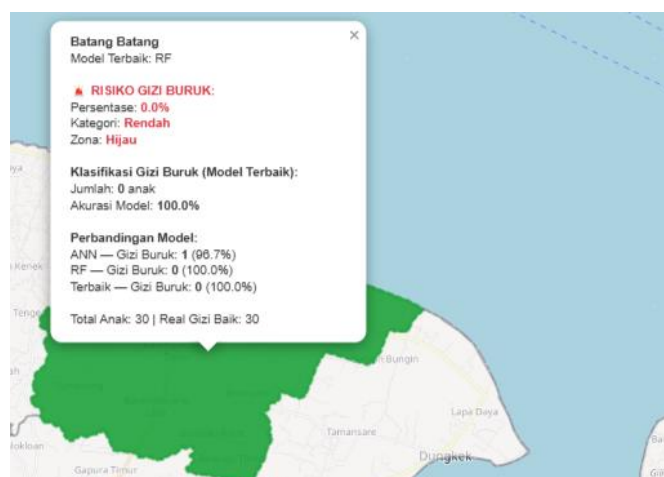


Gambar 6. 15 Peta hasil klasifikasi status gizi kec. Rubaru

Berdasarkan hasil klasifikasi, metode Artificial Neural Network (ANN) memklasifikasi 22 anak gizi buruk dengan tingkat akurasi 93,3%, sedangkan Random Forest (RF) berhasil memklasifikasi 19 anak gizi buruk, sesuai dengan data real, dengan akurasi sempurna sebesar 100%.

6.1.16 Kecamatan Batang-batang

Hasil analisis di wilayah Batang Batang menunjukkan bahwa seluruh anak tergolong dalam kategori gizi baik, dengan risiko gizi buruk sebesar 0,0%, yang termasuk dalam kategori rendah dan ditandai dengan zona berwarna hijau.

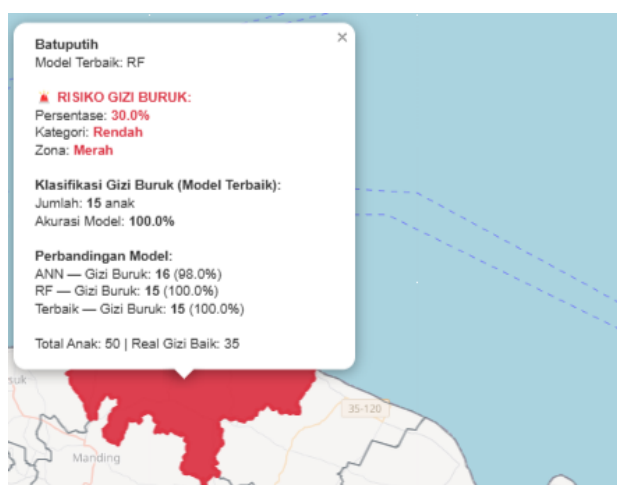


Gambar 6. 16 Peta hasil klasifikasi status gizi kec. Batang-batang

Model Random Forest (RF) memberikan hasil klasifikasi yang sepenuhnya akurat dengan akurasi 100,0%, di mana tidak ada anak yang teridentifikasi mengalami gizi buruk, sesuai dengan kondisi data aktual.

6.1.17 Kecamatan Batuputih

Hasil analisis di wilayah Batuputih menunjukkan bahwa tingkat risiko gizi buruk mencapai 30,0%, yang tergolong kategori rendah, namun berada dalam zona merah karena proporsi anak dengan gizi kurang cukup signifikan.



Gambar 6. 17 Peta hasil klasifikasi status gizi kec. Batuputih

Model Random Forest (RF) memberikan hasil klasifikasi yang paling akurat dengan akurasi 100,0%, di mana sebanyak 15 anak teridentifikasi berisiko gizi buruk, sesuai dengan kondisi aktual di lapangan.

Sementara itu, model Artificial Neural Network (ANN) memperkirakan 16 anak sebagai gizi buruk dengan akurasi sedikit lebih rendah, yaitu 98,0%.

6.1.14. Kecamatan Dungkek

Hasil analisis di wilayah Dungkek menunjukkan tingkat risiko gizi buruk sebesar 12,9%, yang tergolong dalam kategori rendah dengan status zona kuning.



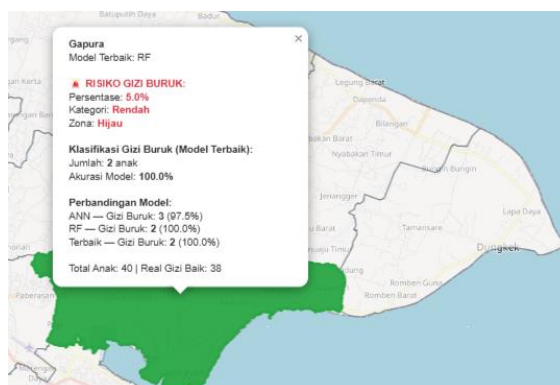
Gambar 6. 18 Peta hasil klasifikasi status gizi kec. Dungkek

Model Random Forest (RF) memberikan hasil terbaik dengan akurasi 100,0%, berhasil memklasifikasi 9 anak berisiko gizi buruk secara tepat sesuai dengan kondisi aktual.

Sebaliknya, model Artificial Neural Network (ANN) memklasifikasi 11 anak sebagai gizi buruk dengan tingkat akurasi 97,1%.

6.1.15. Kecamatan Gapura

Hasil analisis di wilayah Gapura menunjukkan tingkat risiko gizi buruk sebesar 5,0%, yang termasuk dalam kategori rendah dengan status zona hijau.



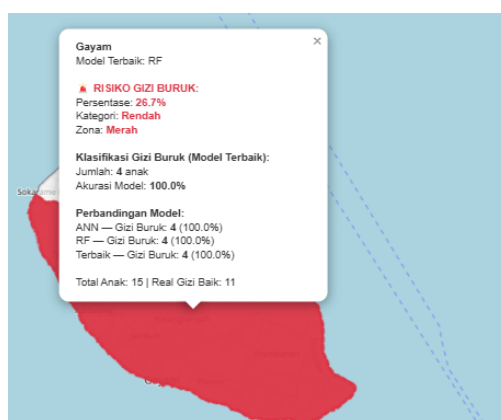
Gambar 6. 19 Peta hasil klasifikasi status gizi kec. Gapura

Model Random Forest (RF) tampil sebagai model terbaik dengan akurasi 100,0%, berhasil memklasifikasi secara tepat 2 anak yang berisiko gizi buruk sesuai dengan data aktual.

Sementara itu, model Artificial Neural Network (ANN) memklasifikasi 3 anak dengan tingkat akurasi 97,5%.

6.1.16. Kecamatan Gayam

Hasil analisis di wilayah Gayam menunjukkan tingkat risiko gizi buruk sebesar 26,7%, yang masih berada dalam kategori rendah, namun secara spasial termasuk ke dalam zona merah karena proporsi anak berisiko relatif tinggi dibanding wilayah lain.



Gambar 6. 20 Peta hasil klasifikasi status gizi kec. Gayam

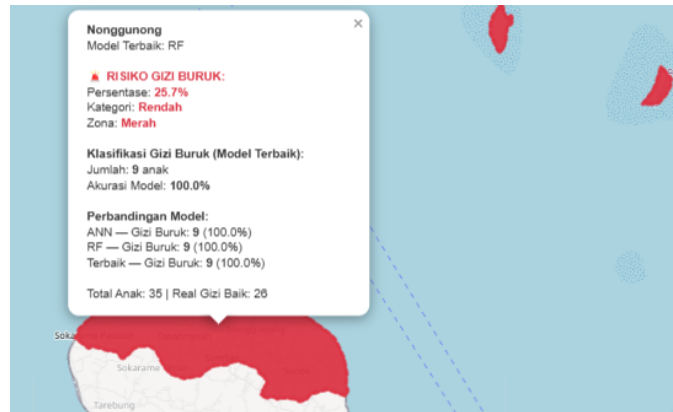
Model Random Forest (RF) menjadi model terbaik dengan akurasi sempurna 100,0%, berhasil memklasifikasi 4 anak yang berisiko gizi buruk secara tepat sesuai dengan data aktual.

Model Artificial Neural Network (ANN) juga menunjukkan hasil identik dengan akurasi 100 %.

6.1.17. Kecamatan Nonggunong

Hasil analisis di wilayah Nonggunong menunjukkan tingkat risiko gizi buruk sebesar 25,7%, yang tergolong kategori rendah, namun secara spasial termasuk

dalam zona merah karena proporsi anak berisiko relatif tinggi dibandingkan wilayah lainnya.



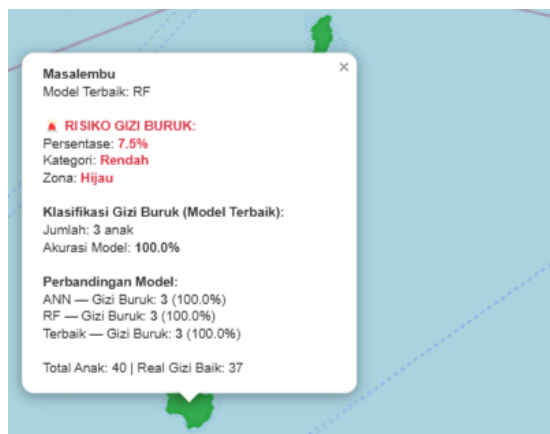
Gambar 6. 21 Peta hasil klasifikasi status gizi kec. Nonggunong

Model Random Forest (RF) menjadi model terbaik dengan akurasi 100,0%, berhasil memklasifikasi secara tepat 9 anak yang tergolong gizi buruk sesuai dengan kondisi aktual di lapangan.

Model Artificial Neural Network (ANN) menghasilkan klasifikasi yang sama dengan tingkat akurasi 100,0%.

6.1.18. Kecamatan Masalembu

Hasil analisis di wilayah Masalembu menunjukkan tingkat risiko gizi buruk sebesar 7,5%, yang termasuk dalam kategori rendah dan secara spasial berada pada zona hijau.



Gambar 6. 22 Peta hasil klasifikasi status gizi kec. Masalembu

Model Random Forest (RF) menjadi model terbaik dengan akurasi sempurna 100,0%, berhasil memklasifikasi secara tepat 3 anak yang tergolong gizi buruk sesuai dengan data aktual.

Model Artificial Neural Network (ANN) juga memberikan hasil yang identik dengan akurasi 100,0%.

6.1.19. Kecamatan Arjasa

Hasil analisis di wilayah Arjasa menunjukkan tidak adanya anak dengan risiko gizi buruk (0,0%), yang menempatkan wilayah ini dalam kategori rendah dengan status zona hijau.



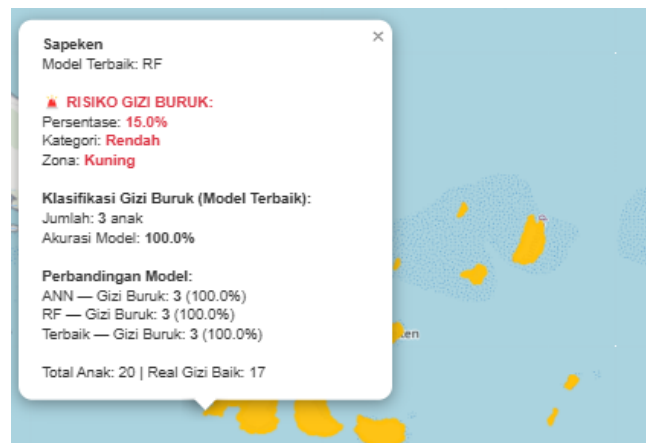
Gambar 6. 23 Peta hasil klasifikasi status gizi kec. Arjasa

Model Random Forest (RF) menjadi model terbaik dengan akurasi sempurna 100,0%, berhasil memklasifikasi bahwa seluruh anak berada dalam kondisi gizi baik, sesuai dengan data aktual.

Sementara itu, model Artificial Neural Network (ANN) memklasifikasi 1 anak sebagai gizi buruk dengan akurasi 93,3%.

6.1.20. Kecamatan Sapeken

Hasil analisis di wilayah Sapeken menunjukkan tingkat risiko gizi buruk sebesar 15,0%, yang tergolong dalam kategori rendah dengan status zona kuning.



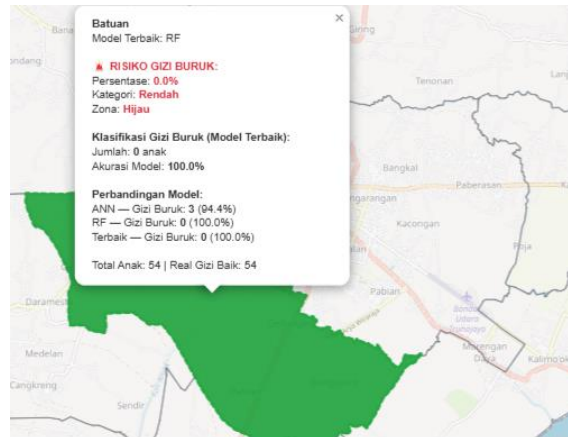
Gambar 6. 24 Peta hasil klasifikasi status gizi kec. Sapeken

Model Random Forest (RF) menjadi model terbaik dengan akurasi sempurna 100,0%, berhasil memklasifikasi secara tepat 3 anak yang berisiko gizi buruk sesuai dengan data aktual.

Model Artificial Neural Network (ANN) juga memberikan hasil yang identik dengan akurasi 100,0%.

6.1.21. Kecamatan Batuan

Hasil analisis di wilayah Batuan menunjukkan tidak ada anak dengan risiko gizi buruk (0,0%), sehingga wilayah ini termasuk dalam kategori rendah dengan zona hijau.



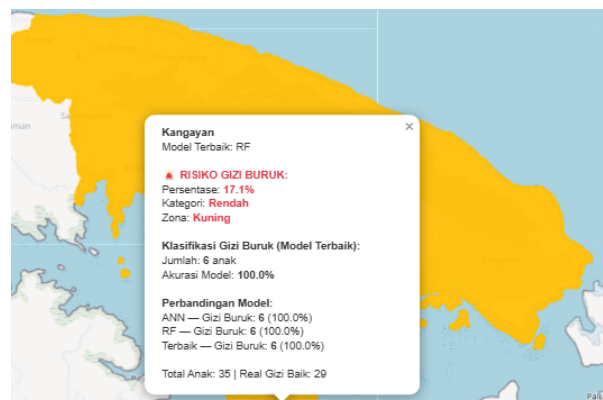
Gambar 6. 25 Peta hasil klasifikasi status gizi kec. Sapeken

Model Random Forest (RF) menjadi model terbaik dengan akurasi 100,0%, berhasil memklasifikasi secara tepat bahwa seluruh anak berada dalam kondisi gizi baik sesuai data aktual.

Sementara itu, model **Artificial Neural Network (ANN)** memklasifikasi **3 anak** sebagai gizi buruk dengan **akurasi 94,4%**

6.1.22. Kecamatan Kangayan

Hasil analisis di wilayah Kangayan menunjukkan tingkat risiko gizi buruk sebesar 17,1%, yang tergolong kategori rendah dengan zona kuning.



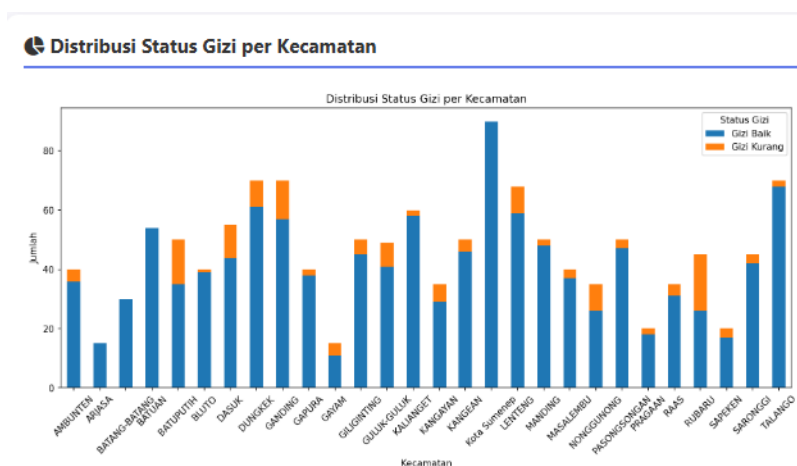
Gambar 6. 26 Peta hasil klasifikasi status gizi kec. Kangayan

Model Random Forest (RF) menjadi model terbaik dengan akurasi 100,0%, berhasil memklasifikasi secara tepat 6 anak yang berisiko gizi buruk sesuai dengan data aktual.

Model Artificial Neural Network (ANN) juga menunjukkan hasil identik dengan akurasi 100,0%, menandakan bahwa pola distribusi gizi pada wilayah ini cukup jelas dan dapat dikenali dengan baik oleh kedua model.

6.2. Distribusi Status Gizi per Kecamatan

Analisis spasial dilakukan untuk memetakan distribusi geografis status gizi anak dan mengidentifikasi pola spasial dari hasil klasifikasi model Random Forest dan ANN. Pendekatan ini bertujuan menilai kesesuaian model dengan kondisi nyata serta mendukung pengambilan keputusan berbasis lokasi.

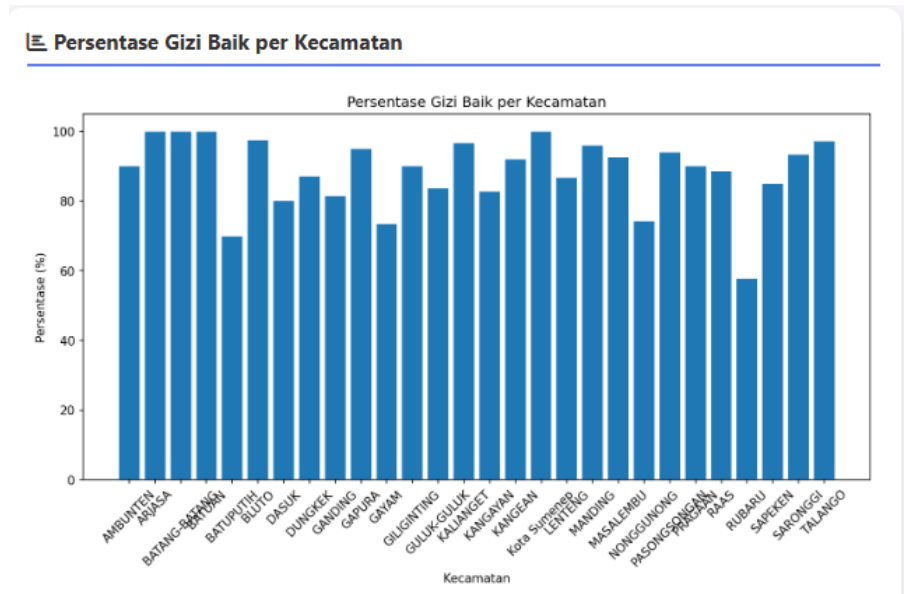


Gambar 6. 27 Distribusi Gizi Per Kecamatan

Visualisasi distribusi status gizi per kecamatan menunjukkan mayoritas anak berada dalam kategori gizi baik, dengan Kota Sumenep, Manding, dan Lenteng memiliki jumlah tertinggi, mencerminkan kondisi kesehatan dan akses gizi yang baik. Sebaliknya, Dasuk, Ganding, dan Batuputih menampilkan proporsi gizi kurang lebih tinggi, mengindikasikan pengaruh faktor spasial seperti fasilitas kesehatan, akses air bersih, dan kondisi sosial ekonomi setempat.

6.3. Persentase Gizi Baik per Kecamatan

Grafik “Persentase Gizi Baik per Kecamatan” menegaskan bahwa sebagian besar kecamatan memiliki tingkat gizi baik di atas 85%, menggambarkan kondisi umum yang positif dan keberhasilan wilayah dalam menjaga status gizi anak di tingkat kabupaten.

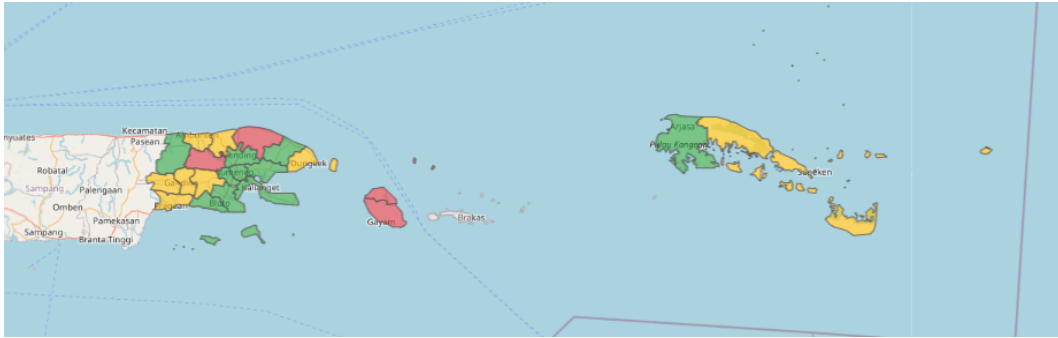


Gambar 6. 28 Persentase Gizi Baik Per Kecamatan

Gambar 6.28 memperlihatkan bahwa kecamatan seperti Dasuk dan Ganding memiliki persentase gizi baik di bawah 80%, menandakan risiko gizi lebih tinggi. Pola spasial ini menunjukkan ketimpangan distribusi gizi yang terkonsentrasi di bagian tengah dan timur wilayah, kemungkinan terkait keterbatasan akses layanan kesehatan..

6.4. Pemetaan Zona Status Gizi Kabupaten Sumenep

Visualisasi peta zonasi pada Gambar 6.29 menunjukkan hasil integrasi antara data spasial dan hasil klasifikasi model machine learning (Random Forest).



Gambar 6. 29 Peta Resiko Gizi Kurang (Model Terbaik)

Setiap kecamatan diklasifikasikan ke dalam tiga zona berdasarkan persentase anak dengan status gizi baik dan gizi kurang, yang direpresentasikan melalui skema warna:

- a. ■ Hijau → wilayah dengan kategori aman (persentase gizi baik > 90%)
- b. ■ Kuning → wilayah dengan kategori waspada (persentase gizi baik 80–90%)
- c. ■ Merah → wilayah dengan kategori rawan gizi kurang (persentase gizi baik < 80%)

Hasil pemetaan menunjukkan bahwa sebagian besar wilayah Kabupaten Sumenep termasuk kategori aman (zona hijau), terutama di barat dan selatan seperti Kota Sumenep, Manding, dan Lenteng, yang memiliki tingkat gizi baik tertinggi. Kondisi ini konsisten dengan analisis spasial sebelumnya, mencerminkan akses kesehatan dan pendidikan yang lebih baik di wilayah tersebut.

Sebaliknya, wilayah tengah dan timur seperti Dasuk, Ganding, Batuputih, serta sebagian Dungkek masuk kategori kuning hingga merah, menandakan risiko gizi kurang yang lebih tinggi. Kecamatan Gayam di kepulauan termasuk zona merah, mencerminkan keterbatasan fasilitas kesehatan dan ketimpangan geografis. Zona kuning (waspada) terdapat di kecamatan transisi seperti Kalianget dan Arjasa, yang meskipun sebagian besar anak bergizi baik, masih memerlukan perhatian

pencegahan. Pola spasial ini memperlihatkan gradien kondisi gizi dari barat (lebih baik) ke timur (lebih rawan), menegaskan pengaruh signifikan faktor spasial terhadap status gizi masyarakat.

Dengan demikian, hasil pemetaan ini menjadi dasar perencanaan intervensi gizi berbasis wilayah, sehingga program peningkatan gizi dapat difokuskan pada zona merah dan kuning untuk memperbaiki pemerataan status kesehatan anak.

BAB VII

PENUTUP

7.1. Kesimpulan

Berdasarkan hasil penelitian yang telah dilakukan dapat disimpulkan bahwa:

- a. Berdasarkan hasil penelitian, dapat disimpulkan bahwa model Artificial Neural Network (ANN) dan Random Forest (RF) mampu mengklasifikasikan status gizi anak dengan baik. Model ANN menunjukkan akurasi 95,8%, presisi 96,4%, recall 95,8%, dan F1-score 95,9%, sedangkan model RF memiliki performa lebih tinggi dengan akurasi 98,7%, presisi 98,7%, recall 98,7%, dan F1-score 98,7%. Hasil ini menunjukkan bahwa RF lebih stabil dan andal dalam menangani variasi data serta ketidakseimbangan kelas, sehingga memberikan klasifikasi yang lebih tepat dan konsisten dibanding ANN.
- b. Analisis spasial menunjukkan distribusi status gizi yang tidak merata di Kabupaten Sumenep. Sebagian besar wilayah, terutama Kota Sumenep, Manding, dan Lenteng, berada dalam kategori aman (zona hijau) dengan proporsi anak bergizi baik yang dominan. Sebaliknya, kecamatan seperti Dasuk, Ganding, Batuputih, dan Gayam berada dalam zona kuning hingga merah, menunjukkan risiko gizi kurang yang lebih tinggi. Pola ini menegaskan adanya gradien geografis dari barat ke timur, di mana faktor spasial seperti akses layanan kesehatan, kondisi sosial ekonomi, dan fasilitas pendidikan berperan penting dalam menentukan status gizi anak.

7.2. Saran

Penelitian ini memberikan pandangan bagi penulis untuk memberi saran seperti berikut:

- a. Penelitian selanjutnya disarankan untuk menambah variabel spasial dan sosial-ekonomi (misalnya tingkat pendapatan, pola konsumsi, dan kepadatan fasilitas kesehatan) agar model lebih komprehensif.
- b. Diperlukan perluasan dataset dan uji validasi lintas wilayah agar model dapat diadaptasi secara nasional.
- c. Disarankan untuk mengembangkan sistem informasi geografis (SIG) berbasis android agar hasil klasifikasi dapat digunakan langsung oleh instansi kesehatan daerah.

DAFTAR PUSTAKA

- Ankalaki, S., G Biradar, V., Naik P, K. K., & S. Hukkeri, G. (2024). A ANN Approach for Malnutrition Detection. *International Journal of Online and Biomedical Engineering (iJOE)*, 20(06), 116–138. <https://doi.org/10.3991/ijoe.v20i06.46919>
- Bitew, F. H., Sparks, C. S., & Nyarko, S. H. (2021). Machine learning algorithms for predicting undernutrition among under-five children in Ethiopia. *Public Health Nutrition*, 1–12. <https://doi.org/10.1017/S1368980021004262>
- BMC Public Health. (2023). Spatial distribution and determinants of stunting, wasting and underweight in children under-five in Ethiopia. *BMC Public Health*, 23(15488), 1–16. <https://doi.org/10.1186/s12889-023-15488-z>
- BMC Nutrition. (2022). Mapping the prevalence of severe acute malnutrition in Papua, Indonesia by using geostatistical models. *BMC Nutrition*, 8(10), 1–14. <https://doi.org/10.1186/s40795-022-00504-z>
- Browne, C., Matteson, D. S., McBride, L., Hu, L., Liu, Y., Sun, Y., Wen, J., & Barrett, C. B. (2021). Multivariate *Random Forest* prediction of poverty and malnutrition prevalence. *PLOS ONE*, 16(9), e0255519. <https://doi.org/10.1371/journal.pone.0255519>
- Chilyabanyama, O. N., Chilengi, R., Simuyandi, M., Chisenga, C. C., Chirwa, M., Hamusonde, K., Saroj, R. K., Iqbal, N. T., Ngaruye, I., & Bosomprah, S. (2022). Performance of Machine Learning Classifiers in Classifying Stunting among Under-Five Children in Zambia. *Children*, 9(7), 1082. <https://doi.org/10.3390/children9071082>
- De Onis, M., Borghi, E., Arimond, M., Webb, P., Croft, T., Saha, K., De-Regil, L. M., Thuita, F., Heidkamp, R., Krusevec, J., Hayashi, C., & Flores-Ayala, R. (2019). Prevalence thresholds for wasting, overweight and stunting in children under 5 years. *Public Health Nutrition*, 22(1), 175–179. <https://doi.org/10.1017/S1368980018002434>
- Dwyer-Lindgren, L., Stubbs, R. W., Bertozzi-Villa, A., Morozoff, C., Kutz, M. J., Huynh, C., Barber, R. M., Shackelford, K. A., Mackenbach, J. P., van Lenthe, F. J., Flaxman, A. D., Naghavi, M., Mokdad, A. H., Murray, C. J. L., & Hay, S. I. (2020). Mapping child growth failure across low- and middle-income countries. *Nature*, 577(7789), 231–234. <https://doi.org/10.1038/s41586-019-1878-8>
- Fenta, H. M., Zewotir, T., & Muluneh, E. K. (2021). A machine learning classifier approach for identifying the determinants of under-five child undernutrition in Ethiopian administrative zones. *BMC Medical Informatics and Decision Making*, 21(1), 291. <https://doi.org/10.1186/s12911-021-01652-1>

- Ghwanmeh, S., Mohammad, A., & Al-Ibrahim, A. (2013). Innovative Artificial Neural Networks-Based Decision Support System for Heart Diseases Diagnosis. *Journal of Intelligent Learning Systems and Applications*, 05(03), 176–183. <https://doi.org/10.4236/jilsa.2013.53019>
- Göl, M., Aktürk, C., Talan, T., Vural, M. S., & Türkbeyler, İ. H. (2025). Predicting malnutrition-based anemia in geriatric patients using machine learning methods. *Journal of Evaluation in Clinical Practice*, 31(2), e14142. <https://doi.org/10.1111/jep.14142>
- Halimah, & Suntin. (2020). Proyeksi dan pemetaan wilayah sebaran balita stunting di Kota Makassar berbasis SIG. *Jurnal Aplikasi Geo-Informatika*, 1(2), 65–72. <https://doi.org/10.30871/jagi.v1i2.6936>
- Islam, Md. M., Rahman, Md. J., Islam, M. M., Roy, D. C., Ahmed, N. A. M. F., Hussain, S., Amanullah, M., Abedin, Md. M., & Maniruzzaman, Md. (2022). Application of machine learning based algorithm for prediction of malnutrition among women in Bangladesh. *International Journal of Cognitive Computing in Engineering*, 3, 46–57. <https://doi.org/10.1016/j.ijcce.2022.02.002>
- Jain, S., Khanam, T., Abedi, A. J., & Khan, A. A. (2022). Efficient Machine Learning for Malnutrition Prediction among under-five children in India. *2022 IEEE Delhi Section Conference (DELCON)*, 1–10. <https://doi.org/10.1109/DELCON54057.2022.9753080>
- Khan, M. N. A., & Yunus, R. M. (2023). A hybrid ensemble approach to accelerate the classification accuracy for predicting malnutrition among under-five children in sub-Saharan African countries. *Nutrition*, 108, 111947. <https://doi.org/10.1016/j.nut.2022.111947>
- Khan, M. U., Saeed, Z., & Raza, A. (2022). *ANN-based Decision Support System for Classification of COVID-19 and Pneumonia Patients*. 6(1).
- Kishore, K. K., Suman, J. V., Mniyamba, I. L., Polamuri, S. R., & Venkatesh, B. (2023). *Prediction of malnutrition in newbornInfants using machine learning techniques*. <https://doi.org/10.21203/rs.3.rs-2958834/v1>
- Kurniawan, F., Aziza, M. R., Hasanah, N. A., Junikhah, A., Alam, L. S., Wibawa, A. P., & Hammad, J. (2024). The Innovative Smart Green Campus as Life-Based Learning Characteristics of Future Learning Efforts to Complete the SDG's. *Journal of Lifestyle and SDGs Review*, 5(2), e02908. <https://doi.org/10.47172/2965-730x.sdgsreview.v5.n02.pe02908>
- Lukita Sari, A. D., Ardiansyah, M., & Handayani, D. (2022). GIS mapping of toddler stunting cases in Bantul Regency. *Procedia of Engineering and Life Science*, 2(3), 200–206. <https://doi.org/10.21070/pels.v2i3.2212>

- Mantero, A., & Ishwaran, H. (2021). Unsupervised random forests. *Statistical Analysis and Data Mining: The ASA Data Science Journal*, 14(2), 144–167. <https://doi.org/10.1002/sam.11498>
- Momand, Z., Zarinkhail, M. S., & Aryan, M. F. (2022). Machine Learning Based Prediction of Edematous Malnutrition in Afghan Children. In M. Al-Emran, M. A. Al-Sharafi, M. N. Al-Kabi, & K. Shaalan (Eds.), *Proceedings of International Conference on Emerging Technologies and Intelligent Systems* (Vol. 322, pp. 235–245). Springer International Publishing. https://doi.org/10.1007/978-3-030-85990-9_20
- Mutammimul Ula, Ananda Faridhatul Ulva, Ilham Saputra, Mauliza Mauliza, & Ivan Maulana. (2022). Implementation Of Machine Learning Using The K-Nearest Neighbor Classification Model In Diagnosing Malnutrition In Children. *Multica Science And Technology (Mst) Journal*, 2(1), 94–99. <https://doi.org/10.47002/mst.v2i1.326>
- Nel, S., Feucht, U. D., Nel, A. L., Becker, P. J., & Wenhold, F. A. M. (2022). A novel screening tool to predict severe acute malnutrition through automated monitoring of weight-for-age growth curves. *Maternal & Child Nutrition*, 18(3), e13364. <https://doi.org/10.1111/mcn.13364>
- Prayudhy, S., & Ahyanti, M. (2022). Determining stunting risk areas using a combined AHP-GIS approach: A case study of Pesawaran Regency, Lampung, Indonesia. *Jurnal Kesehatan Lingkungan*, 14(2), 112–121. <https://doi.org/10.31964/jkesling.v14i2.5046>
- Rahman, S. M. J., Ahmed, N. A. M. F., Abedin, Md. M., Ahammed, B., Ali, M., Rahman, Md. J., & Maniruzzaman, Md. (2021). Investigate the risk factors of stunting, wasting, and underweight among under-five Bangladeshi children and its prediction based on machine learning approach. *PLOS ONE*, 16(6), e0253172. <https://doi.org/10.1371/journal.pone.0253172>
- Ramchoun, H., Amine, M., Idrissi, J., Ghanou, Y., & Ettaouil, M. (2016). Multilayer Perceptron: Architecture Optimization and Training. *International Journal of Interactive Multimedia and Artificial Intelligence*, 4(1), 26. <https://doi.org/10.9781/ijimai.2016.415>
- Ridzki, M. M., Hadijah, I., Mukidin, M., Azzahra, A., & Nurjanah, A. (2023). K-Means Algorithm Method for Clustering Best-Selling Product Data at XYZ Grocery Stores. *International Journal of Social Service and Research*, 3(12), 3354–3367. <https://doi.org/10.46799/ijssr.v3i12.652>
- Sari, L. D., Putri, R., & Amelia, D. (2023). Analisis pemetaan faktor risiko kejadian stunting menggunakan aplikasi GIS di Kabupaten Pasaman. *Health Information: Jurnal Penelitian*, 15(1), 55–63. <https://doi.org/10.36990/hijp.v15i1.925>

- Shahriar, M. M., Iqbal, M. S., Mitra, S., & Das, A. K. (2019). A ANN Approach to Predict Malnutrition Status of 0-59 Month's Older Children in Bangladesh. *2019 IEEE International Conference on Industry 4.0, Artificial Intelligence, and Communications Technology (IAICT)*, 145–149. <https://doi.org/10.1109/ICIAICT.2019.8784823>
- Shamsuddin, S., Rahman, A., & Karim, R. (2022). A review of spatial analysis application in childhood malnutrition studies. *International Journal of Health Sciences Research*, 12(6), 33–42. <https://doi.org/10.52403/ijhsr.20220605>
- Sinaga, K. P., & Yang, M.-S. (2020). Unsupervised K-Means Clustering Algorithm. *IEEE Access*, 8, 80716–80727. <https://doi.org/10.1109/ACCESS.2020.2988796>
- SMERU Research Institute. (2022). Nutrition map: Peta status gizi di Indonesia. The SMERU Research Institute. <https://smeru.or.id/en/nutritionmap>
- Suraya, S., Sholeh, M., & Lestari, U. (2023). Evaluation of Data Clustering Accuracy using K-Means Algorithm. *International Journal of Multidisciplinary Approach Research and Science*, 2(01), 385–396. <https://doi.org/10.59653/ijmars.v2i01.504>
- Talukder, A., & Ahammed, B. (2020). Machine learning algorithms for predicting malnutrition among under-five children in Bangladesh. *Nutrition*, 78, 110861. <https://doi.org/10.1016/j.nut.2020.110861>
- Utomo, S., Nugraha, I. M., & Wulandari, R. (2022). Analysis data spatial for nutrition programs: A review using geographic information system. *International Journal of Public Health Science*, 11(4), 1187–1195. <https://doi.org/10.11591/ijphs.v11i4.23590>
- Widjaja, R. D. (2019). *Using Deep Neural Networks for Lane Change Identification*.
- Yanto, M., Hadi, F., & Arlis, S. (2024). Determination of children's nutritional status with machine learning classification analysis approach. *Indonesian Journal of Electrical Engineering and Computer Science*, 33(1), 303. <https://doi.org/10.11591/ijeecs.v33.i1.pp303-313>
- Zhang, X., Usman, M., Irshad, A. U. R., Rashid, M., & Khattak, A. (2024). Investigating Spatial Effects through Machine Learning and Leveraging Explainable AI for Child Malnutrition in Pakistan. *ISPRS International Journal of Geo-Information*, 13(9), 330. <https://doi.org/10.3390/ijgi13090330>

LAMPIRAN

Program ANN dan Random Forest

```
from flask import Flask, render_template, request, jsonify,
send_from_directory
import pandas as pd
import numpy as np
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import StandardScaler, LabelEncoder,
MinMaxScaler
from sklearn.metrics import accuracy_score, classification_report
from sklearn.utils.class_weight import compute_class_weight
from sklearn.model_selection import GridSearchCV,
RandomizedSearchCV
from sklearn.ensemble import RandomForestClassifier
from sklearn.model_selection import cross_val_score
import tensorflow as tf
from tensorflow import keras
from tensorflow.keras import layers
import matplotlib
matplotlib.use('Agg') # Set backend sebelum import pyplot
import matplotlib.pyplot as plt
import seaborn as sns
from matplotlib.patches import Patch

# Fix untuk Windows threading issues
import threading
import warnings
warnings.filterwarnings('ignore')

# Set matplotlib untuk Windows
plt.ioff() # Turn off interactive mode
import io
import base64
from collections import defaultdict
import json
import os
import random

# Helper function to convert numpy types to Python native types
for JSON serialization
def convert_numpy_types(obj):
    """Recursively convert numpy types to Python native types for
    JSON serialization"""
    import numpy as np

    if isinstance(obj, dict):
        return {str(convert_numpy_types(k)):
convert_numpy_types(v) for k, v in obj.items()}
    elif isinstance(obj, list):
        return [convert_numpy_types(item) for item in obj]
    elif isinstance(obj, tuple):
        return tuple(convert_numpy_types(item) for item in obj)
    elif isinstance(obj, np.ndarray):
        return convert_numpy_types(obj.tolist())
    elif isinstance(obj, np.generic):
        # Handle all numpy scalar types using np.generic base
```

```

class
    # This is compatible with both NumPy 1.x and 2.x
    if isinstance(obj, np.integer):
        return int(obj)
    elif isinstance(obj, np.floating):
        return float(obj)
    elif isinstance(obj, np.bool_):
        return bool(obj)
    elif isinstance(obj, np.complexfloating):
        return complex(obj)
    else:
        # Fallback: try to convert using item() method
        try:
            return obj.item()
        except:
            return str(obj)
    elif obj is None or isinstance(obj, (str, int, float, bool,
complex)):
        return obj
    else:
        # Try to convert if it's a numpy scalar (fallback)
        try:
            if hasattr(obj, 'item'):
                result = obj.item()
                # Recursively convert the result in case it's
still a numpy type
                return convert_numpy_types(result)
        except:
            pass
        return obj

# Base directory for building absolute file paths to avoid
FileNotFoundError
BASE_DIR = os.path.dirname(os.path.abspath(__file__))
DATA_CSV_PATH = os.path.join(BASE_DIR, 'data_gizi_lengkap2.csv')
GEOJSON_DIR = BASE_DIR
BEST_MODEL_PATH = os.path.join(BASE_DIR, 'best_model.keras')
MODEL_METRICS_PATH = os.path.join(BASE_DIR, 'model_metrics.json')

app = Flask(__name__, template_folder=os.path.join(BASE_DIR,
'templates'))

# Set random seeds untuk reproduksi
def set_random_seeds(seed=42):
    """Set random seeds untuk semua library"""
    os.environ['PYTHONHASHSEED'] = str(seed)
    random.seed(seed)
    np.random.seed(seed)
    tf.random.set_seed(seed)
    os.environ['TF_DETERMINISTIC_OPS'] = '1'
    os.environ['TF_CUDNN_DETERMINISTIC'] = '1'

    # Fix untuk Windows matplotlib issues
    os.environ['MPLBACKEND'] = 'Agg'
    matplotlib.rcParams['figure.max_open_warning'] = 0

# Set seeds di awal
set_random_seeds(42)

```

```

# Load dan preprocess data dengan normalisasi yang lebih baik
def load_and_preprocess_data():
    # Load data
    if not os.path.exists(DATA_CSV_PATH):
        raise FileNotFoundError(f>Data CSV not found at:
{DATA_CSV_PATH}")
    df = pd.read_csv(DATA_CSV_PATH)

    # Encode categorical variables
    le_kelamin = LabelEncoder()
    le_status = LabelEncoder()

    df['JENIS_KELAMIN_ENCODED'] =
le_kelamin.fit_transform(df['JENIS_KELAMIN'])
    df['STATUS_ENCODED'] = le_status.fit_transform(df['STATUS'])

    # Tambahkan perhitungan BMI
    df['BMI'] = df['BB'] / ((df['TB'] / 100) ** 2)

    # Handle NaN values TERLEBIH DAHULU sebelum feature
engineering
    # Untuk kolom numerik, gunakan mean
    numeric_columns = ['UMUR', 'BB', 'TB']
    for col in numeric_columns:
        if col in df.columns:
            df[col] = df[col].fillna(df[col].mean())

    # Feature engineering tambahan untuk meningkatkan akurasi
    df['BB_TB_RATIO'] = df['BB'] / df['TB'] # Rasio berat
terhadap tinggi
    df['UMUR_SQUARED'] = df['UMUR'] ** 2 # Umur kuadrat
    df['BMI_SQUARED'] = df['BMI'] ** 2 # BMI kuadrat
    df['BB_SQUARED'] = df['BB'] ** 2 # Berat badan kuadrat
    df['TB_SQUARED'] = df['TB'] ** 2 # Tinggi badan kuadrat

    # Kategori umur - setelah NaN dihandle
    df['UMUR_CATEGORY'] = pd.cut(df['UMUR'],
                                bins=[0, 12, 24, 36, 48, 60, 100],
                                labels=[0, 1, 2, 3, 4, 5])

    df['UMUR_CATEGORY'] =
df['UMUR_CATEGORY'].fillna(0).astype(int)

    # Handle NaN untuk feature engineering yang baru dibuat
    feature_columns = ['BB_TB_RATIO', 'UMUR_SQUARED',
'BMI_SQUARED', 'BB_SQUARED', 'TB_SQUARED']
    for col in feature_columns:
        if col in df.columns:
            df[col] = df[col].fillna(df[col].mean())

    return df, le_kelamin, le_status

# Build Neural Network dengan arsitektur yang dioptimalkan dan
hyperparameter tuning
def build_model(input_dim, num_classes, learning_rate=0.001,
dropout_rate=0.3,
                hidden_layers=[128, 64, 32], activation='relu'):
    """Build model dengan hyperparameter yang dapat diatur"""

    model = keras.Sequential()

```

```

    # Input layer
    model.add(layers.Dense(hidden_layers[0],
activation=activation, input_shape=(input_dim,),
name='dense_input'))
    model.add(layers.BatchNormalization(name='bn_input'))
    model.add(layers.Dropout(dropout_rate, name='dropout_input'))

    # Hidden layers
    for i, units in enumerate(hidden_layers[1:], 1):
        model.add(layers.Dense(units, activation=activation,
name=f'dense_hidden_{i}'))

    model.add(layers.BatchNormalization(name=f'bn_hidden_{i}'))
        model.add(layers.Dropout(dropout_rate * 0.8,
name=f'dropout_hidden_{i}'))

    # Output layer
    model.add(layers.Dense(num_classes, activation='softmax',
name='dense_output'))

    model.compile(
optimizer=keras.optimizers.Adam(learning_rate=learning_rate),
    loss='sparse_categorical_crossentropy',
    metrics=['accuracy']
)

    return model

# Hyperparameter tuning function
def hyperparameter_tuning(X_train, y_train, X_val, y_val):
    """Melakukan hyperparameter tuning menggunakan grid search"""

    # Define parameter grid
    param_grid = {
        'learning_rate': [0.001, 0.0005, 0.0001],
        'dropout_rate': [0.2, 0.3, 0.4],
        'hidden_layers': [[128, 64], [256, 128, 64], [128, 64,
32]],
        'activation': ['relu', 'tanh']
    }

    best_score = 0
    best_params = None
    best_model = None

    print("Melakukan hyperparameter tuning...")

    # Grid search manual untuk Keras models
    for lr in param_grid['learning_rate']:
        for dropout in param_grid['dropout_rate']:
            for hidden in param_grid['hidden_layers']:
                for activation in param_grid['activation']:
                    print(f"Testing: lr={lr}, dropout={dropout},
hidden={hidden}, activation={activation}")

                    # Build model dengan parameter ini
                    model = build_model(

```

```

        input_dim=X_train.shape[1],
        num_classes=len(np.unique(y_train)),
        learning_rate=lr,
        dropout_rate=dropout,
        hidden_layers=hidden,
        activation=activation
    )

    # Train model
    early_stopping =
keras.callbacks.EarlyStopping(
        monitor='val_loss',
        patience=10,
        restore_best_weights=True
    )

    history = model.fit(
        X_train, y_train,
        epochs=50,
        batch_size=32,
        validation_data=(X_val, y_val),
        callbacks=[early_stopping],
        verbose=0
    )

    # Evaluate
    val_loss, val_accuracy = model.evaluate(X_val,
y_val, verbose=0)

    if val_accuracy > best_score:
        best_score = val_accuracy
        best_params = {
            'learning_rate': lr,
            'dropout_rate': dropout,
            'hidden_layers': hidden,
            'activation': activation
        }
        best_model = model

    print(f"Best parameters: {best_params}")
    print(f"Best validation accuracy: {best_score:.4f}")

    return best_model, best_params

# Learning curve analysis
def plot_learning_curves(history, save_path=None):
    """Plot learning curves untuk analisis
    overfitting/underfitting"""

    fig, (ax1, ax2) = plt.subplots(1, 2, figsize=(15, 5))

    # Plot accuracy
    ax1.plot(history.history['accuracy'], label='Training
Accuracy')
    ax1.plot(history.history['val_accuracy'], label='Validation
Accuracy')
    ax1.set_title('Model Accuracy')
    ax1.set_xlabel('Epoch')
    ax1.set_ylabel('Accuracy')

```

```

ax1.legend()
ax1.grid(True)

# Plot loss
ax2.plot(history.history['loss'], label='Training Loss')
ax2.plot(history.history['val_loss'], label='Validation Loss')
ax2.set_title('Model Loss')
ax2.set_xlabel('Epoch')
ax2.set_ylabel('Loss')
ax2.legend()
ax2.grid(True)

plt.tight_layout()

if save_path:
    plt.savefig(save_path, dpi=300, bbox_inches='tight')

# Convert to base64 for web display
img = io.BytesIO()
plt.savefig(img, format='png', bbox_inches='tight')
img.seek(0)
learning_curve_img = base64.b64encode(img.getvalue()).decode()
plt.close()

return learning_curve_img

# Custom ANN parameters (global)
custom_ann_params = {
    'hidden_layers': [128, 64, 32],
    'learning_rate': 0.001,
    'dropout_rate': 0.3,
    'activation': 'relu',
    'batch_size': 32,
    'epochs': 200,
    'optimizer': 'adam',
    'use_custom': False # Flag untuk menggunakan custom params
atau auto-tuning
}

# Train model dengan perbaikan
def train_model(test_size=0.2, custom_params=None):
    # Set seeds untuk reproduksi
    set_random_seeds(42)

    df, le_kelamin, le_status = load_and_preprocess_data()

    # Use custom parameters if provided, otherwise use global
    custom_ann_params if enabled
    if custom_params is None:
        custom_params = custom_ann_params if
custom_ann_params['use_custom'] else None

    # Prepare features dengan feature engineering yang lebih baik
    X = df[['JENIS_KELAMIN_ENCODED', 'UMUR', 'BB', 'TB', 'BMI',
            'BB_TB_RATIO', 'UMUR_SQUARED', 'BMI_SQUARED',
            'BB_SQUARED', 'TB_SQUARED', 'UMUR_CATEGORY']].values
    y = df['STATUS_ENCODED'].values

    print(f"Jumlah features: {X.shape[1]}")

```

```

print(f"Jumlah samples: {X.shape[0]}")
print(f"Distribusi kelas: {np.bincount(y)}")
print(f"Test size: {test_size * 100}% (Train: {(1-test_size) * 100}%)")

# Split data dengan seed tetap
X_train, X_test, y_train, y_test = train_test_split(
    X, y, test_size=test_size, stratify=y, random_state=42
)

# Split validation set
X_train, X_val, y_train, y_val = train_test_split(
    X_train, y_train, test_size=0.2, stratify=y_train,
random_state=42
)

# Compute class weights untuk balancing
class_weights = compute_class_weight(
    'balanced',
    classes=np.unique(y_train),
    y=y_train
)

class_weight_dict = dict(zip(np.unique(y_train),
class_weights))

print(f"Class weights: {class_weight_dict}")

# Normalisasi data dengan StandardScaler
scaler = StandardScaler()
X_train_scaled = scaler.fit_transform(X_train)
X_val_scaled = scaler.transform(X_val)
X_test_scaled = scaler.transform(X_test)

# Hyperparameter tuning atau gunakan custom parameters
if custom_params and custom_params.get('use_custom', False):
    print("="*60)
    print("Menggunakan Custom ANN Parameters:")
    print(f"Hidden Layers: {custom_params['hidden_layers']}")
    print(f"Learning Rate: {custom_params['learning_rate']}")
    print(f"Dropout Rate: {custom_params['dropout_rate']}")
    print(f"Activation: {custom_params['activation']}")
    print(f"Batch Size: {custom_params['batch_size']}")
    print(f"Epochs: {custom_params['epochs']}")
    print(f"Optimizer: {custom_params['optimizer']}")
    print("="*60)

```