

BAB 1

PENDAHULUAN

1.1 Latar Belakang

Stunting merupakan salah satu permasalahan kesehatan masyarakat yang masih menjadi prioritas nasional di Indonesia. Berdasarkan data Survei Status Gizi Indonesia (SSGI) tahun 2023, prevalensi *stunting* nasional masih berada pada angka 21,6%, yang artinya satu dari lima anak balita mengalami gangguan pertumbuhan kronis yang ditandai dengan tinggi badan di bawah standar usianya. Pemerintah Indonesia menargetkan penurunan angka *stunting* menjadi 14% pada tahun 2024 sebagai bagian dari agenda pembangunan jangka menengah nasional. Namun demikian, pencapaian target ini masih menghadapi berbagai tantangan, terutama dalam aspek deteksi dini dan klasifikasi status gizi anak balita secara sistematis dan cepat.

Di tingkat daerah, khususnya Kabupaten Aceh Utara, masalah *stunting* menunjukkan tren yang mengkhawatirkan. Menurut laporan Dinas Kesehatan Aceh Utara tahun 2023, tercatat sekitar 32,1% balita mengalami *stunting*, angka ini jauh di atas rata-rata nasional. Wilayah ini menghadapi tantangan struktural dalam aspek sosio-ekonomi, akses pelayanan kesehatan, dan edukasi gizi yang masih terbatas. Situasi ini menuntut adanya upaya strategis untuk meningkatkan efektivitas intervensi, yang salah satunya dapat dilakukan melalui pendekatan berbasis teknologi dalam klasifikasi status *stunting* secara lebih akurat dan cepat.

Klasifikasi status *stunting* yang dilakukan secara manual seringkali menghadapi kendala seperti keterbatasan sumber daya manusia dan *human error* dalam interpretasi data. Oleh karena itu, diperlukan suatu metode yang mampu membantu proses klasifikasi secara otomatis, efisien, dan dapat diandalkan. Teknologi informasi dan data mining telah terbukti mampu memberikan solusi dalam proses pengambilan keputusan berbasis data, khususnya dalam bidang kesehatan masyarakat. Salah satu pendekatan yang dapat digunakan adalah pemanfaatan algoritma klasifikasi seperti C5.0 yang telah banyak diterapkan dalam penelitian terkait kesehatan anak.

Stunting atau kondisi gagal tumbuh akibat kekurangan gizi kronis merupakan masalah kesehatan global yang berdampak jangka panjang terhadap perkembangan fisik dan kognitif anak. Menurut UNICEF (2023), sekitar 22,3% balita di seluruh dunia mengalami *stunting*, dengan prevalensi tertinggi di Asia Selatan dan Sub-Sahara Afrika. Di Indonesia, hasil Survei Status Gizi Indonesia (SSGI) tahun 2022 menunjukkan bahwa 21,6% anak balita mengalami *stunting*, yang berarti hampir 1 dari 5 anak tidak mencapai potensi pertumbuhan maksimalnya. Kondisi ini menuntut solusi inovatif dalam deteksi dan penanganan yang lebih cepat, efektif, dan berbasis teknologi.

Dalam konteks global, transformasi digital dan revolusi industri 4.0 telah memberikan peluang signifikan untuk memanfaatkan teknologi cerdas, termasuk kecerdasan buatan dan pembelajaran mesin (*machine learning*), dalam sektor kesehatan masyarakat. Penerapan teknologi ini terutama pada deteksi dini *stunting* menjadi semakin relevan, mengingat pentingnya intervensi pada periode 1.000 hari pertama kehidupan. Dengan semakin meningkatnya penetrasi *smartphone* dan akses data kesehatan digital, metode klasifikasi otomatis berbasis algoritma kini menjadi strategi potensial dalam mendeteksi risiko *stunting* sejak dini (Chilyabanyama *et al.*, 2022).

Sejumlah penelitian telah menerapkan berbagai algoritma *machine learning* untuk klasifikasi status gizi anak. Salah satu studi oleh Jain *et al* (2022) menggunakan algoritma *Random Forest* dan menemukan tingkat akurasi tinggi dalam mengklasifikasikan status *stunting* berdasarkan data antropometri. Demikian pula, metode *K-Nearest Neighbor* (KNN) dan *Naïve Bayes* digunakan dalam studi oleh (Almasoud *et al.*, 2022), yang mencatat akurasi hingga 94%, menunjukkan efektivitas pendekatan non-parametrik dalam klasifikasi masalah ini. Namun, tantangan utama masih berkaitan dengan akurasi, interpretabilitas, dan efisiensi komputasi.

Meskipun algoritma seperti *Random Forest*, KNN, dan *Support Vector Machine* (SVM) telah terbukti efektif, masih terdapat keterbatasan dalam hal interpretabilitas dan kebutuhan pemrosesan yang tinggi. Di sinilah algoritma C5.0 hadir sebagai solusi potensial. C5.0, sebagai turunan dari algoritma C4.5,

menawarkan keunggulan dalam kecepatan, penggunaan memori yang lebih efisien, serta kemampuan menghasilkan model pohon keputusan yang lebih sederhana dan mudah dipahami (Asih & Hasibuan, 2024). Keunggulan ini menjadi sangat penting dalam konteks aplikasi *mobile* yang membutuhkan sistem ringan dan *real-time*.

Studi oleh Asih dan Hasibuan (2024) menunjukkan bahwa implementasi algoritma C5.0 dalam aplikasi Android mampu mengklasifikasikan status *stunting* dengan tingkat akurasi mencapai 100%, tanpa perlu kunjungan langsung ke fasilitas kesehatan. Ini menunjukkan potensi besar dalam penerapan sistem deteksi dini berbasis *mobile*. Namun, studi ini masih terbatas dari segi skala data dan validasi silang, serta belum dibandingkan secara komprehensif dengan algoritma lainnya. Oleh karena itu, evaluasi yang lebih luas dan metodologis diperlukan untuk menguji replikasi dan generalisasi temuan tersebut.

Kesenjangan penelitian juga muncul dari kurangnya integrasi teknik praproses data dan fitur seleksi yang optimal dalam studi terdahulu. Syahrial *et al.*, (2022) menunjukkan bahwa pemilihan fitur sangat memengaruhi akurasi model klasifikasi *stunting*, namun banyak studi sebelumnya yang menggunakan semua atribut tanpa seleksi, sehingga meningkatkan *noise* dan kompleksitas model. Penelitian ini bertujuan untuk memperbaiki aspek tersebut dengan pendekatan fitur seleksi berbasis gain ratio atau information gain yang terintegrasi dalam algoritma C5.0.

Lebih jauh, sebagian besar studi sebelumnya hanya membandingkan performa algoritma berdasarkan akurasi, tanpa memperhatikan aspek interpretasi model yang esensial dalam konteks kebijakan publik dan intervensi kesehatan. Model yang mudah dipahami akan lebih diterima oleh tenaga medis dan petugas lapangan kesehatan. C5.0 yang menghasilkan rules dalam bentuk logika IF-THEN memiliki keunggulan ini dibanding algoritma seperti SVM atau *neural networks* yang bersifat “*black box*” (Lenatti *et al.*, 2023).

Melalui penelitian ini, akan dikembangkan sistem klasifikasi status *stunting* berbasis algoritma C5.0 yang tidak hanya akurat, tetapi juga mudah diinterpretasi dan efisien dari sisi komputasi. Sistem ini akan diuji dengan dataset riil. Penelitian

ini juga akan menguji performa model dalam berbagai skenario data, termasuk kondisi data *imbalance* yang umum terjadi di lapangan (Mukhsin *et al.*, 2023).

Secara teoritis, penelitian ini memberikan kontribusi dalam pengembangan ilmu komputer terapan di bidang kesehatan masyarakat, khususnya dalam pengembangan model klasifikasi yang akurat dan *interpretable*. Secara praktis, hasil penelitian ini dapat digunakan untuk membangun sistem deteksi dini *stunting* yang berbasis website, memungkinkan tenaga kesehatan dan masyarakat umum melakukan skrining awal secara mandiri dan efisien. Penelitian ini berpotensi mempercepat intervensi nutrisi dan mengurangi beban sistem kesehatan nasional.

Tujuan utama dari penelitian ini adalah mengembangkan dan mengevaluasi model klasifikasi status *stunting* pada balita menggunakan algoritma C5.0 dari sisi akurasi, interpretabilitas, dan efisiensi. Manfaat teoritis mencakup pengayaan metodologi klasifikasi data kesehatan berbasis *machine learning*, sementara manfaat praktis mencakup pengembangan sistem yang dapat diintegrasikan ke dalam website untuk deteksi dini *stunting* di wilayah dengan keterbatasan akses layanan kesehatan.

1.2 Rumusan Masalah

Berdasarkan latar belakang di atas, maka rumusan masalah yang akan dibahas dalam penelitian ini adalah :

1. Bagaimana kinerja algoritma C5.0 dalam mengklasifikasikan status *stunting* berdasarkan data karakteristik balita di Puskesmas Tanah Jambo Aye?
2. Faktor-faktor apa saja yang paling signifikan dalam mempengaruhi hasil klasifikasi status *stunting* menggunakan algoritma C5.0?
3. Seberapa tinggi tingkat akurasi, presisi, dan *recall* dari model klasifikasi berbasis C5.0 dalam memprediksi status *stunting* pada data lokal Puskesmas Tanah Jambo Aye?

1.3 Tujuan Penelitian

Adapun tujuan yang ingin dicapai pada penelitian ini adalah :

1. Menganalisis dan mengimplementasikan algoritma C5.0 untuk membangun model klasifikasi status *stunting* pada balita berdasarkan data pada Puskesmas Tanah Jambo Aye.
2. Mengidentifikasi atribut-atribut paling berpengaruh dalam klasifikasi status *stunting* berdasarkan hasil pemodelan algoritma C5.0
3. Mengevaluasi performa model C5.0 melalui metrik akurasi, presisi, dan *recall* guna menilai keandalan sistem dalam mendeteksi risiko *stunting* di tingkat daerah.

1.4 Manfaat Penelitian

Manfaat penelitian ini adalah sebagai berikut :

1. Mempermudah dalam menentukan status gizi balita di Puskesmas Tanah Jambo Aye.
2. Mengetahui performa algoritma C5.0 pada sistem klasifikasi.
3. Sebagai studi pustaka pada kegiatan-kegiatan penelitian selanjutnya.

1.5 Ruang Lingkup dan Batasan Penelitian

Penelitian ini dibatasi pada ruang lingkup dan variabel-variabel tertentu untuk menjaga fokus dan kejelasan analisis. Adapun batasan-batasan yang ditetapkan dalam penelitian ini adalah sebagai berikut:

1. Ruang lingkup wilayah penelitian ini dibatasi pada wilayah administratif Kabupaten Aceh Utara, Provinsi Aceh. Data yang digunakan merupakan data sekunder yang diperoleh dari Dinas Kesehatan dan/atau Puskesmas yang berada di wilayah tersebut, yang mencakup karakteristik balita usia 0–59 bulan.
2. Populasi dan sampel populasi dalam penelitian ini adalah seluruh anak balita usia 0–59 bulan yang tercatat dalam data dari Puskesmas Tanah Jambo Aye pada tahun terakhir yang tersedia (misalnya tahun 2024 atau 2025). Sampel penelitian diambil secara *purposive* dengan kriteria: memiliki data antropometri lengkap dan riwayat kesehatan anak serta keluarga.

3. Batasan Variabel :
 - a. Variabel dependen (terikat) dalam penelitian ini adalah status *stunting* yang dikategorikan ke dalam dua kelas, yaitu: normal, dan *stunted*, dan sesuai dengan kriteria WHO berdasarkan nilai z-score tinggi badan menurut umur (*height-for-age z-score/Haz*).
 - b. Variabel independen (bebas) mencakup: usia anak, jenis kelamin, berat badan, tinggi badan, usia saat ukur, tanggal pengukuran.
4. Batasan topik dan teknik analisis penelitian ini hanya membahas implementasi dan evaluasi performa algoritma C5.0 dalam mengklasifikasikan status *stunting*. Penelitian tidak membandingkan algoritma lain secara langsung (seperti C4.5, *Random Forest*, atau *Gradient Boosting*), kecuali dalam kajian pustaka. Teknik evaluasi model dibatasi pada metrik akurasi, *precision*, *recall*, dan *confusion matrix*, tanpa melakukan pengujian statistik inferensial terhadap hubungan antar variabel.
5. Aspek temporal dan teknologi pengolahan data dalam penelitian ini bersifat statis (*cross-sectional*) dengan cakupan waktu satu tahun, sehingga belum mampu menggambarkan dinamika perubahan secara *longitudinal*. Proses pengolahan dan perhitungan data dilakukan secara manual menggunakan perangkat lunak *spreadsheet* (Microsoft Excel), sedangkan implementasi model klasifikasi direalisasikan dalam bentuk sistem berbasis web yang dikembangkan menggunakan bahasa pemrograman PHP, tanpa integrasi langsung dengan sistem informasi kesehatan daerah.