

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

Pada masa balita (usia 0–59 bulan), proses pertumbuhan dan perkembangan anak sangat dipengaruhi oleh kondisi status gizinya. Pemenuhan asupan gizi yang tepat dan seimbang pada periode ini berperan penting dalam menentukan kualitas kesehatan, tingkat kecerdasan, serta produktivitas anak di masa yang akan datang. Sebaliknya, permasalahan gizi seperti gizi buruk, gizi kurang, maupun gizi lebih (*overweight*/obesitas) dapat menimbulkan dampak jangka panjang yang serius, antara lain gangguan pertumbuhan fisik, menurunnya sistem kekebalan tubuh, keterlambatan perkembangan kognitif, serta meningkatnya risiko terjadinya penyakit kronis pada usia dewasa [1].

Data Kementerian Kesehatan Republik Indonesia menunjukkan bahwa meskipun prevalensi stunting dan gizi kurang pada balita mengalami penurunan dalam beberapa tahun terakhir, permasalahan gizi masih menjadi tantangan yang cukup besar. Berdasarkan hasil Studi Status Gizi Indonesia (SSGI) tahun 2022, angka prevalensi stunting tercatat sebesar 21,6%, menurun dibandingkan tahun 2021 yang mencapai 24,4%. Namun demikian, capaian tersebut masih berada di atas target global prevalensi stunting di bawah 20% serta target nasional sebesar 14% pada tahun 2024 [2]. Di samping itu, meningkatnya kasus gizi lebih pada balita juga menjadi perhatian tersendiri, seiring dengan perubahan pola konsumsi dan gaya hidup masyarakat modern [3].

Penanganan masalah gizi balita membutuhkan intervensi yang tepat dan terarah, namun seringkali tenaga kesehatan, khususnya bidan dan ahli gizi di tingkat Posyandu atau Puskesmas, menghadapi kendala dalam memberikan rekomendasi gizi yang spesifik dan personal untuk setiap balita. Kendala tersebut meliputi keterbatasan waktu, banyaknya data yang harus dianalisis (umur, berat badan, tinggi badan, riwayat pemberian vitamin, riwayat pemberian asi), serta variasi kondisi individu balita yang sangat beragam. Kesalahan atau keterlambatan dalam diagnosis status gizi dan penentuan *intervensi* yang tepat dapat memperburuk kondisi gizi balita [4].

Di era pencatatan elektronik (Posyandu, Puskesmas, survei), volume data balita kian besar dan bervariasi. *Machine Learning* (ML) menawarkan pendekatan prediktif untuk klasifikasi status gizi secara lebih awal (*early warning*), memanfaatkan pola multi-faktor (usia, jenis kelamin, ukuran antropometri, riwayat penyakit, faktor lingkungan). Penentuan status gizi balita umumnya

dilakukan dengan indikator antropometri seperti *Z-score* BB/U, TB/U, BB/TB, LiLA dan lingkaran kepala dengan acuan standar dari WHO dan Kemenkes [5] namun, pendekatan manual (menggunakan tabel *Z-score* dan KMS) masih banyak digunakan di lapangan dan rentan terhadap kesalahan pengukuran serta memakan waktu.

Penanganan masalah gizi balita membutuhkan intervensi yang tepat dan terarah, namun seringkali tenaga kesehatan, khususnya bidan dan ahli gizi di tingkat Posyandu atau Puskesmas, menghadapi kendala dalam memberikan rekomendasi gizi yang spesifik dan personal untuk setiap balita. Kendala tersebut meliputi keterbatasan waktu, banyaknya data yang harus dianalisis (umur, berat badan, tinggi badan, riwayat pemberian vitamin, riwayat pemberian asi), serta variasi kondisi individu balita yang sangat beragam. Kesalahan atau keterlambatan dalam diagnosis status gizi dan penentuan *intervensi* yang tepat dapat memperburuk kondisi gizi balita [4].

Di era pencatatan elektronik (Posyandu, Puskesmas, survei), volume data balita kian besar dan bervariasi. Machine Learning (ML) menawarkan pendekatan prediktif untuk klasifikasi status gizi secara lebih awal (*early warning*), memanfaatkan pola multi-faktor (usia, jenis kelamin, ukuran antropometri, riwayat penyakit, faktor lingkungan). Penentuan status gizi balita umumnya dilakukan dengan indikator antropometri seperti *Z-score* BB/U, TB/U, BB/TB, LiLA dan lingkaran kepala dengan acuan standar dari WHO dan Kemenkes [5] namun, pendekatan manual (menggunakan tabel *Z-score* dan KMS) masih banyak digunakan di lapangan dan rentan terhadap kesalahan pengukuran serta memakan waktu.

Upaya untuk mengatasi hal ini yaitu dengan dikembangkannya sistem komputer untuk membantu dalam proses identifikasi status gizi balita secara otomatis. Pendekatan menggunakan metode *machine learning*. (RF) merupakan algoritma *ensemble learning* berbasis *Decision Tree* yang menggabungkan hasil dari banyak pohon keputusan untuk meningkatkan akurasi prediksi. Penelitian oleh [6] menunjukkan bahwa yang dilengkapi teknik *feature selection* RFE dan penyeimbangan data SMOTE mampu mencapai akurasi 95% dalam klasifikasi status gizi balita. Keunggulan RF adalah kemampuannya mengatasi *overfitting*, memberikan estimasi pentingnya variabel (*feature importance*), dan bekerja dengan baik pada dataset berukuran besar., telah terbukti efektif dalam klasifikasi status gizi atau stunting pada balita dengan tingkat akurasi tinggi dan proven dalam praktik di Indonesia[7] Sebagai contoh, penelitian “Machine Learning Method to Predict the Toddlers’ Nutritional Status” oleh Gustriansyah et al. (2024) [8] menunjukkan bahwa memberi akurasi hingga 97,37%, sensitivitas 95,00%, spesifisitas 98,81%, AUC 0,9990,

dan Cohen's Kappa 0,9609, menjadikannya metode yang paling direkomendasikan untuk prediksi gizi balita [9]. Selain itu, penelitian oleh Satria et al. (2024) [6] di Samarinda berhasil meningkatkan akurasi model dari 91,91% menjadi 98,39% melalui penggunaan teknik seleksi fitur (RFE) dan optimasi parameter menggunakan Genetic Algorithm, bahkan mencapai 99,70% pada akurasi akhir. Selain itu, penerapan yang dioptimasi dengan algoritma genetika seperti pada studi kasus Puskesmas Cakru menunjukkan peningkatan kinerja hingga F1-score mencapai 65 %, dan akurasi hampir 90 %. Dengan demikian, pengembangan klasifikasi gizi balita berbasis sangat layak untuk dijadikan solusi menerjemahkan data antropometri menjadi status gizi secara cepat, akurat, dan andal [10].

Pentingnya penelitian ini untuk mengembangkan model klasifikasi yang dapat membantu para petugas kesehatan dan pengambil kebijakan dalam mendeteksi status gizi balita secara lebih cepat dan akurat. Dengan adanya sistem yang dapat mengklasifikasikan status gizi secara otomatis menggunakan algoritma pembelajaran mesin, diharapkan dapat mengurangi beban kerja petugas kesehatan dan mempercepat intervensi yang dibutuhkan pada balita yang memiliki masalah gizi. Sebagai contoh, dalam situasi darurat atau di daerah terpencil, penggunaan model berbasis teknologi ini bisa sangat membantu dalam memberikan diagnosis awal dan rekomendasi tindakan lebih lanjut tanpa harus menunggu hasil pemeriksaan manual. Model klasifikasi ini tidak hanya diharapkan untuk meningkatkan pemantauan status gizi balita, tetapi juga untuk mendukung kebijakan kesehatan masyarakat yang lebih berbasis data. Dengan sistem yang akurat dan efisien, pemantauan status gizi dapat dilakukan lebih rutin dan dengan cakupan yang lebih luas, termasuk daerah-daerah yang selama ini sulit dijangkau oleh tenaga medis. Selain itu, informasi yang dihasilkan dari sistem ini bisa menjadi dasar bagi pengambilan kebijakan yang lebih terarah dan efektif dalam mengatasi masalah gizi pada balita di Indonesia [11].

Pada praktiknya, penentuan status gizi balita di tingkat pelayanan kesehatan seperti Posyandu dan Puskesmas dilakukan dengan mengacu pada standar antropometri WHO melalui perhitungan nilai Z-score berdasarkan indikator berat badan menurut umur (BB/U), tinggi badan menurut umur (TB/U), berat badan menurut tinggi badan (BB/TB), lingkaran lengan atas (LiLA), dan lingkaran kepala. Selain pengukuran antropometri, tenaga kesehatan di Puskesmas juga mempertimbangkan faktor pendukung lainnya, seperti riwayat pemberian ASI dan pemberian vitamin A, sebagai bagian dari evaluasi kondisi gizi balita secara menyeluruh. Riwayat pemberian ASI berperan penting dalam mendukung pertumbuhan dan sistem kekebalan tubuh anak, sementara pemberian vitamin A berkontribusi dalam menjaga kesehatan mata, meningkatkan daya tahan tubuh, serta menurunkan risiko penyakit infeksi. Seluruh proses penilaian status gizi yang dilakukan di

Puskesmas tersebut telah disesuaikan dengan standar dan pedoman WHO, sehingga hasil pengukuran dan interpretasi status gizi tetap berada dalam kerangka acuan global yang berlaku. Standar WHO tersebut bersifat global dan digunakan sebagai acuan baku untuk menggambarkan kondisi pertumbuhan anak secara umum. Namun demikian, penerapan standar WHO di lapangan sering kali dihadapkan pada berbagai keterbatasan, terutama ketika berhadapan dengan data riil Puskesmas yang memiliki karakteristik sangat beragam.

Data antropometri yang diperoleh dari Puskesmas merupakan data lapangan yang dipengaruhi oleh berbagai faktor, seperti perbedaan metode pengukuran, kondisi lingkungan, karakteristik sosial ekonomi, serta variasi pertumbuhan balita di tingkat lokal. Kondisi ini menyebabkan hasil klasifikasi status gizi berdasarkan tabel Z-score WHO tidak selalu mencerminkan pola pertumbuhan balita secara optimal pada wilayah tertentu. Selain itu, data Puskesmas umumnya memiliki permasalahan seperti ketidakseimbangan jumlah data antar kelas status gizi, adanya data pencilan (outlier), serta hubungan antar variabel yang bersifat non-linear.

Oleh karena itu, diperlukan suatu pendekatan komputasional yang tidak hanya mengandalkan aturan baku WHO, tetapi juga mampu mempelajari pola lokal dari data Puskesmas secara adaptif. Pendekatan machine learning menjadi solusi yang relevan karena mampu mengolah data dalam jumlah besar dan kompleks, serta mengintegrasikan berbagai variabel antropometri dan faktor pendukung lainnya. Salah satu algoritma yang sesuai untuk permasalahan ini adalah Random Forest, yang merupakan metode ensemble learning berbasis decision tree.

Random Forest bekerja dengan membangun banyak pohon keputusan yang masing-masing dilatih menggunakan subset data dan fitur yang berbeda. Setiap pohon menghasilkan prediksi status gizi balita, kemudian hasil prediksi tersebut digabungkan menggunakan mekanisme majority voting untuk menentukan keputusan akhir. Mekanisme majority voting ini memungkinkan model menghasilkan klasifikasi yang lebih stabil dan representatif, karena keputusan akhir tidak bergantung pada satu pohon atau satu aturan saja, melainkan merupakan hasil konsensus dari banyak pohon yang mempelajari variasi data Puskesmas.

Dengan menggabungkan standar WHO sebagai dasar perhitungan Z-score dan Random Forest sebagai model klasifikasi berbasis data lokal, penelitian ini diharapkan mampu menjembatani perbedaan antara pendekatan teoretis global dan kondisi lapangan. Pendekatan ini diharapkan dapat menghasilkan sistem klasifikasi status gizi balita yang lebih akurat, adaptif, dan relevan untuk mendukung pengambilan keputusan di Posyandu dan Puskesmas, khususnya di Kecamatan Menganti, Kabupaten Gresik.

## 1.2 Rumusan Masalah

Penentuan status gizi balita merupakan langkah penting dalam pemantauan pertumbuhan dan pencegahan masalah gizi sejak dini. Namun, proses klasifikasi yang masih banyak dilakukan secara manual dengan mengacu pada tabel *Z-score* WHO atau Kartu Menuju Sehat (KMS) sering kali memakan waktu, rawan kesalahan perhitungan, dan tidak efisien ketika volume data sangat besar. Selain itu, perbedaan metode pengukuran di lapangan dapat menyebabkan inkonsistensi data yang memengaruhi akurasi penilaian. Dengan berkembangnya teknologi machine learning, khususnya algoritma, peluang untuk mengotomatisasi dan meningkatkan akurasi klasifikasi gizi balita semakin besar. Oleh karena itu, diperlukan perumusan masalah yang jelas untuk membangun sistem klasifikasi berbasis yang efektif, akurat, mampu menangani data dalam jumlah besar, serta sesuai dengan standar WHO dan Permenkes No. 2 Tahun 2020.

### 1.2.1 Permasalahan

Permasalahan utama dalam penelitian ini adalah bagaimana merancang dan mengimplementasikan model klasifikasi status gizi balita berbasis data antropometri (BB/U, TB/U, BB/TB) dengan menggunakan algoritma agar mampu menghasilkan akurasi tinggi dan sesuai standar klasifikasi *Z-score* WHO serta Permenkes. Tantangan yang dihadapi meliputi optimalisasi parameter, penanganan ketidakseimbangan data antar kelas gizi, serta interpretabilitas model untuk mengetahui faktor antropometri yang paling berpengaruh. Selain itu, diperlukan evaluasi model menggunakan metrik yang lebih komprehensif seperti *Precision*, *recall*, F1-score, dan AUC guna memastikan keandalan model dalam mendeteksi status gizi, khususnya pada kasus gizi buruk.

### 1.2.2 Solusi Permasalahan

Solusi dari permasalahan penelitian ini adalah dengan merancang model klasifikasi status gizi balita menggunakan algoritma yang dioptimalkan secara menyeluruh. Proses optimalisasi dilakukan melalui penentuan parameter terbaik seperti jumlah pohon, kedalaman pohon, dan jumlah fitur yang dipertimbangkan pada setiap pemisahan agar model memiliki akurasi tinggi dan tidak mengalami *overfitting*. Untuk memastikan kesesuaian dengan standar WHO dan Permenkes, data antropometri balita akan diolah dalam bentuk *Z-score* sebelum dilakukan klasifikasi. Masalah ketidakseimbangan data antar kelas gizi ditangani dengan teknik resampling, SMOTE, atau penerapan class weight sehingga distribusi data lebih seimbang dan hasil klasifikasi lebih adil. Interpretabilitas model diperoleh melalui analisis feature importance untuk mengetahui variabel antropometri yang paling memengaruhi hasil klasifikasi. Selanjutnya, performa model akan dievaluasi menggunakan berbagai metrik seperti akurasi, *Precision*, *recall*, F1-score, dan AUC,

sehingga model yang dihasilkan tidak hanya unggul dalam akurasi, tetapi juga andal dalam mendeteksi kasus gizi buruk yang jumlahnya relatif sedikit namun sangat krusial.

### 1.2.3 Pertanyaan Penelitian

Berdasarkan uraian permasalahan dan solusi yang diusulkan, penelitian ini memunculkan beberapa pertanyaan penelitian:

1. Bagaimana performa model dalam klasifikasi status gizi balita jika dievaluasi menggunakan metrik akurasi, *Precision*, *recall*, F1-score?
2. Berapakah jumlah variasi parameter *n\_estimators*, *max\_depth*, *max\_features*, dan *criterion* pada algoritma Random Forest yang memengaruhi tingkat performa klasifikasi status gizi balita berdasarkan data antropometri yang terbaik?

## 1.3 Tujuan Penelitian

### 1.3.1 Tujuan

Tujuan dan manfaat yang dihasilkan dari penelitian ini adalah sebagai berikut:

1. Menganalisis performa model Random Forest dalam melakukan klasifikasi status gizi balita berdasarkan data antropometri dengan mengevaluasi hasil klasifikasi menggunakan metrik akurasi, precision, recall, dan F1-score.
2. Menganalisis pengaruh variasi parameter algoritma Random Forest, meliputi *n\_estimators*, *max\_depth*, *max\_features*, dan *criterion*, terhadap tingkat performa klasifikasi status gizi balita, sehingga diperoleh konfigurasi parameter yang menghasilkan kinerja model optimal.

### 1.3.2 Manfaat

Penelitian ini diharapkan dapat memberikan manfaat akademis dan praktis. Secara akademis, penelitian ini menambah pemahaman mengenai penerapan algoritma Random Forest dan pengaruh pengaturan parameter terhadap performa klasifikasi status gizi balita berbasis data antropometri. Secara praktis, hasil penelitian ini diharapkan dapat menghasilkan model klasifikasi status gizi balita yang akurat dan teruji, sehingga dapat membantu tenaga kesehatan dalam mengidentifikasi status gizi balita secara lebih cepat dan objektif sesuai dengan standar WHO dan Peraturan Menteri Kesehatan, serta mendukung perencanaan dan evaluasi program penanggulangan masalah gizi balita.

### 1.4 Batasan Masalah

Penelitian ini dibatasi pada beberapa hal sebagai berikut:

1. Data yang digunakan adalah data antropometri balita usia 0–59 bulan yang mencakup berat badan, tinggi/panjang badan, umur, Lingkar Lengan Atas, dan lingkar kepala, yang kemudian dikonversi ke *Z-score* sesuai standar WHO.
2. Kategori status gizi yang dianalisis mengacu pada klasifikasi WHO berbasis *Z-score*, yaitu: gizi buruk, gizi kurang, gizi baik/normal, dan gizi lebih.
3. Algoritma yang digunakan pada penelitian ini adalah *Random Forest*.
4. Evaluasi kinerja model dilakukan menggunakan metrik akurasi, presisi, *recall*, dan F1-score.
5. Penanganan outlier dilakukan menggunakan metode *Interquartile Range* (IQR) untuk mengurangi pengaruh data ekstrem terhadap kinerja model klasifikasi.
6. Metode SMOTE digunakan untuk mengatasi ketidakseimbangan data antar kelas dengan menghasilkan data sintetis pada kelas minoritas agar performa klasifikasi lebih optimal.

### 1.5 Sistematika Penulisan

Berikut merupakan sistematika penelitian yang digunakan dalam penyusunan penulisan penelitian tugas akhir

## BAB 1 PENDAHULUAN

Bab ini menjelaskan berbagai aspek mendasar yang melandasi penelitian. Latar belakang memaparkan konteks dan alasan mengapa topik ini penting untuk diteliti, termasuk masalah atau fenomena yang melatarbelakangi studi ini.

## BAB 2 TINJAUAN PUSTAKA

Bab ini menguraikan berbagai teori dasar yang menjadi landasan penelitian. Teori-teori tersebut dipilih karena relevansinya dalam menjelaskan fenomena yang diteliti, sekaligus berfungsi sebagai kerangka pemikiran untuk menganalisis data dan menjawab pertanyaan penelitian.

## BAB 3 METODE USULAN

Bab ini memaparkan langkah demi langkah proses penelitian secara lengkap dan terstruktur. Awalnya dilakukan studi literatur untuk menggali konsep-konsep dasar yang relevan dengan penelitian, sekaligus memahami landasan teoritis yang diperlukan. Setelah itu, peneliti mempersiapkan segala kebutuhan teknis dan konseptual untuk merancang sistem, termasuk menentukan spesifikasi dan alat yang akan digunakan.

## BAB 4 HASIL DAN PEMBAHASAN

Bab ini menyajikan hasil implementasi sistem klasifikasi status gizi balita. Hasil pengolahan data ditampilkan dalam bentuk tabel, grafik, dan confusion matrix untuk menunjukkan kinerja model klasifikasi. Selain itu, dibahas perbandingan hasil sebelum dan sesudah penerapan SMOTE serta pengaruh majority voting terhadap akurasi dan stabilitas hasil klasifikasi. Pembahasan dilakukan dengan mengaitkan hasil penelitian dengan teori dan penelitian terdahulu untuk menjawab rumusan masalah yang telah ditetapkan.

## **BAB 5 PENUTUP**

Bab ini berisi kesimpulan dan saran. Kesimpulan merangkum hasil utama penelitian terkait keberhasilan penerapan metode data mining dalam klasifikasi status gizi balita. Sementara itu, saran berisi rekomendasi untuk pengembangan penelitian selanjutnya, seperti penambahan variabel gizi, penggunaan algoritma lain sebagai pembanding, serta pengembangan sistem agar dapat diimplementasikan secara lebih luas di fasilitas pelayanan kesehatan.

## **DAFTAR PUSTAKA**

Bagian ini memuat seluruh referensi yang digunakan sebagai landasan teori, metode, dan pembahasan dalam penelitian, baik berupa buku, jurnal ilmiah, maupun sumber resmi lainnya yang relevan dengan topik penelitian.

## **LAMPIRAN**

Bagian ini berisi dokumen pendukung penelitian, seperti dataset yang digunakan, kode program, hasil pengujian tambahan, flowchart sistem, serta dokumentasi lain yang menunjang dan memperjelas pelaksanaan penelitian.