

BAB I

PENDAHULUAN

1.1. Latar Belakang

Periode usia di bawah lima tahun memegang peran yang sangat menentukan dalam proses tumbuh kembang anak, baik dari aspek jasmani maupun kemampuan kognitifnya. Rentang waktu yang membentang sejak masa konsepsi sampai anak berusia dua tahun ini lazim dikenal sebagai periode emas perkembangan. (Hardiani & Putri, 2024). Saat ini, kegagalan pertumbuhan yang dipicu oleh kekurangan gizi—yang umum disebut *stunting*—telah berkembang menjadi persoalan gizi berskala dunia dengan urgensi penanganan yang tinggi. (Komalasari et al., 2020). Di Indonesia, tingkat prevalensi *stunting* masih sangat mengkhawatirkan. Berdasarkan data Riset Kesehatan Dasar (2013), persentase *stunting* menyentuh 37,2%, dan meskipun terjadi tren penurunan menjadi 27,5% pada 2016, angka ini belum memenuhi batas aman yang direkomendasikan oleh World Health Organization (WHO), yakni di bawah 20%. Merespons hal tersebut, pemerintah memprioritaskan penanganan *stunting* dengan menetapkan target penurunan prevalensi hingga 14% di tahun 2024 (Fitri et al., 2022).

Tingginya kasus *stunting* mengindikasikan adanya kendala mendasar, khususnya minimnya pemahaman masyarakat mengenai pemicu utama kondisi tersebut (Finda & Utomo, 2024). Hambatan pada tumbuh kembang anak ditentukan oleh sejumlah faktor yang saling berkaitan, di antaranya mutu layanan kesehatan ibu beserta pemeriksaan kehamilan, keadaan sanitasi dan lingkungan tempat tinggal, pengelolaan pemberian air susu ibu, serta kecukupan asupan gizi anak. (Ashuri et al., 2024). Pengaruh *stunting* tidak terbatas pada bentuk tubuh semata, melainkan turut mengganggu perkembangan keterampilan motorik dasar serta memperlambat fungsi kognitif yang ditandai dengan menurunnya tingkat kecerdasan intelektual (IQ). Dengan demikian, *stunting* memberikan konsekuensi yang besar bagi proses tumbuh kembang anak, baik pada ranah jasmani maupun kognitif. (Subadi & Kusri, 2024).

Secara operasional di tingkat Puskesmas, identifikasi status gizi balita umumnya mengandalkan pengukuran antropometri manual yang mencakup parameter usia, bobot, dan tinggi tubuh. Hasil ukur ini dikonversi menjadi z-score sesuai standar

Kementerian Kesehatan RI, di mana ambang batas <-3 standar deviasi mengindikasikan stunting. Sayangnya, evaluasi manual ini memakan waktu yang lama dan memiliki risiko human error yang tinggi. Oleh karena itu, diperlukan sebuah mekanisme klasifikasi cerdas yang mampu memprediksi status stunting balita dengan tingkat akurasi dan kecepatan yang optimal (Lonang & Normawati, 2022).

Implementasi *machine learning* di sektor medis menawarkan solusi inovatif untuk menjembatani keterbatasan tersebut. Melalui pendekatan komputasi ini, dataset riwayat kesehatan dapat dianalisis secara komprehensif guna mendeteksi pola laten yang menjadi faktor risiko stunting (Dwinanto et al., 2024). Klasifikasi dapat dipahami sebagai langkah membangun suatu model atau fungsi yang berperan menggambarkan sekaligus memisahkan kategori data tertentu, dengan sasaran utama menentukan label kelas dari sebuah data. (Hardiani & Putri, 2024). Proses ini berlangsung melalui dua fase, yakni fase pembelajaran yang menelaah data latih melalui algoritma klasifikasi, dan fase pengujian yang mengukur keakuratan hasil klasifikasi menggunakan data uji. (Lonang & Normawati, 2022).

Studi ini mengajukan pemanfaatan algoritma K-Nearest Neighbors (KNN), yaitu metode klasifikasi yang tergolong dalam pendekatan supervised learning. Algoritma tersebut menetapkan kelas pada data baru dengan mengukur kedekatan jarak terhadap sejumlah K data tetangga terdekat yang berasal dari himpunan data pelatihan. (Kamila et al., 2024). Keunggulan utama Algoritma KNN terletak pada kapabilitasnya memproses himpunan data berukuran besar dan kemudahan implementasinya di berbagai sektor (Dwinanto et al., 2024). Namun demikian, KNN juga memiliki kekurangan, khususnya pada aspek recall yang dapat mempengaruhi kemampuan model dalam mendeteksi seluruh data positif secara optimal. Oleh karena itu, diperlukan strategi optimasi pada algoritma KNN, seperti penentuan nilai K yang optimal dan penggunaan metode evaluasi yang tepat, guna meminimalkan kelemahan dari metode K-NN tersebut dan dapat meningkatkan akurasi dalam mengklasifikasikan *stunting*.

Telaah pustaka terdahulu yang membandingkan KNN dengan Naïve Bayes menggunakan dataset posyandu Desa Pasirjengkol memperlihatkan keunggulan

kinerja komputasi pada KNN. Pada konfigurasi parameter $K=1$, KNN memperoleh akurasi 97,90% sehingga jauh melampaui Naïve Bayes yang hanya menyentuh 54,39. (Meriyana et al., 2025). Berdasarkan penelitian terkait sebelumnya, kajian ini memilih metode *K-Nearest Neighbor* (KNN) atas dasar keluwesannya dalam mengolah data yang berpola non-linear sekaligus kemampuannya memberikan hasil klasifikasi yang lebih akurat pada data berstruktur rumit. KNN tergolong algoritma *machine learning* yang andal untuk persoalan klasifikasi *stunting*, sebagaimana ditunjukkan dalam berbagai penelitian terdahulu yang melaporkan kinerja dan tingkat akurasi yang tinggi.

Pada penelitian (Bimawan et al., 2024), melakukan komparasi beberapa algoritma yakni *Random Forest*, *K – Nearest Neighbor*, *Decision Tree*, dan *XGBoost* untuk mendeteksi penyakit *stunting* balita dan menyimpulkan bahwa KNN memberikan performa paling unggul dengan akurasi 99,90% serta F1-score 99,87%. Temuan ini menegaskan bahwa meskipun KNN dikenal sederhana dan praktis untuk diterapkan, hasil yang dihasilkannya tetap memiliki ketepatan yang tinggi. (Bimawan et al., 2024). Selain itu, penelitian oleh Eko Abdul Goffar dkk. mengenai pemodelan prediksi *stunting* pada balita menggunakan pendekatan *machine learning* dan implementasi model pada aplikasi *mobile* juga menerapkan algoritma KNN. Penelitian ini juga memperoleh hasil baik pada KNN dengan akurasi yang tinggi sebesar 84%, nilai *precision* sebesar 82%, *recall* sebesar 84%, serta nilai *F1 – score* sebesar 82% (Goffar et al., 2025). Berdasarkan sejumlah riset terdahulu, KNN terbukti sesuai diterapkan pada data yang polanya dapat dipetakan dengan baik melalui *matriks* jarak tertentu, kendati efektivitasnya sangat ditentukan oleh penetapan nilai k dan jenis perhitungan jarak antar sampel (Goffar et al., 2025).

Berdasarkan urgensi pada ranah klinis serta keunggulan metrik evaluasi dari algoritma tersebut, penelitian ini diusulkan dengan judul "Klasifikasi Stunting Menggunakan Metode K-Nearest Neighbor (KNN)". Penerapan model *machine learning* ini diproyeksikan dapat mengotomasi tahap penyaringan (*screening*) terhadap data rekam medis, sehingga berperan sebagai *decision support system* yang membantu fasilitas kesehatan, pemerintah, maupun institusi terkait dalam menyusun intervensi gizi secara lebih tepat sasaran dan terukur.

1.2. Rumusan Masalah

1.2.1. Permasalahan

Mengacu pada data yang dipublikasikan WHO, Indonesia tercatat menempati posisi ketiga sebagai negara dengan angka prevalensi *stunting* paling tinggi di kawasan Asia Tenggara (SEAR). Temuan Pemantauan Status Gizi (PSG) pada 2017 semakin menegaskan pentingnya persoalan tersebut, di mana rasio balita *stunting* nasional menyentuh persentase 29,6% (Komalasari et al., 2020). Mengingat signifikansi dari mitigasi preventif, ketersediaan sistem penapisan (*screening*) dini yang presisi menjadi sangat kritikal. Oleh karena itu, implementasi model klasifikasi *supervised learning* berbasis algoritma *K-Nearest Neighbor* diusulkan sebagai solusi komputasional guna menghadirkan pendekatan yang sistematis, terukur, dan efisien dalam memetakan label status gizi.

1.2.2. Metode Usulan

Pendekatan komputasi yang diusulkan pada penelitian ini berpusat pada penerapan algoritma klasifikasi *K-Nearest Neighbor* (KNN). Pemodelan algoritma ini dibangun untuk mengekstraksi dan mengklasifikasikan *dataset* rekam medis antropometri balita historis (periode observasi 2023–2024) ke dalam representasi kelas status *stunting* secara terotomatisasi.

1.2.3. Pertanyaan Penelitian

Berdasarkan rumusan masalah di atas, penelitian ini bertujuan untuk menganalisis dan mengevaluasi performa model *K-Nearest Neighbor* (KNN) dalam melakukan klasifikasi *stunting* pada *dataset* rekam medis balita di Puskesmas Tegalsiwalan Probolinggo dengan membandingkan beberapa nilai *K* pada dua skenario pembagian data, yaitu 80:20 dan 70:30. Fokus utama penelitian ini adalah bagaimana performa algoritma *K-Nearest Neighbor* (KNN) dalam mengklasifikasikan status *stunting* pada *dataset* rekam medis balita di Puskesmas Tegalsiwalan Probolinggo berdasarkan perbandingan beberapa nilai *K* pada skenario pembagian data 80:20 dan 70:30?

1.3. Tujuan dan Manfaat

1.3.1. Tujuan Penelitian

Berdasarkan latar belakang dan rumusan masalah yang telah diuraikan diatas, maka tujuan dari penelitian ini adalah untuk menganalisis model *K-Nearest Neighbor* dalam mengklasifikasikan kasus *stunting* pada balita di Puskesmas Tegalsiwalan Probolinggo berdasarkan data yang digunakan.

1.3.2. Manfaat Penelitian

Adanya *prototype* perangkat lunak dapat membantu tenaga kesehatan dalam penanganan *stunting*. Sehingga dapat dilakukan intervensi awal untuk masyarakat dalam mencegah atau mengurangi risiko *stunting*.

1.4. Batasan Masalah

Batasan masalah yang diperlukan agar penelitian tetap fokus dan terarah pada materi yang diteliti, batasan tersebut antara lain:

1. Dataset merupakan data rekap pasien *stunting* pada tahun 2023 - 2024 yang diperoleh dari Puskesmas Tegal Siwalan Probolinggo.
2. Jumlah keseluruhan data yang dianalisis sebanyak 2.098 record.
3. Dataset yang digunakan dalam penelitian ini menggunakan penelitian sebelumnya.
4. Data berupa rekapitulasi kondisi *stunting* balita yang ditinjau dari sejumlah faktor penentu, meliputi jenis kelamin, usia, tinggi badan, dan berat badan, dengan hasil diagnosa sebagai variabel target.
5. Pembagian dataset dilakukan menjadi dua, yaitu data uji dan data latih dengan proporsi 20:80 dan 70:30.

1.5. Sistematika Proposal

Sistematika proposal ini disusun dalam lima bab sebagai berikut.

BAB I PENDAHULUAN

Bagian ini memberikan ikhtisar singkat tentang penelitian yang akan dilakukan. Ini mencakup latar belakang, rumusan masalah, batasan masalah, tujuan dan manfaat dilakukannya penelitian, dan sistematis penyusunan laporan.

BAB II TINJAUAN PUSTAKA

Bagian ini merincikan landasan teori yang menjadi dasar dan panduan untuk penelitian klasifikasi *stunting* menggunakan metode *K - Nearest Neighbors* untuk mengetahui akurasi model dalam bidang klasifikasi medis khususnya pada kasus *stunting*.

BAB III METODOLOGI PENELITIAN

Bagian ini memberikan penjelasan mengenai arsitektur sistem, *preprocessing* data, dan tahap atau sistematisasi dari metode yang diajukan serta skenario pengujian.

BAB IV HASIL DAN PEMBAHASAN

Bagian ini menyajikan hasil implementasi metode K-Nearest Neighbors (KNN) dalam klasifikasi *stunting* pada balita. Tahapan yang dibahas meliputi proses prapemrosesan data, pembagian dataset menggunakan metode *stratified split*, normalisasi data, pelatihan model, serta pengujian menggunakan beberapa nilai K. Selain itu, bab ini menyajikan analisis dan pembahasan mengenai hasil evaluasi model. Penentuan performa terbaik dari model yang dibangun tersebut didasarkan pada metrik *accuracy*, *precision*, *recall*, dan *F1-score*.

BAB V KESIMPULAN DAN SARAN

Bagian ini menyajikan kesimpulan yang diperoleh dari hasil penelitian dan pembahasan secara menyeluruh. Kesimpulan tersebut disusun untuk jawaban atas tujuan penelitian, yaitu menganalisis penerapan algoritma K-Nearest Neighbors (KNN) dalam mengklasifikasikan kondisi *stunting* pada balita berdasarkan karakteristik data yang telah dianalisis. Selain itu, bab ini juga memuat saran yang dapat digunakan sebagai bahan pertimbangan untuk pengembangan penelitian selanjutnya agar diperoleh hasil yang lebih optimal dan komprehensif.