

BAB I

PENDAHULUAN

1.1 Latar Belakang

Setiap orang tua mempunyai pola asuh tersendiri dalam mendidik anak-anak mereka. Pola asuh tersebut berpengaruh terhadap tumbuh kembang anak. Orang tua mempunyai pengaruh yang paling besar terhadap anak, setiap orang tua mempunyai pola asuh tersendiri dari segi mengasah, mengasuh dan mengasih dalam hubungan bersama anak-anak, dan ini berpengaruh terhadap pertumbuhan dan perkembangan anak[1]

Pola asuh yang kurang tepat, terutama dalam praktik pemberian makan, dapat berdampak pada ketidakseimbangan asupan gizi yang berujung pada risiko stunting. Kondisi gizi ibu juga memiliki peran penting, ibu yang mengalami kekurangan nutrisi sejak masa remaja, kehamilan, hingga menyusui cenderung melahirkan anak dengan pertumbuhan tubuh dan perkembangan otak yang tidak optimal. Selain faktor pola asuh dan gizi, penyebab lain seperti infeksi selama kehamilan, usia ibu yang terlalu muda, gangguan kesehatan mental, jarak kelahiran yang terlalu dekat, serta hipertensi turut meningkatkan risiko stunting pada anak. Rendahnya akses terhadap layanan kesehatan, sanitasi, dan air bersih juga menjadi faktor lingkungan yang memperburuk kondisi tersebut, karena dapat menghambat penyerapan nutrisi serta mengganggu pertumbuhan anak secara menyeluruh[2][3].

Stunting merupakan akibat dari kekurangan gizi jangka panjang, dan seringkali menyebabkan keterlambatan perkembangan mental, prestasi sekolah yang buruk, dan penurunan kapasitas[4]. Kondisi ini juga menjadi indikator penting untuk menilai kualitas kesehatan dan kesejahteraan anak, karena berhubungan erat dengan kemampuan belajar, produktivitas di masa depan, serta risiko penyakit kronis saat dewasa[5].

Berdasarkan data dari World Health Organization (WHO), prevalensi balita stunting secara global pada tahun 2024 mencapai sekitar 23,2% atau setara dengan 150,2 juta jiwa[6]. Sementara itu, hasil Survei Kesehatan Indonesia (SKI) 2023 menunjukkan bahwa prevalensi stunting nasional berada pada angka 21,5%, mengalami penurunan dalam satu dekade terakhir (2013–2023). Namun demikian, capaian tersebut masih belum memenuhi target Rencana Pembangunan Jangka Menengah Nasional (RPJMN) 2020–2024, yang menargetkan prevalensi stunting sebesar 14% pada tahun 2024. Kondisi ini menegaskan bahwa upaya percepatan penurunan angka stunting masih memerlukan strategi yang lebih efektif dan berbasis data[7].

Dalam konteks permasalahan stunting, ibu hamil dengan konsumsi asupan gizi yang rendah dan mengalami penyakit infeksi akan melahirkan bayi dengan Berat Badan Lahir Rendah (BBLR) dan atau panjang badan bayi di bawah standar. Asupan gizi yang baik tidak hanya ditentukan oleh ketersediaan pangan ditingkat rumah tangga tetapi juga dipengaruhi oleh pola asuh seperti pemberian kolostrum (ASI yang pertama kali keluar), inisiasi menyusui dini (IMD), pemberian ASI eksklusif, dan pemberian makanan pendamping ASI (MP-ASI) secara tepat[8].

Tingginya prevalensi stunting tersebut memerlukan perhatian lebih terhadap penguatan strategi deteksi dan pencegahan. Untuk itu, dibutuhkan metode analisis yang lebih adaptif dan berbasis data guna memperkuat akurasi deteksi kasus stunting sejak dini. Dalam menangani permasalahan evaluasi stunting pada anak-anak, penerapan teknik inovatif berbasis *machine learning* (ML) dinilai mampu meningkatkan efektivitas dalam memprediksi kemungkinan anak mengalami stunting, sehingga langkah-langkah preventif dapat dilakukan secara lebih tepat dan proaktif [9].

Penelitian ini menerapkan algoritma *machine learning* *K-Nearest Neighbor* (KNN) dan *Random Forest* karena kedua algoritma tersebut memiliki karakteristik yang berbeda dalam proses klasifikasi dan berpotensi menghasilkan kinerja yang berbeda pada data yang sama. KNN merupakan metode klasifikasi berbasis jarak yang bekerja dengan mengelompokkan data berdasarkan kedekatan antar data.

Keunggulan KNN terletak pada kesederhanaan konsep dan kemudahan implementasinya, namun kinerjanya sangat dipengaruhi oleh pemilihan nilai k dan skala fitur.

Sementara itu, *Random Forest* merupakan metode *ensemble* berbasis pohon keputusan yang mampu menangani data dengan jumlah fitur dan kompleksitas yang tinggi serta hubungan *non-linear* antar variabel. *Random Forest* cenderung memiliki performa yang lebih stabil dan tidak mudah mengalami *overfitting*, namun memiliki sejumlah *hyperparameter* yang perlu diatur secara optimal dan membutuhkan sumber daya komputasi yang lebih besar. Perbedaan pendekatan dan tingkat kompleksitas antara KNN dan *Random Forest* menunjukkan bahwa masing-masing algoritma memiliki kelebihan dan keterbatasan yang perlu dievaluasi secara komparatif.

Meskipun penerapan *machine learning* dalam deteksi stunting telah menunjukkan potensi yang signifikan, namun eksplorasinya masih terbatas dalam mengoptimalkan pengolahan data guna mencapai tingkat akurasi yang maksimal. Oleh karena itu, penelitian ini menekankan penerapan *Grid Search* sebagai metode penyetelan *hyperparameter* guna meningkatkan performa model secara optimal. Selain itu melakukan perbandingan antara dua algoritma *machine learning* untuk menentukan model dengan kinerja terbaik dalam mengidentifikasi kasus stunting pada balita.

Perbedaan karakteristik dan sensitivitas terhadap *hyperparameter* pada kedua algoritma tersebut menjadi dasar utama dalam melakukan perbandingan kinerja. Dalam konteks deteksi stunting yang melibatkan data dengan banyak variabel dan karakteristik yang beragam, tidak terdapat jaminan bahwa satu algoritma akan selalu memberikan performa terbaik. Oleh karena itu, penelitian ini menekankan penerapan *hyperparameter tuning* menggunakan metode *Grid Search* untuk memastikan setiap algoritma dievaluasi pada kondisi terbaiknya. Melalui pendekatan ini, perbandingan kinerja yang dilakukan bersifat objektif dan diharapkan dapat menentukan algoritma yang paling optimal dalam mendeteksi status stunting pada balita.

1.2 Rumusan Masalah

1. Bagaimana performa algoritma *Random Forest* dalam memprediksi *stunting* pada balita?
2. Bagaimana performa algoritma *K-Nearest Neighbor* (KNN) dalam memprediksi *stunting* pada balita?
3. Algoritma mana yang memberikan kinerja terbaik dalam klasifikasi *stunting* berdasarkan metrik evaluasi *accuracy*, *precision*, *recall*, dan *f1-score*, dengan fokus utama pada metrik *recall*?

1.3 Batasan Masalah

1. Penelitian ini hanya menggunakan dataset mengenai data *stunting* dari platform *kaggle* yang menyediakan data secara online dari tahun 2023 - 2025
2. Penelitian ini hanya menguji algoritma *machine learning*, yaitu *Random Forest* dan *K-Nearest Neighbor* (KNN) dengan penerapan teknik *Grid Search* sebagai metode *hyperparameter tuning*
3. Penelitian ini menggunakan enam fitur utama sebagai variabel pemodelan, yaitu jenis kelamin (*Gender*), berat lahir (*Birth Weight*), panjang lahir (*Birth Length*), berat badan saat ini (*Body Weight*), panjang badan saat ini (*Body Length*), serta status pemberian ASI (*Breastfeeding*).

1.4 Tujuan Penelitian

Tujuan dari penelitian ini adalah

1. Mengimplementasikan dan membandingkan algoritma *K-Nearest Neighbor* (KNN) dan *Random Forest* untuk menentukan algoritma dengan kinerja terbaik dalam mendeteksi *stunting* pada balita berdasarkan metrik evaluasi, dengan fokus utama pada *recall*..
2. Untuk menganalisis pengaruh penyetelan parameter secara otomatis menggunakan *hyperparameter tuning* dengan teknik *grid search* terhadap peningkatan kinerja algoritma dalam memprediksi penyakit *stunting*.

1.5 Manfaat Penelitian

Manfaat dari penelitian ini adalah

1. Untuk mengklasifikasikan kasus *stunting* dengan fokus pada pencegahan dan penanganan, sehingga sumber daya kesehatan dapat dialokasikan dengan lebih tepat dan efisien.
2. Untuk mengembangkan sistem diagnosa berbasis algoritma *machine learning* yang canggih dan efektif dalam memprediksi status *stunting*.
3. Untuk mendukung pengambilan kebijakan dalam upaya pencegahan dan pengendalian *stunting*. Dengan mengidentifikasi area prioritas untuk intervensi dan alokasi sumber daya.

1.6 Sistematika Penulisan

Struktur penulisan skripsi ini terdiri atas lima bab, yang masing-masing bab diuraikan sebagai berikut:

BAB I PENDAHULUAN

Bab ini berisi latar belakang masalah, rumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian, serta sistematika penulisan skripsi.

BAB II TINJAUAN PUSTAKA

Bab ini menyajikan studi literatur serta dasar-dasar teori yang mendukung penelitian, yang mencakup pemaparan hasil penelitian terdahulu dan uraian konsep yang digunakan. Ruang lingkup teori meliputi pembahasan mengenai *stunting*, *data mining*, *machine learning*, *hyperparameter tuning*, *grid search*, *cross validation*, *label encoding*, serta normalisasi data. Selain itu, bab ini juga menguraikan konsep klasifikasi, algoritma *K-Nearest Neighbor (KNN)* dan *Random Forest*, serta metode evaluasi performa model yang digunakan dalam penelitian.

BAB III METODE PENELITIAN

Pada bab ini menjelaskan alur penelitian, metode yang digunakan, serta alat dan bahan yang mendukung proses penelitian. Secara umum, penelitian ini dilakukan melalui serangkaian tahapan mulai dari pengumpulan data, *preprocessing data*, pemodelan algoritma dan proses *tuning*, hingga evaluasi menggunakan metrik seperti *accuracy*, *precision*, *recall* dan *f1-score*.

BAB IV HASIL DAN PEMBAHASAN

Bab ini menyajikan hasil penelitian yang dihasilkan melalui proses analisis dan pengujian model, meliputi penjelasan mengenai dataset, tahapan *preprocessing data*, pemodelan algoritma, serta evaluasi kinerja model.

BAB V PENUTUP

Bab ini memuat kesimpulan yang merangkum temuan penting dari hasil analisis penelitian. Selain itu, bab ini juga memberikan saran untuk penelitian berikutnya, termasuk peluang untuk mengkaji algoritma lain atau menggunakan variasi dataset yang lebih luas.