

BAB I

PENDAHULUAN

1.1. Latar Belakang

Kondisi gizi yang dikenal sebagai *stunting* diidentifikasi melalui indeks PB/U atau TB/U, yang menilai status gizi anak menggunakan pengukuran *antropometri*. Ketika skor Z berada antara -2 dan -3 SD, hal itu dianggap sebagai *stunting*; ketika skor Z berada di bawah -3 SD, hal itu dianggap sebagai *stunting* berat [1]. *stunting* pada balita telah menjadi permasalahan serius di Indonesia, dengan prevalensi dan dampak yang signifikan terhadap pertumbuhan fisik dan kognitif anak-anak [2]. *World Health Organization* (WHO) mengidentifikasi bahwa Indonesia memiliki peringkat yang tinggi dalam tingkat prevalensi *stunting*, menempatkannya sebagai salah satu negara dengan masalah *stunting* yang signifikan di kawasan Asia [3].

Penelitian sebelumnya telah mencoba mempelajari faktor apa saja yang berpengaruh terhadap kejadian *stunting*. Faktor-faktor seperti status gizi ibu, pola asuh, ketersediaan pangan di rumah tangga, dan akses terhadap pelayanan kesehatan telah diidentifikasi sebagai penyebab utama *stunting* [2]. Pemerintah Indonesia telah melaksanakan Studi Status Gizi Indonesia sejak tahun 2021 dan 2022 sesuai dengan upaya tersebut. Studi ini mencakup survei status gizi sampel balita di seluruh Indonesia, termasuk Kota dan Kabupaten Bekasi. Dengan angka prevalensi *stunting* sebesar 13,8 persen pada tahun 2021, wilayah Bekasi menempati peringkat kedua terendah di Jawa Barat; namun, pada tahun 2022, prevalensi *stunting* di wilayah tersebut turun secara signifikan menjadi 6 persen, menunjukkan efektivitas inisiatif pencegahan *stunting* lokal. Selain itu, data tahun 2022 menunjukkan bahwa prevalensi berat badan rendah (*underweight*) sebesar 4,8 persen, atau 6.374 anak, dan prevalensi *wasting* (berat badan rendah) sebesar 3,9 persen, atau 5.145 anak. Meskipun demikian, masih ada tantangan besar yang perlu diatasi untuk mencapai target nasional yang ditetapkan [4].

Penelitian ini bertujuan untuk mengembangkan model prediksi pertumbuhan *stunting* pada balita menggunakan teknik *machine learning*, khususnya *Support Vector Machine (SVM)* dan *Gradient Boosting*. *Support Vector*

Machine (SVM) merupakan teknik *machine learning* yang sangat efektif dalam tugas klasifikasi dan regresi, namun *SVM* menunjukkan *recall* yang tinggi dan presisi yang rendah [5][6], di sisi lain *Gradient Boosting* merupakan teknik yang menunjukkan akurasi yang unggul dibandingkan dengan algoritma lain dalam tugas klasifikasi [7]. Pemilihan kedua teknik ini bertujuan untuk membandingkan performa algoritma klasifikasi yang memiliki karakteristik berbeda dalam memprediksi kejadian *stunting* pada balita.

Penelitian ini akan menggunakan data BBLR (Berat Badan Lahir Rendah) dan PBLR (Panjang Badan Lahir Rendah) di Kabupaten Bekasi sebagai studi kasus. Kabupaten Bekasi dipilih karena memiliki prevalensi *stunting* yang cukup tinggi dan merupakan representasi dari masalah *stunting* di daerah perkotaan di Indonesia. Berdasarkan eksplorasi awal terhadap *dataset* tersebut, ditemukan adanya ketidakseimbangan distribusi kelas antara balita yang mengalami *stunting* dan yang tidak. Dengan memanfaatkan data yang tersedia, diharapkan penelitian ini dapat memberikan wawasan yang lebih dalam tentang faktor-faktor risiko yang berkontribusi terhadap *stunting* pada anak-anak di wilayah tersebut.

Penelitian ini akan menggabungkan pemahaman yang mendalam tentang pentingnya *preprocessing* data dalam hal kesehatan anak dan berbagai masalah yang terkait dengannya. Penelitian ini akan memprioritaskan keakuratan prediksi sebagai pijakan untuk intervensi dini yang efektif dalam menghadapi *stunting*, dengan fokus pada mengatasi kesalahan yang berpotensi fatal akibat data yang tidak terproses dengan baik. Untuk mengatasi masalah ketidakseimbangan kelas dalam data, penelitian ini menerapkan pendekatan *SMOTE-ENN* (*Synthetic Minority Over-sampling Technique + Edited nearest neighbor*). Metode *SMOTE* secara umum efektif dalam meningkatkan performa model *machine learning* pada *dataset* dengan distribusi kelas yang tidak seimbang [8]. Di sisi lain, *SMOTE-ENN* dapat menghasilkan data pelatihan yang lebih bersih dan seimbang sambil juga secara signifikan meningkatkan kemampuan model untuk generalisasi pada *dataset* yang sangat tidak seimbang [9]. Penelitian diharapkan menggunakan metode ini untuk menghasilkan model prediksi yang lebih akurat, yang dapat digunakan sebagai dasar untuk mendukung upaya yang lebih tepat sasaran untuk menangani *stunting*.

1.2. Perumusan Masalah

Rumusan masalah dari penelitian ini adalah:

1. Bagaimana kinerja algoritma *Support Vector Machine (SVM)* dalam memprediksi status *stunting* pada balita?
2. Bagaimana kinerja algoritma *Gradient Boosting* dalam memprediksi status *stunting* pada balita?
3. Bagaimana perbandingan performa prediksi antara algoritma *Support Vector Machine (SVM)* dan *Gradient Boosting* berdasarkan metrik evaluasi akurasi, presisi, *recall*, *F1-score*, dan *AUC ROC*?

1.3. Tujuan

Tujuan dari penelitian ini adalah:

1. Mengevaluasi kinerja algoritma *Support Vector Machine (SVM)* dalam memprediksi status *stunting* pada balita.
2. Mengevaluasi kinerja algoritma *Gradient Boosting* dalam memprediksi status *stunting* pada balita.
3. Membandingkan performa prediksi kedua algoritma berdasarkan metrik akurasi, presisi, *recall*, *F1-score*, dan *AUC ROC*.

1.4. Batasan Masalah

Berikut adalah batasan yang terdapat pada penelitian ini:

1. Penelitian ini hanya mengambil data balita dari Kabupaten Bekasi, tidak mencakup wilayah lain di Indonesia.
2. Model prediksi yang digunakan dalam penelitian ini hanya terbatas pada dua metode *machine learning*, yaitu *Support Vector Machine (SVM)* dan *Gradient Boosting*.
3. Evaluasi kinerja model dilakukan menggunakan metrik akurasi, presisi, *recall*, *F1-score*, serta *ROC-AUC* untuk memberikan gambaran lebih menyeluruh terhadap performa klasifikasi.