

BAB 1. PENDAHULUAN

1.1 Latar Belakang

Stunting adalah suatu permasalahan kesehatan serius dan berdampak signifikan terhadap pertumbuhan serta perkembangan anak di seluruh dunia. Di Indonesia, *stunting* masih menjadi tantangan serius dalam bidang kesehatan anak karena tingkat prevalensinya yang tinggi. Menurut Kurniasih, prevalensi *stunting* di Indonesia mencapai 21,6% pada pertengahan tahun 2023. Prevalensi ini harus menjadi perhatian serius, sehingga di tahun 2024 ini pemerintah memiliki target untuk menurunkan prevalensi *stunting* hingga 14% (emedia.dpr.go.id, 2024).

Stunting adalah suatu kondisi dimana terdapat kegagalan pada pertumbuhan serta perkembangan yang diakibatkan oleh kekurangan gizi pada 1000 Hari Pertama Kehidupan (1000 HPK) (Achadi et al., 2020). Faktor utama yang menyebabkan kondisi ini adalah asupan gizi yang tidak memadai pada ibu selama masa kehamilan dan tidak terpenuhinya kebutuhan nutrisi anak. *Stunting* dapat menyebabkan anak gagal tumbuh karena kurang asupan gizi serta memiliki dampak negatif dalam jangka panjang seperti kesulitan belajar, terkena penyakit jantung, serta daya tahan tubuh yang lemah (Lim et al., 2024).

Stunting bukan hanya menjadi masalah di Indonesia, tetapi juga merupakan isu kesehatan global. Berdasarkan perkiraan WHO tahun 2020, sekitar 22% anak-anak di seluruh dunia, atau setara dengan 149,2 juta, mengalami *stunting*. Di Indonesia sendiri, menurut *Asian Development Bank*, sebanyak 31,8% balita mengalami *stunting* pada tahun 2022, menjadikan Indonesia berada di peringkat ke-10 di Asia Tenggara. Meski begitu, data dari Kementerian Kesehatan pada tahun 2023 menunjukkan adanya penurunan prevalensi *stunting* menjadi 21,6% (Deviana, 2023).

Puskesmas Maja yang terletak di Kabupaten Lebak, Provinsi Banten, mencatat tingkat prevalensi *stunting* pada anak dan balita yang tergolong cukup tinggi. Berdasarkan data Survei Status Gizi Indonesia (SSGI) tahun 2022 dari Kementerian Kesehatan Republik Indonesia, sekitar 27,5% balita di wilayah tersebut mengalami *stunting*. Persentase ini berada di atas rata-rata nasional yang hanya sebesar 21%.

Kondisi ini menandakan perlunya upaya intervensi yang lebih efektif dan terarah untuk menurunkan kasus *stunting* di daerah tersebut (Lantara, 2023).

Pemetaan wilayah rawan *stunting* sangat penting untuk memprioritaskan intervensi kesehatan yang efektif dan tepat sasaran. Deteksi daerah rawan *stunting* memungkinkan pihak-pihak terkait, seperti pemerintah daerah dan Puskesmas, untuk memfokuskan pemanfaatan sumber daya yang terbatas ke area dengan prevalensi tinggi. Selain itu, identifikasi daerah berisiko tinggi dapat membantu meningkatkan upaya pencegahan dan penanganan *stunting* secara menyeluruh (Fathurrahman et al., 2022).

Dari permasalahan tersebut diperlukan adanya identifikasi serta diagnosa dini untuk menekan laju peningkatan *stunting*. Salah satu metode yang bisa diterapkan untuk mendiagnosis *stunting* adalah dengan memanfaatkan algoritma *Random Forest*. Algoritma ini, yang merupakan bagian dari teknik *machine learning*, telah banyak digunakan dalam berbagai aplikasi prediktif di bidang medis. *Random Forest* sendiri merupakan metode yang berbasis pada teknik klasifikasi dan regresi. Dengan menggabungkan beberapa pohon keputusan, algoritma ini mampu meminimalkan risiko *overfitting* serta efektif dalam menangani data dengan distribusi kelas yang tidak seimbang (Reza & Rohman, 2024).

Data kesehatan kerap kali menghadapi masalah ketidakseimbangan kelas, di mana kasus positif (*stunting*) hanya merupakan sebagian kecil dari keseluruhan data. Untuk mengatasi permasalahan ini, salah satu pendekatan yang diterapkan adalah teknik *oversampling* seperti SMOTE (*Synthetic Minority Over-sampling Technique*). SMOTE merupakan metode *oversampling* yang bekerja dengan menghasilkan data sintetis untuk menambah jumlah sampel pada kelas minoritas. Data sintetis tersebut dibuat berdasarkan kemiripan dengan data asli dari kelas minoritas (Sutoyo & Fadlurrahman, 2020). Ketidakseimbangan distribusi kelas dapat mengakibatkan model hanya memberikan prediksi yang akurat pada kelas mayoritas, sementara performa pada kelas lainnya menurun. Oleh karena itu, teknik SMOTE digunakan untuk penanganan dengan membangkitkan data sintetis dari kelas minor (Syukron et al., 2020).

Dalam penelitian ini, selain teknik SMOTE, juga dilakukan perbandingan dengan teknik SMOTE-Tomek, yaitu kombinasi *oversampling* SMOTE dan *undersampling Tomek Links* yang dapat meningkatkan kualitas data sintetis dan menghilangkan data yang tumpang tindih (*noise*). Perbandingan ini bertujuan untuk mendapatkan model terbaik dalam mendiagnosa *stunting*.

Berdasarkan hal tersebut, penelitian ini tidak hanya bertujuan untuk mendiagnosa *stunting* pada individu, tetapi juga untuk mendeteksi daerah rawan *stunting* menggunakan algoritma *Random Forest* yang ditingkatkan dengan teknik SMOTE maupun variasinya untuk mendapatkan model terbaik dalam proses klasifikasi. Dengan ini, diharapkan hasil penelitian dapat berkontribusi pada penurunan angka *stunting* khususnya di wilayah Puskesmas Maja. Hasil pemetaan ini dapat dijadikan dasar intervensi program kesehatan yang lebih akurat dan strategis. Selain itu, penelitian ini bertujuan untuk mengetahui hasil evaluasi dari algoritma *Random Forest* dalam mendiagnosa *stunting* menggunakan teknik SMOTE serta variasinya dengan mempertimbangkan akurasi, prediksi dan matrik evaluasi lainnya, serta menganalisis kelebihan dan kekurangan metode dalam konteks diagnosa *stunting*.

1.2 Rumusan Masalah

Berdasarkan permasalahan tersebut, berikut adalah beberapa rumusan masalah yang akan diuraikan pada penelitian ini, yaitu:

1. Bagaimana implementasi algoritma *Random Forest* dengan teknik SMOTE serta variasinya dalam mendiagnosa *stunting* dan daerah rawan *stunting* pada anak dan balita di wilayah Puskesmas Maja berdasarkan hasil evaluasi dengan *confusion matrix*?
2. Bagaimana pengembangan hasil algoritma *Random Forest* dengan hasil dari model terbaik dalam bentuk GUI untuk mendiagnosa *stunting* dan daerah rawan *stunting* di Puskesmas Maja?

1.3 Batasan Masalah

Batasan masalah yang digunakan dalam penelitian ini adalah sebagai berikut:

1. Penelitian ini berfokus terhadap penggunaan algoritma *Random Forest* dalam mendiagnosa *stunting*.
2. Penelitian ini akan menggunakan teknik SMOTE serta variasinya seperti dengan mengkombinasikan teknik *undersampling* untuk menangani data *stunting* yang memiliki kelas tidak seimbang.
3. Penelitian ini akan menguji performa model dengan tiga skenario berbeda, yaitu tanpa menggunakan teknik SMOTE, menggunakan teknik SMOTE, dan menggunakan kombinasi SMOTE dengan *undersampling* untuk membandingkan pengaruh teknik penanganan data tidak seimbang terhadap hasil prediksi model.
4. Data *stunting* yang digunakan pada penelitian ini berupa data tabular dengan jumlah 16.515 data yang diambil secara langsung dari Puskesmas Maja.
5. Data yang digunakan adalah data *stunting* pada anak dan balita di Puskesmas Maja dari bulan Agustus tahun 2022 – bulan Agustus tahun 2024.
6. Penelitian ini juga membatasi deteksi daerah rawan *stunting* hanya pada wilayah Puskesmas Maja dengan 14 Desa/Kelurahan.
7. Model yang dikembangkan dalam penelitian ini hanya digunakan untuk memprediksi daerah rawan *stunting* berdasarkan data yang tersedia, sehingga tidak mencakup intervensi atau solusi kesehatan di lapangan.
8. Penelitian ini berfokus pada pengembangan dan evaluasi model, sementara prototipe *Graphical User Interface* (GUI) hanya digunakan untuk keperluan pengujian dan demonstrasi.
9. Hasil dari *Graphical User Interface* (GUI) hanya untuk memprediksi anak-anak yang tergolong *stunting* tidak sampai pada prediksi gizi anak.
10. *Graphical User Interface* (GUI) yang dihasilkan tidak menggunakan *database* untuk menyimpan *input*, melainkan hanya memanfaatkan model yang telah dimasukkan untuk melakukan proses prediksi.

1.4 Tujuan dan Manfaat Penelitian

Penelitian ini memiliki tujuan sebagai berikut:

1. Untuk dapat mengetahui hasil dari algoritma *Random Forest* serta teknik SMOTE dan variasinya dalam mendiagnosa *stunting* serta daerah rawan *stunting* dengan menggunakan *confusion matrix* yaitu nilai akurasi, presisi, *recall*, dan *F1-score*.
2. Mengembangkan antarmuka pengguna grafis (GUI) berdasarkan hasil algoritma *Random Forest* dengan pemilihan dari hasil model terbaik untuk mendiagnosa *stunting* dan identifikasi daerah rawan *stunting* di Puskesmas Maja.

Adapun manfaat dari penelitian ini yaitu:

1. Bagi Peneliti

Penelitian ini dapat menjadi sarana bagi peneliti untuk lebih memahami algoritma *Random Forest* serta penggunaannya dengan teknik SMOTE serta variasinya dan meningkatkan kesadaran dalam memahami bahaya *stunting* baik bagi diri sendiri maupun lingkungan sekitar.

2. Bagi Pembaca

Melalui penelitian ini, diharapkan pembaca memperoleh pemahaman yang lebih mendalam mengenai penerapan algoritma *Random Forest* serta cara penggunaan teknik SMOTE dan variasinya dalam mengatasi masalah ketidakseimbangan data terutama dalam bidang kesehatan, khususnya dalam mendiagnosa *stunting* pada anak dan balita, serta untuk meningkatkan kesadaran tentang bahaya *stunting* dan upaya pencegahannya di lingkungan sekitar.

3. Bagi Peneliti Lain

Penelitian ini memberikan kontribusi bagi pengembangan model prediksi kesehatan khususnya mengenai diagnosa *stunting* dan daerah yang rawan terkena *stunting* sehingga diharapkan dapat diterapkan di daerah lain dengan permasalahan serupa.

4. Bagi Institusi Terkait

Penelitian ini menyediakan hasil berupa diagnosa *stunting* serta daerah rawan *stunting* yang dapat membantu Puskesmas Maja dalam memfasilitasi serta mengidentifikasi daerah rawan *stunting*, sehingga dapat menjadi dasar bagi intervensi kesehatan yang lebih tepat sasaran.

1.5 Sistematika Penulisan

Berikut merupakan sistematika penulisan yang digunakan dalam penelitian ini:

BAB I

PENDAHULUAN

Bab 1 dalam penelitian ini memuat latar belakang masalah, perumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian, serta susunan sistematika penulisan.

BAB II

TINJAUAN PUSTAKA

Bab 2 pada penelitian ini berisi kajian teori yang terkait dengan penelitian serta pengembangan yang relevan, yang menjadi landasan dalam penyusunan penelitian ini.

BAB III

METODE PENELITIAN

Bab 3 pada penelitian ini berisikan tahapan-tahapan penelitian, jadwal penelitian, uraian, dan segala metode yang digunakan untuk pengembangan model algoritma dari penelitian ini.

BAB IV

HASIL DAN PEMBAHASAN

Bab 4 dalam penelitian ini menyajikan tahapan penelitian secara sistematis, serta memaparkan hasil dan analisis yang diperoleh dari implementasi algoritma *Random Forest* serta skenario teknik yang digunakan dalam mendiagnosa *stunting* hingga penelitian selesai.

BAB V

PENUTUP

Bab 5 dalam penelitian ini memuat kesimpulan yang disusun berdasarkan hasil dari penelitian. Selain itu, bab ini juga memberikan rekomendasi untuk penelitian lanjutan yang berfokus pada deteksi dan pencegahan *stunting* dengan pendekatan yang serupa.

DAFTAR PUSTAKA Pada bagian ini merupakan kumpulan referensi yang digunakan

sebagai dasar teori, metode, dan literatur pendukung lainnya dalam penyusunan penelitian. Referensi yang digunakan berasal dari buku, jurnal ilmiah, prosiding, laporan penelitian, serta sumber lain yang relevan dan kredibel.

LAMPIRAN

Pada bagian ini berisi dokumen serta dokumentasi pendukung yang relevan dengan penelitian, seperti hasil wawancara, foto kegiatan, atau dokumen lain yang digunakan dalam penyusunan penelitian. Lampiran ini bertujuan untuk menguraikan informasi tambahan yang lebih rinci dan mendukung transparansi dari hasil penelitian yang disampaikan.