SMARTICS Journal, Vol.12 No.02, 2026. p135-143
ISSN online: 2476-9754, ISSN print: 2623-0429
DOI : https://doi.org/10.21067/smartics.v12i2.14592

Prediksi Risiko Stunting pada Anak Berdasarkan Sanitasi Lingkungan Tinggal menggunakan Random Forest

Rusina Widha Febriana[a*], Rony Kriswibowo[b], Agung Budi Setyawan[c], Purwanto[d]
[a,b,c,d]Universitas Anwar Medika, Sidoarjo, Indonesia
[*]correspondence email : rusina.widha@uam.ac.id

Abstract*— Stunting is a health problem influenced by various factors, one of which is the residential environmental sanitation. While stunting has been linked to growth development and nutritional intake, environmental sanitation factors tend to receive less attention. Environmental sanitation in this study includes the availability of clean water, proper sanitation facilities, waste disposal systems, and the overall cleanliness of the living environment. This research aims to build a predictive model for stunting risk in children based on environmental sanitation variables using the Random Forest method. This method is a classification algorithm that constructs multiple decision trees to improve prediction accuracy. The dataset consists of environmental condition data and children's nutritional status, which have undergone preprocessing and variable transformation. The modeling process was conducted by splitting the data into training and testing sets, then evaluated using accuracy and F1-score metrics. The results indicate that the model can classify stunting risk with an accuracy rate of 85% and identify the most influential sanitation variables. This model is expected to support decision making tool for data driven stunting prevention efforts.

***Index Terms*— Environmental Sanitation; Random Forest; Risk Prediction; Stunting.**

Abstrak*– Stunting pada anak merupakan permasalahan kesehatan yang dipengaruhi oleh banyak faktor, salah satunya adalah kondisi sanitasi lingkungan tempat tinggal. Selama ini stunting lebih sering dikaitkan dengan aspek tumbuh kembang dan asupan gizi, sementara faktor sanitasi lingkungan cenderung kurang mendapat perhatian. Sanitasi lingkungan dalam penelitian ini mencakup ketersediaan air bersih, fasilitas jamban sehat, sistem pembuangan limbah, dan kondisi kebersihan tempat tinggal. Penelitian ini bertujuan untuk membangun model prediksi risiko stunting pada anak berdasarkan variabel sanitasi lingkungan menggunakan metode Random Forest. Metode ini merupakan algoritma klasifikasi yang digunakan dengan membangun sejumlah decision tree untuk meningkatkan akurasi prediksi. Dataset yang digunakan terdiri dari data kondisi lingkungan dan status gizi anak yang telah melalui tahap preprocessing dan transformasi variabel. Proses pemodelan dilakukan dengan membagi data menjadi data training dan testing, kemudian dievaluasi menggunakan metrik akurasi, precision, recall dan F1-score. Hasil penelitian menunjukkan bahwa model mampu mengklasifikasikan risiko stunting dengan tingkat akurasi 85%, serta mampu mengidentifikasi variabel sanitasi yang paling berpengaruh. Model ini diharapkan dapat menjadi alat bantu dalam pengambilan keputusan untuk upaya pencegahan stunting berbasis data.

***Kata Kunci*—Prediksi Risiko; Random Forest; Sanitasi Lingkungan; Stunting.**

I. PENDAHULUAN

Stunting merupakan salah satu permasalahan kesehatan yang hingga saat ini masih menjadi perhatian utama di negara berkembang termasuk Indonesia. Stunting menggambarkan kondisi gagal tumbuh akibat kekurangan gizi kronis yang terjadi dalam jangka waktu panjang dan berdampak pada perkembangan fisik serta kognitif anak. Selain faktor gizi, stunting juga dipengaruhi oleh berbagai faktor lain, seperti kondisi lingkungan, sosial ekonomi, dan kesehatan masyarakat [1].

Dalam beberapa tahun terakhir penanganan stunting tidak lagi hanya berfokus pada pemenuhan asupan gizi, tetapi mulai mempertimbangkan faktor lingkungan sebagai penentu penting. Lingkungan tempat tinggal beserta aspek sanitasi memiliki peran dalam mempengaruhi kesehatan dan status gizi anak. Sanitasi yang buruk dapat meningkatkan paparan penyakit infeksi, seperti diare dan infeksi cacing, yang dapat menghambat penyerapan nutrisi dan pertumbuhan anak [2].

Berbagai penelitian menunjukkan bahwa kondisi sanitasi lingkungan memiliki hubungan yang kuat dengan kejadian stunting. Faktor-faktor seperti kualitas air minum, ketersediaan jamban sehat, pengelolaan limbah, serta kondisi kebersihan tempat tinggal berpengaruh terhadap risiko stunting pada

anak. Anak yang tinggal di lingkungan dengan sanitasi buruk memiliki risiko stunting yang jauh lebih tinggi dibandingkan dengan anak yang hidup di lingkungan yang sehat [3].

Penelitian terbaru menunjukkan bahwa kualitas air minum yang tidak aman, penggunaan jamban yang tidak layak, serta pengelolaan limbah yang buruk, berhubungan dengan peningkatan jumlah stunting. Anak yang tinggal di rumah tangga dengan sanitasi yang tidak memenuhi standar memiliki kemungkinan lebih besar mengalami stunting dibandingkan dengan mereka yang hidup dalam kondisi sanitasi yang baik [4].

Banyak penelitian terkait stunting masih berfokus pada satu atau dua variabel secara terpisah, sehingga belum memberikan gambaran mengenai pengaruh kombinasi faktor sanitasi lingkungan terhadap risiko stunting. Pendekatan analisis yang digunakan umumnya seperti regresi statistik, memiliki keterbatasan dalam menangkap pola kompleks antar variabel [5].

Seiring dengan perkembangan teknologi, *machine learning* mulai banyak digunakan dalam bidang kesehatan untuk melakukan prediksi dan klasifikasi. Salah satu metode yang cukup populer adalah Random Forest, algoritma klasifikasi yang memiliki kemampuan menangani data kompleks dan meningkatkan akurasi prediksi melalui kombinasi beberapa *decision tree*. Metode ini dinilai mampu memberikan hasil prediksi yang lebih stabil dibandingkan metode lain.

Penelitian ini bertujuan untuk membangun model prediksi risiko stunting pada anak berdasarkan variabel sanitasi lingkungan tinggal menggunakan metode Random Forest. Metode ini mampu memberikan kontribusi dalam mengidentifikasi faktor sanitasi yang paling berpengaruh serta mendukung pengambilan keputusan berbasis data dalam upaya pencegahan stunting secara lebih efektif dan terarah.

Berdasarkan latar belakang yang telah diuraikan, maka rumusan masalah dalam penelitian ini adalah:

1. Bagaimana pengaruh variabel sanitasi lingkungan tempat tinggal terhadap risiko stunting pada anak?
2. Bagaimana membangun model prediksi risiko stunting pada anak menggunakan metode Random Forest berdasarkan variabel sanitasi lingkungan?
3. Seberapa baik kinerja model Random Forest dalam mengklasifikasikan risiko stunting yang diukur menggunakan metrik akurasi, *precision*, *recall*, dan *F1-score*?
4. Variabel sanitasi lingkungan apa saja yang paling berpengaruh dalam menentukan risiko stunting pada anak berdasarkan hasil pemodelan?

II. METODE

Penelitian ini menggunakan pendekatan kuantitatif yang menekankan pada pengolahan data numerik serta pengujian hubungan antar variabel secara terukur. Menggunakan metode Random Forest yang tidak hanya berfokus pada analisis hubungan antar variabel, tetapi juga pada kemampuan model dalam melakukan prediksi terhadap data baru [6]. Hal ini penting dalam permasalahan stunting yang memiliki banyak faktor penyebab dan pola yang tidak selalu linier, sehingga membutuhkan pendekatan analisis yang fleksibel.

Penggunaan metode Random Forest dalam bidang kesehatan berbasis data untuk prediksi memberikan peluang dalam pengambilan keputusan yang lebih cepat [7]. Model yang dihasilkan dapat digunakan sebagai alat bantu dalam mengidentifikasi kelompok risiko tinggi secara dini, sehingga pencegahan dapat dilakukan lebih tepat sasaran. Pendekatan ini tidak hanya berkontribusi pada pengembangan metode analisis, tetapi juga memiliki dampak dalam mendukung kebijakan kesehatan masyarakat.

Metode penelitian dirancang secara sistematis untuk memastikan proses pengolahan data dan pembangunan model prediksi dilakukan secara terstruktur dan menghasilkan output yang optimal. Pendekatan yang digunakan mencakup beberapa tahapan utama, mulai dari pengumpulan dan pengolahan data, pembangunan model menggunakan metode Random Forest, hingga evaluasi kinerja model. Setiap tahapan dilakukan secara berurutan untuk memastikan bahwa model yang dihasilkan tidak hanya memiliki tingkat akurasi yang baik, tetapi juga relevan terhadap faktor-faktor yang mempengaruhi risiko stunting pada anak.

A. Data

Data dalam penelitian ini merupakan data sekunder yang diperoleh dalam kurun waktu 2023 – 2025. Data ini mencakup informasi mengenai sanitasi lingkungan tempat tinggal serta status gizi anak yang digunakan untuk mengidentifikasi kondisi stunting. Penggunaan data sekunder dipilih karena memiliki cakupan yang lebih luas serta telah melalui proses pencatatan dan validasi oleh instansi terkait.

Jumlah data yang diperoleh, kemudian dilakukan proses seleksi dan preprocessing untuk menghilangkan data yang tidak lengkap atau tidak konsisten. Setelah melalui tahap pembersihan data, diperoleh sebanyak 200 data yang digunakan dalam proses pemodelan. Proses ini penting untuk memastikan bahwa data yang digunakan memiliki kualitas yang baik sehingga dapat meningkatkan akurasi model yang dihasilkan.

Data yang digunakan dalam penelitian ini dikumpulkan dalam bentuk dataset. Setiap data mewakili atribut atau variabel yang dapat diolah lebih lanjut dalam proses pemodelan. Kualitas data menjadi faktor penting, sehingga perlu dilakukan proses seleksi untuk memastikan bahwa data yang digunakan relevan, konsisten, dan tidak mengandung informasi yang bias [8].

Variabel yang digunakan dalam penelitian ini meliputi:

1. Ketersediaan air bersih
2. Kepemilikan jamban sehat
3. Sistem pembuangan limbah
4. Kondisi kebersihan lingkungan
5. Status stunting anak

Variabel sanitasi lingkungan dipilih berdasarkan indikator sanitasi sehat yang umum digunakan dalam studi kesehatan masyarakat. Kombinasi variabel yang digunakan diharapkan mampu menggambarkan kondisi sanitasi secara luas [9]. Ketersediaan air bersih menjadi indikator utama variabel independen karena berhubungan langsung dengan kebutuhan dasar rumah tangga. Kepemilikan jamban sehat dapat mencerminkan sanitasi yang baik, sistem pembuangan limbah dan kondisi kebersihan lingkungan menunjukkan perilaku pengelolaan lingkungan yang bersih, dan tidak menyebarkan infeksi penyakit.

Variabel dependen yang digunakan berupa status kondisi anak yang diklasifikasikan ke dalam kategori stunting dan tidak stunting. Variabel ini menjadi target utama dalam proses klasifikasi menggunakan metode Random Forest. Dengan adanya pemisahan yang jelas antara variabel, model yang dibangun diharapkan mampu mengelola pola hubungan antar variabel dan menghasilkan prediksi yang akurat terhadap risiko stunting pada [10].

B. Tahapan Penelitian

Tahapan penelitian dilakukan melalui beberapa proses yang saling terintegrasi. Tahap awal dimulai dengan proses data *preprocessing*, bertujuan untuk meningkatkan kualitas data sebelum digunakan dalam pemodelan. Proses ini meliputi (1) pembersihan data (*data cleaning*) untuk mengatasi nilai yang hilang (*missing values*) atau data yang tidak konsisten, (2) transformasi data melalui proses pengkodean untuk variabel kategorik, (3) normalisasi atau standarisasi data. Tahap *preprocessing* menjadi langkah penting karena kualitas data yang baik sangat mempengaruhi performa model yang dihasilkan [11].

Dataset yang telah diproses dibagi menjadi data *training* dan data *testing*. Pembagian ini bertujuan untuk mengukur kemampuan model dalam melakukan prediksi terhadap data baru [12]. Proporsi pembagian data dilakukan dengan komposisi 80% data *training* dan 20% data *testing*. Data *training* digunakan untuk membangun model, sedangkan data *testing* digunakan untuk mengevaluasi kinerja model.

Penelitian ini menggunakan metode Random Forest sebagai algoritma klasifikasi utama. Random Forest memungkinkan menghasilkan prediksi yang lebih stabil dan akurat dibandingkan metode klasifikasi tunggal, serta telah terbukti efektif dalam berbagai studi, termasuk pada bidang kesehatan [13]. Metode ini digunakan dengan membangun sejumlah *decision tree* dari subset data yang berbeda, kemudian menggabungkan hasil prediksi dari setiap pohon menggunakan mekanisme *majority voting*. Pada setiap *node* dilakukan pemilihan fitur secara acak untuk meningkatkan keberagaman model dan mengurangi risiko *overfitting*.

Random Forest bekerja dengan cara:

1. Membuat sejumlah pohon keputusan (*decision tree*) dari subset data. Setiap pohon keputusan mengikuti aturan pemisahan berdasarkan entropy menggunakan persamaan (1).

$$H(S) = \sum_{i=1}^{c} pi \log 2pi \qquad (1)$$

   S = himpunan data
   c = jumlah kelas dalam dataset
   pi = probabilitas suatu kelas i

2. Melakukan pemilihan fitur secara acak pada setiap node menggunakan *gini impurity* dengan persamaan (2). Semakin rendah nilai *gini*, maka data dalam *node* semakin murni (tidak heterogen).

$$Gini(S) = 1 - \sum_{i=1}^{c} pi2 \quad (2)$$

3. Menggabungkan hasil prediksi dari seluruh pohon menggunakan mekanisme *majority voting* (3).

$$\gamma = \text{mod}\, e\{T1(x), T2(x), ..., TN(x)\} \quad (3)$$

Keterangan:
γ = hasil akhir prediksi
$T1(x)$ = prediksi dari pohon ke-i

Kinerja model yang dihasilkan dievaluasi menggunakan beberapa metrik evaluasi, yaitu akurasi pada persamaan, *precision*, *recall,* dan *F1-score* ditunjukkan pada persamaan berikut.

$$accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (4)$$

$$precision = -\frac{TP}{TP + FP} \quad (5)$$

$$recall = \frac{TP}{TP + FN} \quad (6)$$

$$F1 = 2 \times \frac{precision \times recall}{precision + recall} \quad (7)$$

TP (*True Positive*) = prediksi positif dan benar (stunting)
TN (*True Negative*) = prediksi negatif dan benar (tidak stunting)
FP (*False Positive*) = prediksi positif, tetapi salah
FN (*False Negative*) = prediksi negatif, tetapi salah

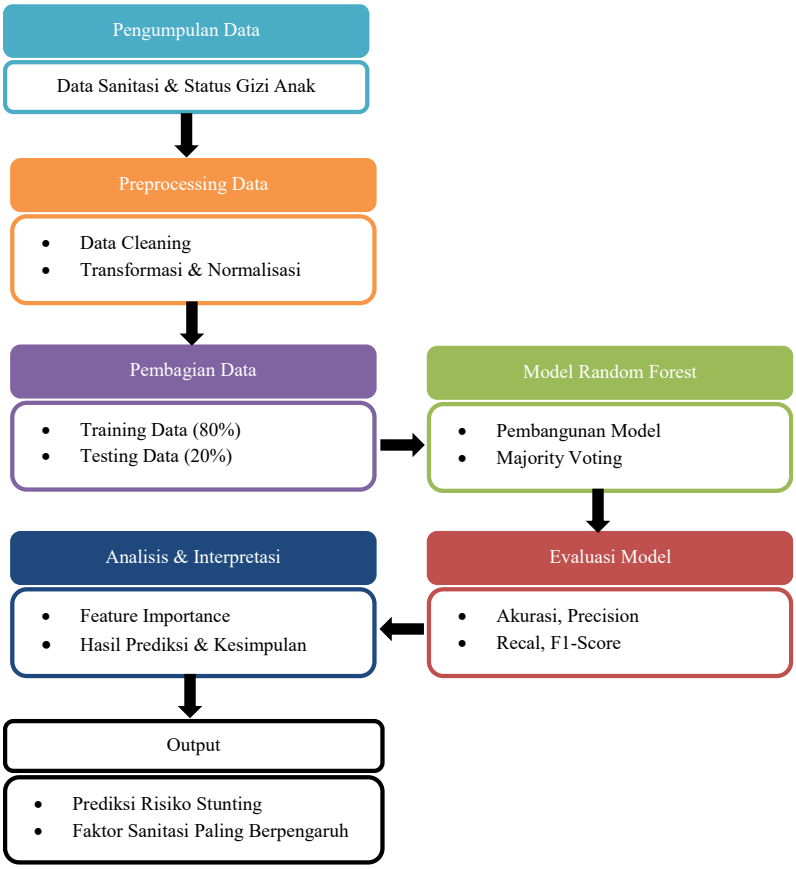
Akurasi digunakan untuk mengukur tingkat ketepatan prediksi secara keseluruhan, *precision* dan *recall* digunakan untuk mengevaluasi performa model dalam mengklasifikasikan kelas. *F1-score* digunakan sebagai ukuran keseimbangan antara *precision* dan *recall* terhadap performa model, khususnya pada dataset yang memiliki distribusi kelas yang tidak seimbang [14].

Keunggulan Random Forest dalam menyediakan informasi terkait kontribusi masing-masing variabel menjadikannya sebagai metode yang akurat dalam prediksi, tetapi juga informatif dalam analisis hasil [15]. Analisis tingkat kepentingan variabel (*feature importance*) yang dihasilkan oleh Random Forest bertujuan untuk mengidentifikasi variabel sanitasi lingkungan yang memiliki kontribusi paling besar terhadap risiko stunting pada anak.

C. Alur Penelitian

Prediksi merupakan sebuah proses untuk menemukan pola tertentu dari suatu data. Pola tersebut dapat diketahui dari berbagai variabel yang ada pada data. Ketika sebuah pola sudah didapat maka bisa digunakan untuk memprediksi kejadian yang akan datang, yang belum diketahui nilai ataupun jenisnya. Fungsi prediksi juga bisa digunakan untuk memprediksi sebuah variabel tertentu yang tidak ada dalam suatu data [16].

Untuk memberikan gambaran yang lebih sistematis mengenai tahapan penelitian yang dilakukan diperlukan alur penelitian yang terstruktur. Alur ini menggambarkan proses mulai dari pengumpulan data hingga menghasilkan output berupa prediksi risiko stunting serta identifikasi faktor sanitasi lingkungan yang paling berpengaruh. Penyusunan alur penelitian ini bertujuan untuk memperjelas tahapan kerja yang dilakukan serta menunjukkan hubungan antar proses dalam penelitian secara keseluruhan. Alur penelitian yang digunakan dalam penelitian ini dapat dilihat pada Gambar 1, yang menyajikan secara visual tahapan-tahapan proses penelitian dari awal hingga akhir.



Gambar 1. Alur Penelitian Terstuktur Menggunakan Random Forest

III. HASIL DAN PEMBAHASAN

Penelitian ini menghasilkan model prediksi risiko stunting pada anak menggunakan metode Random Forest berdasarkan variabel sanitasi lingkungan. Model dibangun menggunakan data yang telah melalui tahap *preprocessing* dan pembagian data menjadi *training* dan *testing* dengan perbandingan 80:20. Proses pelatihan model menunjukkan bahwa Random Forest mampu mempelajari pola hubungan antara kondisi sanitasi lingkungan dan status stunting dengan baik.

Implementasi model dalam penelitian ini dilakukan menggunakan bahasa pemrograman Python, dengan memanfaatkan beberapa pustaka yang mendukung proses pengolahan data dan klasifikasi. Python dipilih karena memiliki fleksibilitas tinggi serta didukung oleh berbagai pustaka yang banyak digunakan dalam analisis data dan pengembangan model prediksi.

A. Data Penelitian

Data yang digunakan dalam penelitian ini merupakan data sekunder dalam kurun waktu 2023 – 2025. Dataset mencakup variabel sanitasi lingkungan dan status stunting anak. Setelah melalui proses seleksi dan pembersihan data, diperoleh sebanyak 200 data yang digunakan dalam proses pemodelan. Data yang digunakan merupakan kondisi nyata di lapangan, sehingga dapat memberikan gambaran yang lebih akurat mengenai hubungan antara sanitasi lingkungan dan risiko stunting pada anak.

B. Preprocessing dan Pembagian Data

Proses ini meliputi pembersihan data dari nilai yang hilang, transformasi variabel, serta normalisasi data. Pembersihan data dilakukan dengan menghapus data yang tidak lengkap atau tidak konsisten, sehingga dataset menjadi lebih valid dan siap digunakan. Transformasi data dilakukan untuk mengubah variabel kategorik menjadi bentuk numerik agar dapat diproses oleh algoritma Random Forest. Tahap ini sangat penting karena kualitas data sangat mempengaruhi hasil prediksi yang dihasilkan.

Proses *preprocessing* data dilakukan menggunakan pustaka *Pandas* dan *NumPy* untuk pengolahan data, seperti pembersihan data, transformasi variabel, serta pengelolaan dataset. Pembangunan model

klasifikasi dilakukan menggunakan pustaka *Scikit-learn*. Pustaka ini banyak digunakan karena kemudahan penggunaan serta kemampuannya dalam menangani berbagai permasalahan klasifikasi dan prediksi [17].

Dataset yang telah melalui tahap *preprocessing* kemudian dibagi menjadi data *training* dan data *testing* dengan perbandingan 80:20. Data *training* digunakan untuk membangun model, sedangkan data *testing* digunakan untuk menguji kinerja model. Pembagian data ini bertujuan untuk memastikan model yang dibangun memiliki kemampuan yang baik dan tidak hanya mampu mengenali data yang telah dilatih.

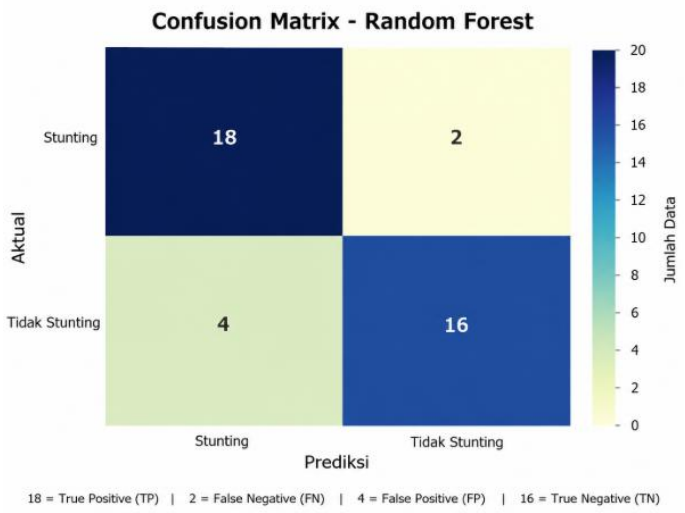
C. Hasil Pemodelan Random Forest

Model prediksi dibangun menggunakan metode Random Forest yang bekerja dengan menggabungkan beberapa *decision tree* untuk meningkatkan akurasi prediksi. Proses pelatihan model dilakukan menggunakan data *training* yang telah disiapkan. Algoritma Random Forest diimplementasikan dengan memanfaatkan fungsi *RandomForestClassifier* yang tersedia pada *Scikit-learn*. Parameter model ditentukan secara *default* dan disesuaikan dengan kebutuhan penelitian untuk menghasilkan performa yang optimal.

Tabel 1 Confusion Matrix

	Prediksi Stunting	Prediksi Tidak Stunting
Aktual Stunting	18 (TP)	2 (FN)
Aktual Tidak Stunting	4 (FP)	16 (TN)

Untuk mengetahui kinerja model secara lebih detail, dilakukan evaluasi menggunakan *confusion matrix*. *Confusion matrix* digunakan untuk melihat jumlah prediksi yang benar dan salah pada masing-masing kelas stunting dan kelas tidak stunting. Hasil *confusion matrix* menunjukkan bahwa jumlah data yang termasuk kategori *True Positive* (TP) dan *True Negative* (TN) lebih dominan dibandingkan dengan kesalahan klasifikasi yaitu *False Positive* (FP) dan *False Negative* (FN). Hasil *confusion matrix* ditunjukkan pada Tabel 1 dan Gambar 3, model memiliki jumlah prediksi benar yang lebih tinggi dibandingkan kesalahan prediksi. Nilai ini mendukung hasil evaluasi akurasi yang mencapai 85%.



Gambar 2. Confusion Matrix Random Forest Menggunakan Scikit-Learn

Evaluasi model dilakukan menggunakan data *testing* untuk mengukur kinerja model dalam melakukan klasifikasi. Berdasarkan hasil pengujian, diperoleh nilai akurasi sebesar 85%, yang menunjukkan bahwa model memiliki tingkat ketepatan yang baik dalam mengklasifikasikan risiko stunting pada anak. Evaluasi juga dilakukan menggunakan *matriks precision*, *recall*, dan *F1-score* untuk memberikan gambaran yang lebih lengkap terhadap kinerja model ditunjukkan pada Tabel 2.

Tabel 2 Hasil Evaluasi Random Forest

Matriks	Nilai
Akurasi	85.0%
Precision	81.8%
Recall	90.4%
F1-Score	85.7%

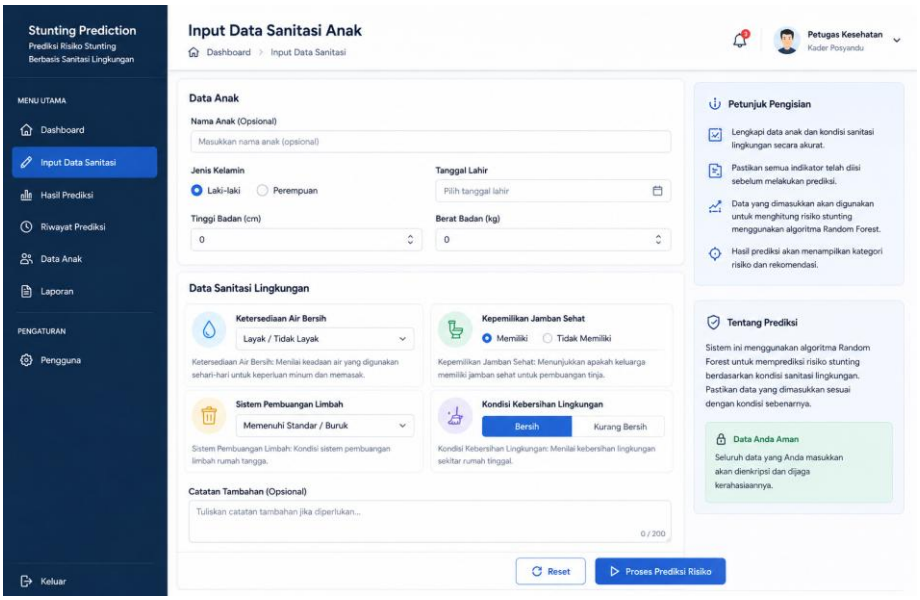
Random Forest juga digunakan untuk mengidentifikasi tingkat pengaruh masing-masing variabel sanitasi lingkungan terhadap risiko stunting pada Tabel 3. Analisis *feature importance* menunjukkan bahwa ketersediaan air bersih merupakan variabel yang memiliki kontribusi terbesar dalam proses prediksi. Variabel berikutnya adalah kepemilikan jamban sehat, diikuti oleh sistem pembuangan limbah, dan kondisi kebersihan lingkungan. Hasil tersebut menunjukkan bahwa akses terhadap air bersih dan kepemilikan jamban sehat merupakan faktor sanitasi yang paling dominan dalam menentukan risiko stunting pada anak.

Tabel 3 Feature Importance

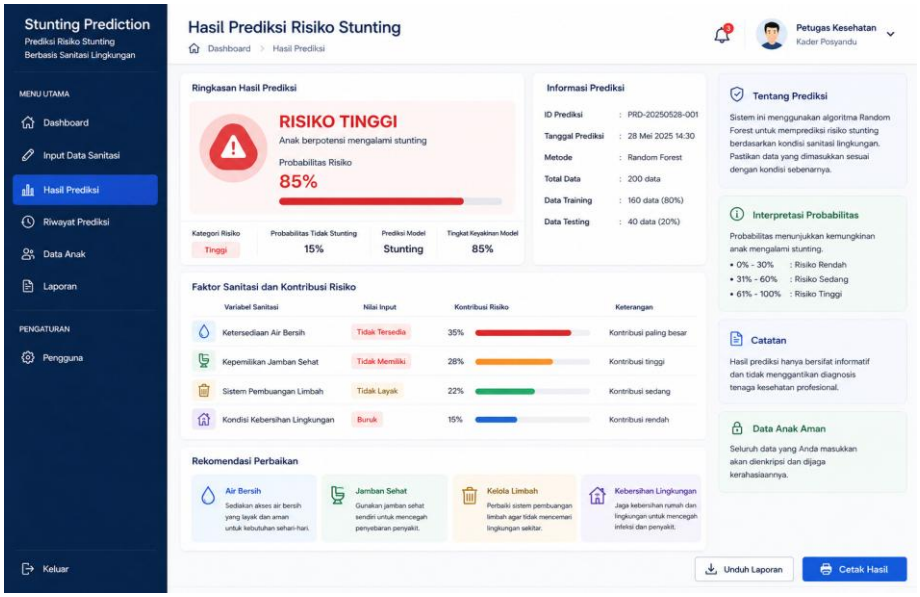
Variabel	Importance
Ketersediaan Air Bersih	0.35
Kepemilikan Jamban Sehat	0.28
Sistem Pembuangan Limbah	0.22
Kondisi Kebersihan Lingkungan	0.15

D. Rancangan Antarmuka Pengguna

Random Forest diimplementasikan ke dalam sebuah antarmuka pengguna berbasis web, sebagai sistem pengambilan keputusan yang praktis. Sistem ini didesain agar mudah digunakan oleh petugas kesehatan atau kader posyandu yang menangani data stunting. Sistem memungkinkan proses identifikasi risiko stunting dilakukan secara lebih cepat dibandingkan metode manual [18]. Pengguna hanya perlu memasukkan kondisi sanitasi lingkungan yang tersedia pada form input, kemudian sistem akan memberikan hasil prediksi secara otomatis. Selain menghasilkan kategori risiko, sistem juga menampilkan faktor sanitasi yang paling mempengaruhi hasil prediksi. Informasi ini dapat digunakan sebagai dasar pengambilan keputusan oleh petugas kesehatan dalam menentukan prioritas pencegahan stunting.



Gambar 3. Rancangan Antarmuka Input Data



Gambar 4. Rancangan Antarmuka Hasil Prediksi

IV. KESIMPULAN

Penelitian ini berhasil membangun model prediksi risiko stunting pada anak berdasarkan variabel sanitasi lingkungan menggunakan metode Random Forest. Variabel yang digunakan meliputi ketersediaan air bersih, kepemilikan jamban sehat, sistem pembuangan limbah, dan kondisi kebersihan lingkungan.

Hasil pengujian menunjukkan bahwa model mampu mengklasifikasikan risiko stunting dengan tingkat akurasi sebesar 85.0%, precision 81.8%, recall 90.4%, dan F1-score 85.7%. Hasil tersebut menunjukkan bahwa metode Random Forest memiliki kemampuan yang baik dalam mengenali pola hubungan antara sanitasi lingkungan dan risiko stunting.

Analisis menunjukkan bahwa ketersediaan air bersih dan kepemilikan jamban sehat merupakan faktor yang paling berpengaruh terhadap risiko stunting. Implementasi model dalam bentuk sistem prediksi berbasis web dapat membantu petugas kesehatan melakukan identifikasi dini terhadap anak yang berisiko stunting sehingga upaya pencegahan dapat dilakukan lebih cepat dan tepat sasaran.

V. DAFTAR PUSTAKA

[1] Kementerian Kesehatan Republik Indonesia, Profil Kesehatan Indonesia Tahun. Jakarta, 2022.
[2] Kementerian PPN/Bappenas, Strategi Nasional Percepatan Pencegahan Stunting 2020–2024. Jakarta, 2020.
[3] R. Torlesse, H., Cronin, A. A., Sebayang, S. K., & Nandy, Determinants of stunting in Indonesian children: Evidence from a cross-sectional survey, 21 (1). BMC Public Health, 2021.
[4] W. W. Danaei, G., Andrews, K. G., Sudfeld, C. R., Fink, G., McCoy, D. C., Peet, E., & Fawzi, Risk factors for childhood stunting in 137 developing countries, 8 (2). The Lancet Global Health.
[5] J. H. Prendergast, A. J., & Humphrey, The stunting syndrome in developing countries. Paediatrics and International Child Health, 41 (3). 2021.
[6] M. Rahman, M., Hasan, M., & Islam, "Machine learning approaches for predicting child malnutrition," Healthc. Inform. Res., vol. 28 (3), pp. 210–218, 2022.
[7] Y. Kurniawan, D., & Sari, "Implementation of Random Forest for stunting classification in Indonesia," J. Inf. Syst. Eng. Bus. Intell., vol. 9(2), pp. 123–130, 2023.
[8] L. Putri, N. E., & Handayani, "Analysis of environmental sanitation factors on stunting incidence in children under five," J. Kesehat. Masy., vol. 19 (1), pp. 45–53, 2024.
[9] D. Support, E. On, and R. Adolescent, "Indonesian Journal of Global Health Research," Indones. J. Glob. Heal. Res., vol. 2, no. 4, pp. 1075–1086, 2019, doi: 10.37287/ijghr.v2i4.250.
[10] N. Januari, K. Mande, and K. Cianjur, "SENTRI : Jurnal Riset Ilmiah," vol. 3, no. 1, pp. 339–348,

2024.
[11] N. Cicci, H. Hamzah, F. K. Masyarakat, U. S. Barat, and R. K. Majene, “Identifikasi Risiko dan Evaluasi Intervensi Spesifik pada Anak Stunting Berdasarkan Audit Kasus Stunting Berbasis Surveilans di Wilayah Lokus Kabupaten Majene,” vol. 7, no. 1, pp. 90–96, 2024.
[12] S. R. Nadhiroh, “Faktor yang berhubungan dengan kejadian,” 2010.
[13] T. K. Ho, “Random decision forests,” in Proceedings of the Third International Conference on Document Analysis and Recognition, pp. 278–282.
[14] S. A. Ningrum, M. A. Barata, and N. Mahmudah, “Klasifikasi Status Stunting Pada Balita di Kecamatan Singgahan dengan Algoritma Support Vector Machine,” vol. 11, no. 1, pp. 28–33, 2025.
[15] and K. H. R. C. Fink, G. Günther, “The effect of water and sanitation on child health: Evidence from demographic and health surveys,” Int. J. Epidemiol., vol. 40, pp. 1196–1204, 2011.
[16] R. W. Febriana, R. Kriswibowo, J. S. Prayogo, P. A. Alia, and S. Budi, “Decision Support System Diagnosis Penyakit Stroke Menggunakan Metode Composite Performance Index (CPI),” vol. 4, no. September, pp. 111–121, 2025.
[17] L. Breiman, “Random Forests,” Machine Learning. 2001.
[18] dkk B. J. D. Sitompul, Teknik Dasar Pembuatan Website. Jawa Tengah: Penamuda Media, 2024.