



## *Identification of Stunting Risk Based on Maternal Prenatal Factors Using Random Forest Approach: A Preliminary Retrospective Study*

**Kaysa Naisy Khosina<sup>1</sup>, Pramesti Kusumaningtyas<sup>2\*</sup>, Mohammad Rofi'i<sup>3</sup>**

<sup>1-3</sup>Program Studi D-3 Teknik Elektomedis, Sekolah Tinggi Ilmu Kesehatan Semarang, Indonesia

Email: [kaysanaisy@gmail.com](mailto:kaysanaisy@gmail.com) [mohammadrofii@stikesemarang.ac.id](mailto:mohammadrofii@stikesemarang.ac.id)

\*Penulis Korespondensi: [emailkantoresti@gmail.com](mailto:emailkantoresti@gmail.com)

**Abstract.** *Stunting is a multifactorial public health problem influenced by various risk factors that may emerge during the prenatal period. Early identification of stunting risk during pregnancy is important to support preventive interventions. This study aimed to develop a stunting risk prediction model based on maternal prenatal factors using the Random Forest algorithm. Secondary data from 172 pregnant women, consisting of 83 stunting cases and 89 non-stunting cases, were analyzed. The predictor variables included maternal age during pregnancy, height, hemoglobin level, mid-upper arm circumference (MUAC), smoking history, hypertension, asthma, and diabetes mellitus. The research stages consisted of data preprocessing, model training using Stratified 5-Fold Cross Validation, performance evaluation, external testing, and feature importance analysis. Internal evaluation results showed an accuracy of 60%, precision of 60.6%, recall of 57.3%, F1-score of 58.9%, and AUC of 0.6688. External testing yielded an accuracy of 70% and an AUC of 0.6167. Feature importance analysis identified maternal age during pregnancy as the most influential variable in the prediction process. The findings indicate that maternal prenatal factors have potential for early stunting risk identification, although the predictive performance remains moderate. This approach may serve as a foundation for developing early screening tools to support targeted interventions among high-risk pregnancies.*

**Keywords:** *Maternal Factors; Machine Learning; Predictive Model; Random Forest; Stunting.*

**Abstrak.** Stunting merupakan masalah kesehatan masyarakat yang bersifat multifaktorial dan dapat dipengaruhi oleh berbagai faktor yang telah muncul sejak masa prenatal. Identifikasi risiko stunting sejak kehamilan diperlukan untuk mendukung upaya pencegahan yang lebih dini dan preventif. Penelitian ini bertujuan mengembangkan model prediksi risiko stunting berdasarkan faktor maternal prenatal menggunakan algoritma Random Forest. Penelitian menggunakan data sekunder sebanyak 172 ibu hamil yang terdiri atas 83 kasus stunting dan 89 non-stunting. Variabel yang digunakan meliputi usia ibu saat hamil, tinggi badan, kadar hemoglobin (Hb), Lingkar Lengan Atas (LILA), riwayat merokok, hipertensi, asma, dan diabetes melitus (DM). Tahapan penelitian meliputi pra-pemrosesan data, pelatihan model menggunakan Stratified 5-Fold Cross Validation, evaluasi performa model, pengujian eksternal, dan analisis feature importance. Hasil evaluasi internal menunjukkan nilai akurasi sebesar 60%, presisi 60,6%, recall 57,3%, F1-score 58,9%, dan AUC 0,6688. Pada pengujian eksternal, model memperoleh akurasi sebesar 70% dan AUC sebesar 0,6167. Analisis feature importance menunjukkan bahwa usia ibu saat hamil merupakan variabel yang memberikan kontribusi terbesar dalam proses prediksi. Hasil penelitian menunjukkan bahwa faktor maternal prenatal memiliki potensi untuk dimanfaatkan dalam identifikasi dini risiko stunting meskipun kemampuan prediksi model masih berada pada kategori sedang. Pendekatan ini dapat menjadi dasar pengembangan sistem skrining awal guna mendukung intervensi yang lebih tepat sasaran pada kelompok kehamilan berisiko.

**Kata kunci:** Faktor maternal; pembelajaran mesin; pemodelan prediktif; Random Forest; stunting.

### **1. LATAR BELAKANG**

Masa kehamilan merupakan fase kritis yang mempengaruhi pertumbuhan dan perkembangan janin. Kondisi fisiologis ibu selama kehamilan berperan penting dalam menentukan kecukupan nutrisi, oksigen dan keseimbangan metabolik yang dibutuhkan janin (Joshi et al., 2023). Dengan demikian, setiap gangguan pada kondisi kesehatan ibu, baik yang bersifat kronik maupun akut, berpotensi mempengaruhi pertumbuhan dan perkembangan janin

sejak tahap paling awal kehidupan. Berbagai indikator kesehatan ibu seperti status gizi, kadar hemoglobin, serta riwayat penyakit mencerminkan kapasitas tubuh ibu dalam mendukung pertumbuhan janin (Hendraswari et al., 2025; Marine, Haile, & Zewde, 2026a, 2026b; Nakshine & Jogdand, 2023). Lingkar Lengan Atas (LILA) misalnya, tidak hanya menggambarkan kondisi gizi tetapi juga cadangan energi yang tersedia selama kehamilan (Salih, Omar, AlHabardi, & Adam, 2023). Demikian pula anemia pada ibu hamil dapat menghambat distribusi oksigen ke jaringan janin, yang dalam jangka panjang dapat mengganggu proses pertumbuhan intrauterin (Díaz-Torres, Díaz-López, & Arija, 2024). Faktor lain seperti tinggi badan ibu dan jarak kehamilan juga merepresentasikan kondisi biologis dan reproduktif yang berkontribusi terhadap kualitas kehamilan (Ni et al., 2023). Dampak kondisi tersebut tidak selalu terlihat secara langsung saat kelahiran, tetapi dapat muncul dalam bentuk gangguan pertumbuhan kronis pada anak, termasuk gangguan pertumbuhan stunting (Muglia, Benhalima, Tong, & Ozanne, 2022).

Menurut standar WHO, stunting didefinisikan sebagai kondisi gagal tumbuh yang ditandai dengan tinggi badan menurut umur berada di bawah -2 standar deviasi. Kondisi ini menunjukkan kekurangan gizi jangka panjang akibat pola makan anak dan juga merupakan hasil akumulasi dari berbagai faktor risiko yang telah terjadi sejak masa prenatal (Kustanto, Rachmat, & Setyadi, 2024; Ratnawati, Irene Kartasurya, Kartini, & Noe, 2025). Pendekatan penanganan stunting selama ini cenderung berfokus pada fase pascakelahiran, seperti intervensi nutrisi pada bayi dan balita (Aulia & Dewi Purnamawati, 2025; Dedeng Nurkholik Sidik Permana, Nilawaty Uly, & Andi Alim, 2025; Suharni, Nurtanny, & Ardian, 2025). Namun, pendekatan tersebut seringkali terlambat, mengingat sejumlah faktor risiko stunting telah muncul sejak periode prenatal. Oleh karena itu, diperlukan pendekatan yang mampu mengidentifikasi risiko stunting sejak dini, bahkan sebelum anak dilahirkan, sehingga intervensi dapat dilakukan secara lebih preventif.

Di sisi lain, hubungan antara berbagai faktor maternal dan kejadian stunting memiliki potensi hubungan non-linier dan interaksi antarvariabel yang kompleks (Rao et al., 2026). Interaksi antar variabel seperti status gizi, kondisi kesehatan dan faktor reproduksi sulit dianalisis secara optimal dengan metode statistik konvensional. Dalam konteks ini, perkembangan teknologi machine learning menawarkan pendekatan baru yang mampu menangkap pola hubungan yang lebih kompleks (Karypidis, Mouslech, Skoulariki, & Gazis, 2022). Salah satu metode yang banyak digunakan adalah Random Forest, yang memiliki kemampuan dalam mengelola data dengan karakteristik non-linier serta mengidentifikasi kontribusi variabel terhadap hasil (Sheth, Tripathi, & Sharma, 2022). Random Forest dipilih

karena mampu menangani data numerik dan kategorik secara bersamaan, tidak memerlukan asumsi linieritas, relatif tahan terhadap overfitting pada dataset berukuran terbatas, serta menyediakan interpretasi melalui feature importance (Salman, Kalakech, & Steiti, 2024a). Pendekatan ini memungkinkan pengembangan model prediksi risiko stunting berbasis data maternal yang lebih adaptif dan berpotensi meningkatkan kemampuan identifikasi pola risiko. Dengan memanfaatkan data kesehatan ibu yang tersedia dalam layanan kesehatan primer, seperti buku Kesehatan Ibu dan Anak (KIA), penelitian ini berupaya mengimplementasikan *machine learning* dengan pendekatan algoritma Random Forest untuk mengembangkan model prediksi risiko stunting. Model ini diharapkan dapat mendukung pengembangan pendekatan skrining awal guna membantu tenaga kesehatan dalam mengidentifikasi ibu hamil dengan risiko tinggi, sehingga intervensi dapat dilakukan lebih dini dan tepat sasaran.

## 2. KAJIAN TEORITIS

### Faktor Maternal terhadap Risiko Stunting

Stunting merupakan kondisi gagal tumbuh kronis pada anak yang ditandai dengan tinggi badan menurut umur berada di bawah -2 standar deviasi berdasarkan standar *World Health Organization* (WHO). Dampak stunting tidak hanya bersifat jangka pendek seperti gangguan perkembangan otak, penurunan daya tahan tubuh, dan peningkatan risiko kesakitan, tetapi juga berdampak jangka panjang berupa penurunan kemampuan kognitif, peningkatan risiko penyakit tidak menular, serta penurunan produktivitas di masa dewasa (Amran, Nurwiyeni, Pratama, & Wahyuni, 2025; Laily & Indarjo, 2023). Kondisi ini berkembang akibat akumulasi berbagai faktor risiko, baik sebelum maupun setelah kelahiran, termasuk status gizi ibu, kondisi kesehatan selama kehamilan, asupan nutrisi anak, serta faktor lingkungan dan sosial ekonomi. Periode 1.000 Hari Pertama Kehidupan (HPK) menjadi fase penting dalam pencegahan stunting karena gangguan pertumbuhan yang terjadi pada periode ini dapat berdampak jangka panjang terhadap perkembangan fisik dan kognitif anak (Meliati, Sudarmi, & Halimatusyaadiah, 2025; Nurdivanuha & Wibawani, 2025). Oleh karena itu, identifikasi faktor risiko sejak masa kehamilan menjadi penting sebagai bagian dari upaya pencegahan stunting secara dini.

Kondisi maternal selama kehamilan memiliki keterkaitan dengan pertumbuhan dan perkembangan janin. Berbagai parameter kesehatan ibu sering digunakan dalam kajian risiko stunting karena dapat mencerminkan status gizi, kondisi fisiologis, serta kualitas lingkungan intrauterin selama masa kehamilan. Tinggi badan ibu dapat merepresentasikan kondisi gizi

jangka panjang yang berhubungan dengan pertumbuhan janin (Alam Fajar, Liberty, Sri Ananingsih, & Nur Amalia, 2025), sedangkan usia ibu hamil berkaitan dengan kesiapan biologis dan risiko komplikasi kehamilan. Selain itu, Lingkar Lengan Atas (LILA) digunakan sebagai indikator status gizi dan risiko kekurangan energi kronis pada ibu hamil (Irma, Sugirah Nour Rahman, Hasra Ryska, & Veny Hadju, 2026).

Kondisi kesehatan ibu selama kehamilan juga dapat mempengaruhi kecukupan nutrisi dan oksigen yang diterima janin. Kadar Hemoglobin (Hb) yang rendah dapat mengindikasikan anemia yang berpotensi menghambat distribusi oksigen ke jaringan janin. Riwayat merokok pada ibu hamil dapat meningkatkan paparan zat toksik yang mempengaruhi fungsi plasenta dan pertumbuhan janin (Irma et al., 2026). Sementara itu, riwayat penyakit seperti hipertensi, asma dan diabetes militus (DM) dapat mempengaruhi proses metabolik dan sirkulasi selama kehamilan sehingga berpotensi mengganggu pertumbuhan intrauterin (Alejandra, Ricardo, & Guadalupe, 2025; Valverde-Pérez, Olea, Rocher, Aaronson, & Prieto-Lloret, 2025; Vlachou et al., 2025). Oleh karena itu, parameter maternal tersebut dinilai memiliki keterkaitan terhadap risiko stunting dan digunakan sebagai variabel dalam penelitian ini.

### **Random Forest dalam prediksi risiko stunting**

Random Forest merupakan algoritma yang dikembangkan dari kumpulan pohon keputusan untuk menghasilkan model klasifikasi yang lebih stabil. Sejumlah pohon keputusan dibentuk menggunakan kombinasi data dan fitur yang dipilih secara acak. Setiap pohon akan menghasilkan prediksi tersendiri, kemudian hasil akhir ditentukan berdasarkan jumlah prediksi terbanyak dari seluruh pohon yang terbentuk (Salwa Alexita, Kusumaningtyas, & Rofi'i, 2025). Pendekatan tersebut membuat Random Forest mampu mengurangi ketergantungan terhadap satu model pohon keputusan sehingga klasifikasi menjadi lebih konsisten.

Penggunaan Random Forest dalam penelitian kesehatan cukup banyak dilakukan karena algoritma ini mampu menangani hubungan antarvariabel yang tidak selalu linier, seperti hubungan faktor maternal prenatal terhadap kasus stunting (Salman, Kalakech, & Steiti, 2024b). Variabel yang digunakan terdiri atas beberapa karakteristik ibu dengan kondisi kesehatan yang berbeda-beda, sehingga kemungkinan terdapat interaksi antarvariabel yang cukup kompleks. Random Forest dinilai efektif digunakan pada data numerik maupun kategorik. Random Forest juga memiliki ketahanan yang baik terhadap noise dan risiko overfitting. Metode ini juga memungkinkan identifikasi variabel yang memiliki kontribusi lebih besar terhadap hasil prediksi melalui analisis feature importance, sehingga dapat membantu interpretasi faktor-faktor yang berpengaruh dalam suatu model klasifikasi (Khan, Ali, Khan, Ullah, & Faheem, 2025).

### 3. METODE PENELITIAN

Penelitian ini merupakan penelitian kuantitatif dengan pendekatan prediktif yang bertujuan untuk membangun model klasifikasi risiko stunting berdasarkan faktor maternal menggunakan algoritma Random Forest. Tahapan penelitian dilakukan melalui proses pengolahan dan pra-pemrosesan data, pembangunan model latih dan model prediksi, evaluasi model dan analisis feature importance.

#### **Dataset Variabel Ibu Hamil**

Dataset yang digunakan dalam penelitian ini merupakan data sekunder yang diperoleh dari penelitian sebelumnya (Setyaningsih, Noor Wijayanti, Widayati, Susanti, & Respati Yogyakarta, 2024) yang menggunakan desain case-control dengan pendekatan retrospektif di tiga wilayah di Indonesia. Berbeda dengan penelitian sebelumnya yang berfokus pada analisis asosiasi antara faktor maternal prenatal dengan kejadian stunting, penelitian ini melakukan analisis lanjutan terhadap dataset yang sama dengan pendekatan machine learning untuk mengembangkan model prediksi risiko stunting secara individual.

**Tabel 1.** Variabel Penelitian.

Variabel	Skala	Deskripsi
Usia	Numerik	Usia ibu dalam tahun
Tinggi	Numerik	Tinggi badan ibu dalam centimeter
LILA	Numerik	Lingkar Lengan Atas ibu dalam centimeter
Hb	Numerik	Kadar hemoglobin dalam darah dalam g/dL
Merokok	Kategorik	Riwayat merokok pada ibu hamil (ya/tidak)
Hipertensi	Kategorik	Riwayat hipertensi ibu hamil(ya/tidak)
Asma	Kategorik	Riwayat asma ibu hamil (ya/tidak)
DM	Kategorik	Riwayat diabetes militus ibu hamil (ya/tidak)
Stunting	Kategorik	Target (ya/tidak)

Jumlah dataset yang digunakan sebanyak 172 data ibu hamil dengan sebaran 83 kasus stunting dan 89 kasus non-stunting. Variabel yang digunakan mencakup output target status stunting dan variabel input (prediktor) meliputi usia ibu saat hamil, tinggi badan ibu, Lingkar Lengan Atas (LILA), kadar hemoglobin (Hb), riwayat merokok, riwayat hipertensi, riwayat asma dan riwayat diabetes militus (DM). Variabel yang digunakan dalam penelitian ini ditentukan berdasarkan ketersediaan data pada sumber data retrospektif yang digunakan. Data tersebut memiliki tipe data numerik dan kategorik sesuai dengan Tabel 1. Data tersebut selanjutnya akan diolah di tahap pra-pemrosesan sebagai pelatihan model dan prediksi.

## Pra-pemrosesan Data

Sebelum proses pemodelan dilakukan, data terlebih dulu melalui tahap pra-pemrosesan yang bertujuan untuk meningkatkan kualitas data yang digunakan. Tahap ini meliputi pemeriksaan kelengkapan data (*imputasi missing value*), transformasi variabel kategorik dalam format numerik melalui encoding, serta pemisahan fitur (x) dan label (y) (Vinita Yadav & Kumar, 2022). Tahap pra-pemrosesan bertujuan memastikan kualitas data sehingga dapat digunakan secara optimal pada proses pelatihan model. Diketahui terdapat 2 missing data, yaitu 1 data variabel Hb dan 1 data variabel LILA. Penanganan missing value dengan imputansi data yaitu nilai Hb rata-rata 12,2 dan nilai LILA rata-rata 26,0. Dataset yang telah melalui tahap pra-pemrosesan kemudian dibagi menjadi 150 data internal untuk pelatihan dan validasi model dan 22 data pengujian independen.

## Implementasi dan Pengujian Random Forest

Pada penelitian ini, model random forest diimplementasikan menggunakan pustaka Scikit-Learn dengan parameter yang tertera pada **Tabel 2**.

**Tabel 2.** Parameter pada Random Forest.

Parameter	Keterangan
n_estimators	100
max_depth	5
min_samples_leaf	5
min_samples_split	10
class_weight	balanced
random_state	42
n_jobs	-1

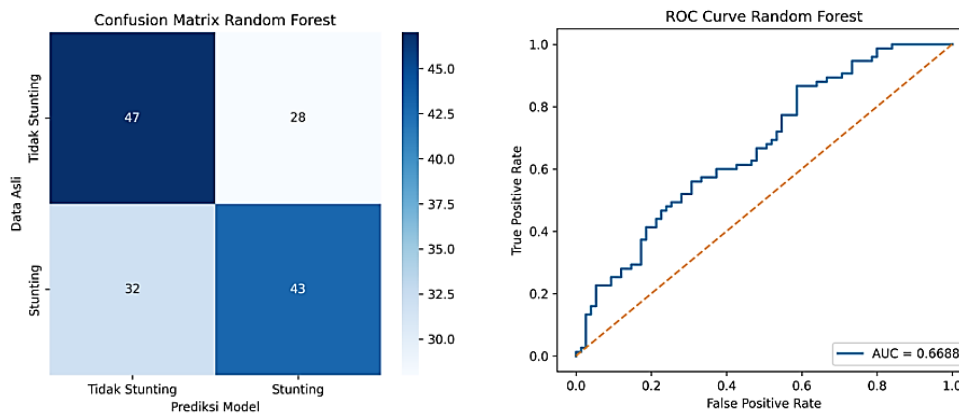
Penggunaan parameter n\_jobs bernilai -1 untuk mengoptimalkan pemrosesan komputasi dan parameter class\_weight balances bertujuan menjaga keseimbangan kontribusi masing-masing kelas selama proses pelatihan model. Model juga dilengkapi dengan analisis *feature importance* untuk mengetahui variabel yang paling berpengaruh dalam proses prediksi. Nilai *feature importance* dihitung berdasarkan kontribusi suatu variabel dalam meningkatkan kualitas pemisahan data pada seluruh pohon keputusan yang terbentuk. Pengaturan parameter pada feature importance yaitu n\_repeat = 10, random\_state = 42, n\_jobs = -1. Semakin tinggi nilai feature importance, semakin besar peran variabel tersebut dalam menghasilkan prediksi model (Yuan, Liu, Feng, & Dauphin, 2023). Selanjutnya, proses pelatihan model menggunakan Stratified K-Fold Cross validation dengan nilai performa yang dilaporkan merupakan rata-rata

hasil evaluasi dari 5 fold untuk mengurangi variasi pembagian data terhadap performa model. Pendekatan ini dipilih karena dataset yang digunakan relatif terbatas. Akhir tahapan, kinerja model diukur menggunakan confusion matrix yang menghasilkan beberapa metrik evaluasi yaitu accuracy, precision, recall dan F1-score. Kemampuan model dalam membedakan kelas dianalisis menggunakan *Area Under Curve (AUC)* dari kurva *Receiver Operating Characteristic (ROC)*. Setelah proses validasi selesai dilakukan, model selanjutnya diuji menggunakan data eksternal untuk mengetahui kemampuan generalisasi model terhadap data yang tidak digunakan selama proses pelatihan.

#### 4. HASIL DAN PEMBAHASAN

##### Kinerja Model dalam Identifikasi Risiko Stunting

Pengujian performa model menggunakan dataset sejumlah 150 data terdiri dari 75 data kasus stunting dan 75 data kasus non-stunting. Hasil pengujian diperoleh 90 data diklasifikasikan dengan benar dan 60 data salah diklasifikasikan, detail nilai yang diperoleh digambarkan dengan confusion matrix pada Gambar 1. dan hasil perhitungan parameter kinerja model terdapat pada Tabel 3.



**Gambar 1.** Confusion Matrix dan Kurva ROC-AUC Data Internal.

**Tabel 3.** Kinerja Model Prediksi Stunting.

Metrik Parameter	Nilai
Akurasi	60%
Presisi	60,6%
Recall	57,3%
F1-Score	58,9%
AUC	0,6688

Dari seluruh sampel, model berhasil mengklasifikasikan sekitar 60% data dengan benar. Nilai presisi sebesar 60,6% menunjukkan bahwa dari seluruh sampel yang diprediksi berisiko stunting, sekitar 60,6% di antaranya benar-benar termasuk kelompok stunting. Sementara itu, nilai recall sebesar 57,3% mengindikasikan bahwa model mampu mengidentifikasi sekitar 57,3% kasus stunting yang sebenarnya terdapat dalam data. Nilai F1-score sebesar 58,9% menunjukkan keseimbangan yang masih moderat antara kemampuan model dalam mendeteksi kasus stunting dan ketepatan prediksi yang dihasilkan.

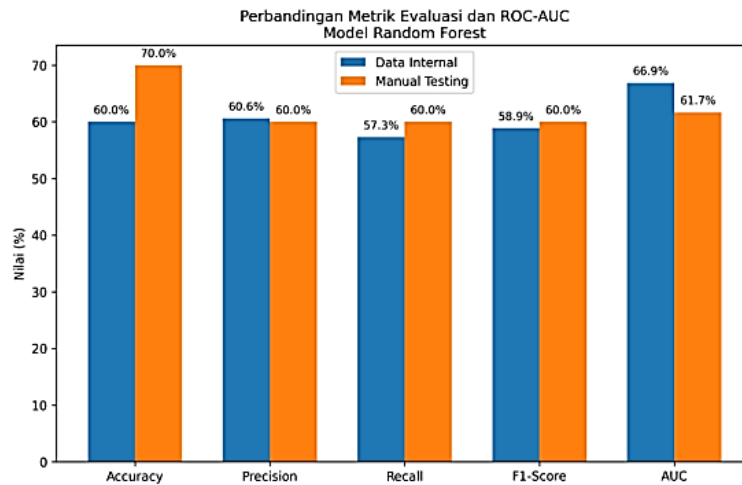
Nilai AUC sebesar 0,6688 menunjukkan model mampu membedakan kelompok stunting dan non-stunting dengan kemampuan diskriminasi dibanding klasifikasi acak ( $AUC=0,5$ ). Nilai ini menunjukkan jika dipilih secara acak satu individu dari kelompok stunting dan satu individu dari kelompok non-stunting, model memiliki probabilitas sekitar 66,88% untuk memberikan skor risiko yang lebih tinggi pada individu yang benar-benar mengalami stunting. Namun demikian, nilai tersebut juga mengindikasikan bahwa faktor maternal prenatal yang digunakan belum sepenuhnya mampu menjelaskan variasi kejadian stunting. Kondisi ini dapat dipahami mengingat stunting merupakan masalah multifaktoral yang tidak hanya dipengaruhi oleh kondisi ibu selama kehamilan, tetapi juga oleh berbagai faktor pasca-kelahiran seperti pola pemberian makan, kondisi lingkungan, status sosial ekonomi dan riwayat penyakit pada anak.

Meskipun performa model belum menunjukkan tingkat akurasi yang tinggi, hasil penelitian tetap memberikan informasi penting bahwa data yang tersedia selama masa kehamilan berpotensi dimanfaatkan untuk proses identifikasi dini risiko stunting. Dengan demikian, pendekatan prediktif berbasis faktor maternal dapat dipertimbangkan sebagai alat bantu skrining awal untuk menentukan kelompok ibu hamil yang memerlukan pemantauan dan intervensi lebih intensif selama periode kehamilan maupun setelah persalinan.

### **Kemampuan Generalisasi Model**

Pengujian dilakukan menggunakan dataset eksternal yang tidak terlibat dalam proses pelatihan untuk menguji kemampuan generalisasi model. Hasil pengujian menunjukkan bahwa model memperoleh accuracy sebesar 70% dengan nilai AUC sebesar 0,6167 seperti yang tertera pada Grafik Perbandingan Metrik Evaluasi pada Gambar 2. Jika dibandingkan dengan hasil evaluasi internal yang menghasilkan akurasi 60% dan AUC 0,6688, terlihat bahwa model masih menunjukkan kemampuan klasifikasi pada data yang berbeda, meskipun diperlukan pengujian pada dataset yang lebih besar untuk memastikan kemampuan generalisasinya. Penurunan nilai AUC menunjukkan bahwa pola hubungan antara faktor maternal prenatal terhadap kejadian stunting pada data eksternal tidak sepenuhnya identik dengan data pelatihan,

yang dapat dipengaruhi oleh variasi karakteristik sampel maupun faktor-faktor lain yang tidak tercakup dalam model.



**Gambar 2.** Perbandingan Metrik Evaluasi Data Internal dan Data Eksternal.

### Kontribusi Faktor Maternal dan Implikasinya terhadap Pencegahan Stunting

Untuk memahami kontribusi masing-masing variabel dalam proses prediksi risiko stunting, dilakukan analisis feature importance pada model Random Forest yang dibangun. Hasil analisis feature importance tertera pada Tabel 4.

**Tabel 4.** Analisis Feature Importance.

Variabel	Importance
Usia	0,1439
LILA	0,0693
Tinggi badan	0,0646
Hb	0,0306
Hipertensi	0,0133
Asma	0,0093
Merokok	0
DM	0

Hasil analisis feature importance menunjukkan bahwa usia ibu saat hamil merupakan variabel yang memberikan kontribusi terbesar dalam proses identifikasi risiko stunting. Temuan ini sejalan dengan berbagai penelitian yang melaporkan bahwa kehamilan pada usia terlalu muda maupun terlalu tua berhubungan dengan peningkatan risiko gangguan pertumbuhan pada anak. Pada usia yang terlalu muda, kebutuhan nutrisi ibu yang masih berada dalam fase pertumbuhan dapat berkompetisi dengan kebutuhan nutrisi janin, sedangkan pada usia yang lebih tua risiko komplikasi kehamilan dan penyakit penyerta cenderung meningkat. Dari perspektif kesehatan masyarakat, temuan ini menunjukkan pentingnya pemantauan yang lebih intensif terhadap kelompok ibu hamil dengan usia berisiko sebagai bagian dari upaya pencegahan stunting sejak masa prenatal. Temuan ini juga mendukung pentingnya program pendewasaan usia perkawinan dan perencanaan kehamilan sebagai bagian dari strategi

pengecahan stunting jangka panjang. Dengan mendorong kehamilan pada rentang usia reproduksi yang optimal, risiko gangguan pertumbuhan yang berasal dari faktor prenatal diharapkan dapat dikurangi.

LILA merupakan variabel dengan kontribusi terbesar kedua dalam model. LILA sering digunakan sebagai indikator status gizi ibu hamil dan risiko kekurangan energi kronis. Implikasi praktis dari temuan ini adalah perlunya penguatan program pemantauan status gizi ibu hamil melalui pemeriksaan rutin LILA, pemberian makanan tambahan bagi ibu dengan risiko kekurangan energi kronis, serta edukasi mengenai pemenuhan kebutuhan nutrisi selama kehamilan.

Selanjutnya, tinggi badan ibu menjadi variabel ketiga paling berkontribusi dalam model prediksi. Tinggi badan sering dikaitkan dengan status gizi jangka panjang, genetik, serta kondisi kesehatan pada generasi sebelumnya. Meskipun faktor ini tidak dapat dimodifikasi selama kehamilan, informasi tinggi badan ibu dapat dimanfaatkan sebagai indikator awal untuk mengidentifikasi kelompok kehamilan yang memerlukan pemantauan lebih intensif. Temuan ini menunjukkan bahwa pencegahan stunting tidak hanya perlu dilakukan selama kehamilan, tetapi juga memerlukan pendekatan lintas generasi melalui perbaikan status gizi perempuan sejak masa kanak-kanak dan remaja.

Secara keseluruhan, hasil analisis menunjukkan bahwa usia ibu, status gizi saat kehamilan, dan status gizi jangka panjang merupakan faktor maternal prenatal yang paling berkontribusi dalam identifikasi risiko stunting. Oleh karena itu, strategi pencegahan stunting perlu difokuskan pada peningkatan kesehatan reproduksi, perbaikan status gizi perempuan sebelum dan selama kehamilan, serta identifikasi dini kelompok ibu hamil yang memiliki faktor risiko tersebut.

Dalam konteks layanan kesehatan ibu dan anak, model prediksi ini berpotensi diintegrasikan dengan data rutin yang tersedia pada Buku Kesehatan Ibu dan Anak (KIA) maupun sistem informasi kesehatan difasilitas pelayanan primer. Hasil prediksi dapat dimanfaatkan sebagai alat bantu skrining untuk mengidentifikasi ibu hamil dengan risiko lebih tinggi sehingga tenaga kesehatan dapat memberikan pemantauan yang lebih intensif, edukasi gizi, konseling kehamilan, maupun intervensi preventif secara lebih terarah.

### **Keterbatasan Penelitian**

Penelitian ini memiliki beberapa keterbatasan yang perlu diperhatikan dalam menginterpretasikan hasil. Pertama, jumlah data yang digunakan relatif terbatas, yaitu sebanyak 172 sampel, sehingga kemampuan model dalam mempelajari pola yang lebih kompleks masih dapat ditingkatkan melalui penggunaan dataset yang lebih besar.

Pengujian eksternal menggunakan 22 data sampel, jumlah ini relatif sedikit, sehingga kemampuan generalisasi model pada populasi lebih luas masih perlu dikonfirmasi lebih lanjut. Validasi eksternal menggunakan dataset yang berasal dari wilayah geografis, karakteristik populasi dan periode pengumpulan yang berbeda perlu dilakukan pada penelitian berikutnya agar diperoleh model yang lebih robust dan memiliki tingkat keandalan yang lebih tinggi untuk implementasi pada berbagai setting layanan kesehatan.

Kedua, penelitian ini menggunakan data retrospektif yang diperoleh dari penelitian sebelumnya, sehingga variabel yang dianalisis terbatas pada informasi yang tersedia dalam dataset tersebut. Ketiga, model hanya memanfaatkan faktor maternal prenatal, sedangkan kejadian stunting merupakan kondisi multifaktorial yang juga dipengaruhi oleh faktor pasca-kelahiran seperti pola pemberian makan, status sosial ekonomi, sanitasi lingkungan, riwayat infeksi, dan akses pelayanan kesehatan. Oleh karena itu, hasil penelitian ini lebih tepat dipandang sebagai upaya identifikasi dini risiko stunting berdasarkan kondisi ibu selama kehamilan, bukan sebagai alat diagnosis yang dapat menggantikan penilaian klinis maupun surveilans pertumbuhan anak.

## **5. KESIMPULAN DAN SARAN**

Penelitian ini berhasil mengembangkan model prediksi risiko stunting berbasis faktor maternal prenatal menggunakan algoritma Random Forest. Model yang dibangun menggunakan delapan variabel maternal, yaitu usia ibu saat hamil, tinggi badan, LILA, kadar hemoglobin, riwayat merokok, hipertensi, asma, dan diabetes melitus. Hasil evaluasi menunjukkan bahwa model memiliki kemampuan klasifikasi pada kategori sedang dengan akurasi sebesar 60%, F1-score sebesar 58,9%, dan AUC sebesar 0,6688 pada data internal. Pengujian menggunakan data eksternal menghasilkan akurasi sebesar 70% dan AUC sebesar 0,6167, yang menunjukkan bahwa model masih mampu mempertahankan kemampuan generalisasi pada data yang berbeda.

Analisis feature importance menunjukkan bahwa usia ibu saat hamil merupakan faktor yang memberikan kontribusi terbesar dalam proses identifikasi risiko stunting. Temuan ini memperkuat pentingnya faktor maternal prenatal sebagai sumber informasi awal untuk mendeteksi risiko stunting sebelum kelahiran. Meskipun model yang dikembangkan belum dapat digunakan sebagai alat diagnosis, hasil penelitian menunjukkan bahwa informasi kesehatan ibu selama kehamilan berpotensi dimanfaatkan sebagai dasar pengembangan sistem skrining awal risiko stunting untuk mendukung upaya pencegahan yang lebih dini, preventif, dan tepat sasaran pada kelompok berisiko.

Berdasarkan hasil penelitian, pengembangan model prediksi risiko stunting berbasis faktor maternal prenatal masih memiliki peluang untuk ditingkatkan. Walaupun menunjukkan potensi dalam identifikasi dini risiko stunting, performa model yang masih berada pada kategori sedang menunjukkan bahwa model ini masih memerlukan pengembangan lebih lanjut sebelum dapat diterapkan secara luas. Peningkatan jumlah data, perluasan variasi prediktor, serta validasi pada populasi yang lebih beragam diperlukan untuk meningkatkan kemampuan prediksi dan keandalan model. Penelitian selanjutnya disarankan menggunakan jumlah data yang lebih besar dan lebih beragam agar model mampu mempelajari pola hubungan yang lebih kompleks serta menghasilkan performa prediksi yang lebih baik. Selain itu, penambahan variabel yang mencerminkan faktor pascakelahiran, seperti pola pemberian makan anak, status sosial ekonomi keluarga, kondisi sanitasi lingkungan, riwayat infeksi, dan akses terhadap layanan kesehatan, perlu dipertimbangkan mengingat stunting merupakan kondisi yang dipengaruhi oleh berbagai faktor sepanjang periode 1.000 Hari Pertama Kehidupan. Penggunaan dan perbandingan dengan algoritma machine learning lainnya juga dapat dilakukan untuk memperoleh model dengan kemampuan klasifikasi yang lebih optimal. Dari sisi implementasi, hasil penelitian ini dapat menjadi dasar pengembangan sistem skrining dini berbasis digital yang terintegrasi dengan layanan kesehatan ibu dan anak. Selain itu, temuan bahwa usia ibu saat hamil merupakan faktor yang paling berkontribusi dalam prediksi risiko stunting menunjukkan pentingnya penguatan program kesehatan reproduksi, pendewasaan usia perkawinan, dan perencanaan kehamilan sebagai bagian dari strategi pencegahan stunting sejak masa prenatal.

#### DAFTAR REFERENSI

- Alam Fajar, N., Liberty, I., Sri Ananingsih, E., & Nur Amalia, R. (2025). Utilizing maternal height as a predictor for childhood stunting prevention: A health promotion strategy rooted in early risk identification. *Media Publikasi Promosi Kesehatan Indonesia (MPPKI)*, 8(11), 1497–1505. <https://doi.org/10.56338/mppki.v8i11.8217>
- Alejandra, M.-H., Ricardo, E.-T., & Guadalupe, D.-L. (2025). Crosstalk between hypertension and diabetes: Focusing on pregnancy and offspring. A systematic review. *Frontiers in Physiology*, 16. <https://doi.org/10.3389/fphys.2025.1519410>
- Amran, R., Nurwiyeni, Pratama, R. R., & Wahyuni, S. (2025). Stunting sebagai ancaman kualitas sumber daya manusia: Perspektif gizi, lingkungan, dan sosial. *Scientific Journal*, 4(4), 233–240. <https://doi.org/10.56260/sciena.v4i4.234>
- Aulia, F. O., & Purnamawati, D. (2025). Kontribusi kader posyandu dalam intervensi stunting: Kajian literatur tahun 2023–2025. *Jurnal Ilmu Medis Indonesia*, 4(2), 155–168. <https://doi.org/10.35912/jimi.v4i2.4687>

- Díaz-Torres, S., Díaz-López, A., & Arija, V. (2024). Effect of prenatal iron supplementation adapted to hemoglobin levels in early pregnancy on fetal and neonatal growth: ECLIPSES study. *Nutrients*, *16*(3). <https://doi.org/10.3390/nu16030437>
- Hendraswari, C. A., Pramatirta, A. Y., Susiarno, H., Hilmanto, D., Zuhairini, Y., & Sahiratmadja, E. (2025). The correlation of calcium and hemoglobin level among pregnant women in  $\geq 35$  weeks with the anthropometry of newborn babies. *International Journal of Women's Health*, *17*, 507–515. <https://doi.org/10.2147/IJWH.S491476>
- Irma, Rahman, S. N., Ryska, H., & Hadju, V. (2026). Determinants of chronic energy deficiency (CED) in pregnant women in stunting. *Jurnal Ilmiah Kesehatan Sandi Husada*, *14*(1), 201–210. <https://doi.org/10.35816/jiksh.v14i1.324>
- Joshi, S., Sharma, L., Barde, L., Tare, M., Baheti, D., Dama, G., & Tare, H. (2023). The nutritional needs of mothers and babies: A review. *International Journal of Pharmaceutical Quality Assurance*. <https://doi.org/10.25258/ijpqa.14.2.30>
- Karypidis, E., Mouslech, S. G., Skoulariki, K., & Gazis, A. (2022). Comparison analysis of traditional machine learning and deep learning techniques for data and image classification. *WSEAS Transactions on Mathematics*, *21*, 122–130. <https://doi.org/10.37394/23206.2022.21.19>
- Khan, A., Ali, A., Khan, J., Ullah, F., & Faheem, M. (2025). Using permutation-based feature importance for improved machine learning model performance at reduced costs. *IEEE Access*, *13*, 36421–36435. <https://doi.org/10.1109/ACCESS.2025.3544625>
- Kustanto, A., Rachmat, O., & Setyadi, S. (2024). The prevalence of stunting in Indonesia: An examination of the health, socioeconomic status, and environmental determinants. *Journal of Iranian Medical Council*. <https://doi.org/10.18502/jimc.v8i1.17062>
- Laily, L. A., & Indarjo, S. (2023). Literature review: Dampak stunting terhadap pertumbuhan dan perkembangan anak. *HIGEIA (Journal of Public Health Research and Development)*, *7*(3), 354–364. <https://doi.org/10.15294/higeia.v7i3.63544>
- Marine, B. T., Haile, Y. A., & Zewde, M. G. (2026). Maternal nutrition practices and its implications for child growth and development in Jimma Town: A community-based cross-sectional study. *Scientific Reports*, *16*(1). <https://doi.org/10.1038/s41598-026-37151-4>
- Meliati, L., Sudarmi, S., & Halimatusyaadiah, S. (2025). Influence of the first thousand days of life on stunting incidence. *Journal of Holistic Nursing and Midwifery*, *35*(3), 161–167. <https://doi.org/10.32598/jhnm.35.3.2622>
- Muglia, L. J., Benhalima, K., Tong, S., & Ozanne, S. (2022). Maternal factors during pregnancy influencing maternal, fetal, and childhood outcomes. *BMC Medicine*, *20*. <https://doi.org/10.1186/s12916-022-02632-6>
- Nakshine, V. S., & Jogdand, S. D. (2023). A comprehensive review of gestational diabetes mellitus: Impacts on maternal health, fetal development, childhood outcomes, and long-term treatment strategies. *Cureus*. <https://doi.org/10.7759/cureus.47500>
- Ni, W., Gao, X., Su, X., Cai, J., Zhang, S., Zheng, L., ... Zeng, F. (2023). Birth spacing and risk of adverse pregnancy and birth outcomes: A systematic review and dose-response meta-analysis. *Acta Obstetrica et Gynecologica Scandinavica*. <https://doi.org/10.1111/aogs.14648>

- Nurdivanuha, I., & Wibawani, S. (2025). The implementation and impact of the first 1000 days of life program on stunting prevention. *Journal of Social Studies Arts and Humanities (JSSAH)*, 5(1), 13–19. <https://doi.org/10.33751/jssah.v5i1.12840>
- Rao, B., Hasan, M. G., Putturaya, B., Kamath, A., Aatif, M., & Elmosaad, Y. M. (2026). Multidimensional correlates of childhood stunting in India: A spatial machine learning and explainable AI approach. *Stats*, 9(2), 34. <https://doi.org/10.3390/stats9020034>
- Ratnawati, R., Kartasurya, F. M. I., Kartini, A., & Noe, E. R. (2025). Prenatal risk factors for stunting in children aged 12–24 months in Central Java, Indonesia. *Bangladesh Journal of Medical Science*, 24(2). <https://doi.org/10.3329/bjms.v24i2.81535>
- Salih, Y., Omar, S. M., AlHabardi, N., & Adam, I. (2023). The mid-upper arm circumference as a substitute for body mass index in the assessment of nutritional status among pregnant women: A cross-sectional study. *Medicina*, 59(6). <https://doi.org/10.3390/medicina59061001>
- Salman, H. A., Kalakech, A., & Steiti, A. (2024). Random forest algorithm overview. *Babylonian Journal of Machine Learning*, 69–79. <https://doi.org/10.58496/BJML/2024/007>
- Salwa Alexita, A. C., Kusumaningtyas, P., & Rofi'i, M. (2025). Optimalisasi algoritma Random Forest menggunakan PSO untuk klasifikasi kanker payudara dengan citra mammograms. *Teknika STTKD: Jurnal Teknik, Elektronik, Engine*, 11(1), 47–54. <https://doi.org/10.56521/teknika.v11i1.1346>
- Setyaningsih, D., Wijayanti, H. N., Widayati, T., Susanti, S., & Universitas Respati Yogyakarta. (2024). Pengaruh karakteristik ibu terhadap kejadian stunting pada balita. *Jurnal Untuk Masyarakat Sehat (JUKMAS)*, 8(2). <https://doi.org/10.52643/jukmas.v8i2.4420>
- Sheth, V., Tripathi, U., & Sharma, A. (2022). A comparative analysis of machine learning algorithms for classification purpose. *Procedia Computer Science*, 215, 422–431. <https://doi.org/10.1016/j.procs.2022.12.044>
- Suharni, S., Nurtanny, N., & Ardian, R. (2025). Intervensi gizi pada anak indikasi kurang gizi berbasis pangan lokal di wilayah Kayu Jati Tembilahan Hulu. *Jurnal Pengabdian Kesehatan Komunitas*, 5(1), 1–10. <https://doi.org/10.25311/jpkk.Vol5.Iss1.2180>
- Valverde-Pérez, E., Olea, E., Rocher, A., Aaronson, P. I., & Prieto-Lloret, J. (2025). Effects of gestational intermittent hypoxia on the respiratory system: A tale of the placenta, fetus, and developing offspring. *Journal of Sleep Research*, 34(4). <https://doi.org/10.1111/jsr.14435>
- Vinita Yadav, K., & Kumar, R. (2022). Data preprocessing techniques. *Phoenix: International Multidisciplinary Research Journal*, 1.
- Vlachou, M., Kyrkou, G., Georgakopoulou, V., Kapetanaki, A., Vivilaki, V., Spandidos, D., & Diamanti, A. (2025). Smoke signals in the genome: Epigenetic consequences of parental tobacco exposure (Review). *Biomedical Reports*, 23(3), 1–9. <https://doi.org/10.3892/br.2025.2024>
- Yuan, X., Liu, S., Feng, W., & Dauphin, G. (2023). Feature importance ranking of Random Forest-based end-to-end learning algorithm. *Remote Sensing*, 15(21), 5203. <https://doi.org/10.3390/rs15215203>