

Prediksi Risiko Stunting pada Balita menggunakan Algoritma Logistic Regression dan Decision Tree berbasis Data Terbuka

Diana Yusuf^{1*}, Fahrul Razi², Shevti Arbekti Arman³, Vany Terisia⁴, Revi Nurjayanti⁵

^{1,5} Program Studi Sistem Informasi Fakultas Teknik dan Desain

^{2,3,4} Program Studi Teknologi Informasi Fakultas Teknik dan Desain

Institut Teknologi dan Bisnis Ahmad Dahlan, Jl. Ir. H. Juanda No. 77, Ciputat, Kota Tangerang Selatan

*Corresponding Author : dianayusuf01@gmail.com

Abstrak

Stunting masih menjadi tantangan kesehatan utama yang berdampak pada pertumbuhan dan perkembangan anak di Indonesia. Upaya deteksi dini risiko stunting memerlukan pendekatan berbasis data yang akurat dan praktis. Penelitian ini bertujuan mengembangkan dan membandingkan dua model klasifikasi, yaitu Logistic Regression dan Decision Tree, untuk memprediksi risiko stunting pada balita dengan memanfaatkan data kesehatan terbuka. Dataset yang digunakan mencakup variabel usia, berat lahir, berat badan, panjang badan, dan jenis kelamin. Proses penelitian meliputi preprocessing data, pemabagian data menjadi data latih dan data uji, penerapan model, serta evaluasi performa menggunakan metrik akurasi, precision, recall, dan F1-Score. Hasil penelitian menunjukkan bahwa Logistic Regression memberikan performa lebih stabil dengan akurasi 84,4% pada data latih dan uji. Decision Tree memiliki akurasi lebih tinggi pada data latih (96,5%) namun menurun pada data uji (78,7%), menunjukkan kecenderungan overfitting. Visualisasi Decision Tree mengungkapkan bahwa usia dan berat badan menjadi fitur paling dominan dalam klasifikasi risiko stunting. Berdasarkan hasil tersebut, Logistic Regression direkomendasikan sebagai model yang lebih andal untuk implementasi prediksi stunting di tingkat layanan kesehatan masyarakat. Temuan ini diharapkan dapat memberikan kontribusi bagi pengembangan sistem pendukung keputusan berbasis data dalam mitigasi stunting.

Kata kunci: Stunting, Logistic Regression, Decision Tree, Prediksi Risiko, Machine Learning

Abstract

Stunting remains a major public health issue impacting children's growth and development in Indonesia. Early detection of stunting risk requires accurate and practical data-driven approaches. This study aims to develop and compare two classification models, Logistic Regression and Decision Tree, to predict stunting risk among toddlers using open-source health data. The dataset includes age, birth weight, body weight, body length, and gender. The research process involved data preprocessing, data splitting into training and testing sets, applying both models, and evaluating performance using accuracy, precision, recall, and F1-Score. The result show that Logistic Regression demonstrated more stable performance, achieving 84,4% accuracy on both training and testing datasets. In contrast, the Decision Tree model achieved 96,5% accuracy on training data but decreased to 78,7% on testing data, indicating overfitting. The Decision Tree visualization revealed that age and body weight were the most influential features in classifying stunting risk. Based on these findings, Logistic Regression is recommended as a more reliable model for implementing stunting risk prediction at the community healthcare level. These result are expected to contribute to the development of data-driven decision support system in stunting prevention programs.

Keywords : Stunting, Logistic Regression, Decision Tree, Risk Prediction, Machine Learning

PENDAHULUAN

Latar Belakang

Stunting merupakan kondisi gagal tumbuh pada anak balita akibat kekurangan gizi kronis dan infeksi berulang, terutama pada periode 1.000 hari pertama kehidupan. Secara global, menurut Join Child Malnutrition Estimates (JME) 2023, sekitar 149 juta anak dibawah usia lima tahun mengalami stunting, dengan prevalensi tertinggi di Asia Selatan dan Sub-Sahara Afrika. Di Indonesia, meskipun terdapat penurunan prevalensi stunting dari 24,4% pada tahun 2021 menjadi 21,6% pada tahun 2022, angka tersebut masih berada diatas ambang batas yang ditetapkan oleh WHO, yaitu 20%.

Stunting tidak hanya berdampak pada pertumbuhan fisik anak, tetapi juga berpengaruh terhadap perkembangan kognitif, produktivitas ekonomi di masa depan, serta meningkatkan risiko penyakit degeneratif. Dampak jangka panjang ini menjadikan stunting sebagai ancaman serius terhadap kualitas sumber daya manusia dan pembangunan nasional. Oleh karena itu, diperlukan upaya deteksi dini dan intervensi yang tepat untuk mencegah dan mengurangi prevalensi stunting.

Berbagai penelitian telah dilakukan untuk memprediksi risiko stunting menggunakan pendekatan machine learning. Salah satu studi oleh Purwati dan Sulisty (2023) menggunakan algoritma K-Nearest Neighbor (K-NN) untuk klasifikasi stunting pada balita dan berhasil mencapai akurasi sebesar 92,4%. Penelitian lain oleh Angriani et al. (2023) menerapkan Support Vector Machine (SVR) dalam analisis big data stunting dan menunjukkan bahwa pendekatan ini efektif dalam mengidentifikasi faktor utama stunting, yang dapat membantu pengambilan keputusan klinis dan program intervensi.

Ditingkat internasional, Rao et al (2025) melakukan meta-analisis terhadap aplikasi machine learning dalam memprediksi malnutrisi anak menggunakan data Demographic and Health Surveys (DHS). Hasilnya menunjukkan bahwa model machine learning memiliki akurasi yang moderat hingga baik dalam memprediksi stunting pada anak di bawah lima tahun. Selain itu, Zemariam et al. (2025) menggunakan berbagai algoritma machine learning untuk memprediksi stunting pada

remaja putri di Ethiopia dan menemukan bahwa Random Forest memiliki kinerja terbaik dengan akurasi 77% dan AUC 85%.

Meskipun berbagai studi telah menerapkan algoritma machine learning untuk prediksi stunting, penggunaan kombinasi algoritma Decision Tree dan Logistic Regression dengan data terbuka masih jarang dilakukan. **Penelitian ini bertujuan** untuk mengisi kesenjangan tersebut dengan membangun model prediksi risiko stunting yang sederhana namun efektif, serta dapat diinterpretasikan dengan baik. Penggunaan data terbuka juga memungkinkan replikasi dan validasi oleh peneliti lain, sehingga meningkatkan transparansi dan akuntabilitas penelitian.

Mengacu pada permasalahan dan tinjauan pustaka di atas, penelitian ini memiliki urgensi yang tinggi dalam upaya pencegahan dan penanggulangan stunting di Indonesia. Tujuan utama dari penelitian ini adalah membangun model prediksi risiko stunting pada balita menggunakan algoritma Decision Tree dan Logistic Regression berbasis data terbuka. Diharapkan hasil dari penelitian ini dapat memberikan kontribusi dalam pengambilan keputusan kebijakan kesehatan, serta menjadi dasar bagi pengembangan sistem deteksi dini stunting yang lebih akurat dan efisien.

Rumusan Masalah

Berdasarkan uraian latar belakang diatas maka rumusan masalah dari penelitian ini ialah:

1. Bagaimana membangun model prediksi risiko stunting pada balita menggunakan algoritma Decision Tree dan Logistic Regression dengan memanfaatkan data terbuka?
2. Sejauh mana tingkat akurasi dan kinerja dari masing-masing algoritma Decision Tree dan Logistic Regression dalam memprediksi status pada balita?
3. Faktor-faktor apa saja yang paling berpengaruh terhadap risiko stunting berdasarkan hasil analisis dari model prediktif yang dikembangkan?

Rencana Pemecahan Masalah

Masalah stunting pada balita merupakan isu kesehatan yang hingga saat ini masih menjadi perhatian serius di Indonesia dan di banyak negara berkembang. Stunting tidak hanya berdampak pada pertumbuhan fisik anak, tetapi juga pada perkembangan kognitif dan kesehatannya di masa depan. Dalam upaya membantu mitigasi risiko stunting secara preventif, diperlukan metode prediksi yang akurat dan berbasis data, sehingga dapat menjadi alat bantu pengambilan keputusan bagi tenaga kesehatan dalam skrining awal.

Penelitian ini bertujuan untuk mengembangkan model prediksi risiko stunting pada balita menggunakan algoritma Logistic Regression dan Decision Tree dengan memanfaatkan data terbuka yang tersedia. Kedua model dipilih karena memiliki karakteristik yang berbeda, di mana Logistic Regression dikenal sebagai model prediktif yang sederhana dan mampu menghasilkan generalisasi yang baik, sementara Decision Tree unggul dalam interpretabilitas dan kemudahan visualisasi aturan klasifikasi. Dengan membandingkan kedua model tersebut, diharapkan diperoleh rekomendasi model yang paling sesuai untuk digunakan dalam konteks prediksi risiko stunting berbasis data kesehatan fisik balita.

Metode yang digunakan dalam penelitian ini meliputi pengumpulan data sekunder dari sumber terbuka, pengolahan data meliputi pembersihan, encoding, dan normalisasi (untuk model Logistic Regression), pembagian data menjadi data latih dan data uji, serta penerapan algoritma Logistic Regression dan Decision Tree. Selanjutnya dilakukan evaluasi performa model menggunakan metrik akurasi, precision, recall, dan F1-Score, serta analisis visualisasi pohon keputusan untuk memahami pola klasifikasi yang terbentuk.

Landasan Teori

Stunting adalah kondisi gagal tumbuh pada anak balita akibat kekurangan gizi kronis menyebabkan tinggi badan anak lebih rendah dibandingkan standar usianya (Kemenkes RI, 2021). Kondisi ini berdampak jangka panjang terhadap kesehatan, pendidikan, dan produktivitas ekonomi seseorang di masa dewasa.

Algoritma Logistic Regression merupakan salah satu metode klasifikasi statistik yang banyak digunakan dalam prediksi

biner, seperti risiko penyakit atau kondisi kesehatan tertentu. Menurut Kusnadi et al. (2022), Logistic Regression memiliki keunggulan dalam mengolah data dengan variabel dependen biner dan memberikan interpretasi koefisien yang mudah dipahami. Selain itu, penelitian oleh Fitriani et al. (2023) menunjukkan bahwa Logistic Regression mampu memberikan prediksi risiko stunting dengan akurasi yang cukup baik pada studi kasus di wilayah Sumatera Barat.

Di sisi lain, Decision Tree merupakan sebuah teknik mengubah data ke dalam bentuk pohon keputusan berdasarkan rule untuk membagi beberapa populasi yang beragam menjadi lebih kecil dengan tetap memperhatikan variabel tujuannya (Diana, 2020). Decision Tree memiliki kelebihan dalam membentuk model yang transparan dan dapat dipahami dengan mudah oleh pengguna non-teknis. Penelitian yang dilakukan oleh Wulandari dan Putra (2023) menunjukkan bahwa Decision Tree mampu mengidentifikasi faktor risiko utama stunting pada balita, meskipun memiliki kecenderungan overfitting jika tidak dilakukan pruning atau pengaturan kedalaman pohon yang tepat.

Dalam konteks prediksi stunting berbasis data terbuka, penelitian oleh Nurhayati dan Rahmawati (2022) menekankan pentingnya memilih model yang tidak hanya akurat, tetapi juga mampu memberikan generalisasi yang baik terhadap data baru. Penelitian tersebut merekomendasikan kombinasi model interpretable dan model prediktif dalam pengembangan sistem pendukung keputusan mitigasi stunting.

Dengan mengacu pada beberapa penelitian sebelumnya, penelitian ini bertujuan untuk menguji kedua model tersebut menggunakan dataset yang telah disesuaikan dan diolah secara komprehensif, sehingga dapat memberikan rekomendasi model yang paling layak digunakan dalam mendukung program pencegahan stunting di lapangan.

METODE

Penelitian ini menggunakan pendekatan kuantitatif dengan metode deskriptif-prediktif, yang bertujuan untuk membangun model klasifikasi guna memprediksi risiko stunting pada balita. Proses penelitian ini diawali dengan pengumpulan data, dilanjutkan dengan

pengolahan dan analisis data menggunakan algoritma machine learning, yaitu Decision Tree dan Linear Regression (Hastomo et al., 2022).

Pengumpulan Data dan Sumber Data

Data yang digunakan dalam penelitian bersifat sekunder dan diperoleh dari sumber terbuka, yakni Kaggle, platform yang menyediakan berbagai dataset untuk keperluan penelitian dan pengembangan model analitik. Dataset yang dipilih berisi data karakteristik dan status gizi balita, terdiri dari 10.000 entri dan mencakup variabel seperti usia, jenis kelamin, berat lahir, panjang lahir, berat badan saat ini, tinggi badan saat ini, serta status stunting.

Pemilihan dataset ini didasarkan pada pertimbangan ketersediaan variabel yang relevan, ukuran data yang cukup besar untuk melatih model prediksi, serta kemudahan akses yang memungkinkan replikasi oleh peneliti lain. Tidak dilakukan pengumpulan data primer dalam bentuk observasi langsung, wawancara, maupun kusioner.

Proses Analisis Data

Analisis data dilakukan melalui beberapa tahap, dimulai dari pra-pemrosesan data (data preprocessing) (Yulianto et al., 2023), yang meliputi pembersihan, penanganan nilai kosong (missing values) (Hastomo et al., 2021), konversi variabel kategorikal menjadi bentuk numerik, dan pemisahan data ke dalam data latih (training set) dan data uji (testing set) dengan proporsi 70:30. Setelah data siap, dua algoritma klasifikasi diterapkan, yaitu:

1. Decision Tree

Digunakan untuk membangun model berbasis aturan keputusan yang mudah dipahami dan divisualisasikan

2. Logistic Regression

Digunakan untuk memprediksi kemungkinan kejadian stunting berdasarkan kombinasi variabel masukan secara statistik

Model yang dihasilkan kemudian dievaluasi menggunakan beberapa metrik pengukuran, yaitu akurasi, precision, recall, dan F1-score (Aini et al., 2023), untuk mengetahui

sejauh mana tingkat keberhasilan model dalam melakukan klasifikasi.

Hasil dari kedua model akan dibandingkan guna melihat algoritma mana yang memiliki performa terbaik dalam konteks prediksi risiko stunting. Selain itu, analisis lebih lanjut juga dilakukan untuk mengidentifikasi variabel-variabel mana yang paling berkontribusi terhadap klasifikasi risiko stunting, sehingga dapat memberikan wawasan praktis bagi upaya pencegahan stunting di masyarakat.

HASIL DAN PEMBAHASAN

Penelitian ini menghasilkan dua model klasifikasi untuk memprediksi resiko stunting pada balita, yaitu model Decision Tree dan Logistic Regression. Kedua model diuji pada dataset yang telah dibersihkan dan diproses, terdiri dari fitur-fitur seperti usia, jenis kelamin, berat lahir, berat badan, panjang badan, dan panjang lahir.

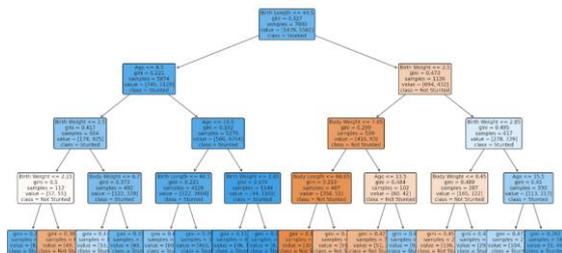
Model Decision Tree menunjukkan performa yang sangat tinggi pada data pelatihan dengan akurasi sebesar 96,5%, namun menurun menjadi 78,7% saat diuji pada data pengujian. Hal ini menunjukkan bahwa Decision Tree cenderung mengalami overfitting, yaitu terlalu menyesuaikan diri dengan data pelatihan sehingga kurang mampu melakukan generalisasi terhadap data baru. Metrik lainnya seperti precision dan recall juga mengalami penurunan dari data latih ke data uji.

Sebaliknya, model Logistic Regression menunjukkan performa yang lebih stabil antara data pelatihan dan pengujian. Akurasi pada data latih dan uji sama-sama berada di angka 84,4%, yang menunjukkan bahwa model ini memiliki kemampuan generalisasi yang baik. Nilai precision dan recall yang dihasilkan cukup seimbang, menjadikan Logistic Regression sebagai model yang lebih andal untuk konteks prediksi risiko stunting dalam data serupa. Perbandingan lengkap hasil evaluasi kedua model ditunjukkan pada Tabel 1 berikut ini:

Tabel 1. Perbandingan Hasil Evaluasi Model Logistic Regression dan Decision Tree

Model	Akurasi	Precision	Recall	F1 Score
Logistic Regression (Train)	84,4%	85,44%	96,85%	90,79%
Logistic Regression (Uji)	84,4%	85,83%	96,37%	90,79%
Decision Tree (Train)	96,5%	98,61%	97,01%	97,81%
Decision Tree (Uji)	78,7%	87,84%	85,09%	86,44%

Visualisasi model Decision Tree disajikan pada Gambar 1. Gambar ini memperlihatkan bagaimana model membentuk aturan keputusan dari setiap fitur untuk memisahkan antara balita yang stunted dan tidak stunted. Cabang pohon yang terbentuk memperlihatkan logika model dalam membagi data berdasarkan fitur usia, berat badan, dan panjang badan.



Gambar 1. Visualisasi Model Decision Tree

Analisis lebih lanjut terhadap hasil Decision Tree memperlihatkan bahwa fitur usia dan berat badan menjadi atribut yang paling dominan dalam membentuk aturan klasifikasi. Ini sejalan dengan penelitian sebelumnya yang menunjukkan bahwa usia dan berat badan balita menjadi indikator kuat dalam menilai status gizi (Nurhayati, 2022). Namun dalam penelitian ini ditemukan bahwa pengaruh jenis kelamin cenderung minor dalam pengambilan keputusan, yang sesuai dengan temuan dari Putri dan Ahmad (2023) yang menyatakan bahwa variabel jenis kelamin bukan prediktor utama stunting pada anak usia dini.

Jika dibandingkan dengan penelitian dari Setiawan dan Kusnadi (2021), akurasi model Logistic Regression yang dihasilkan dalam penelitian ini sedikit lebih tinggi. Setiawan dan Kusnadi (2021) melaporkan akurasi sebesar 82% dalam prediksi stunting menggunakan Logistic Regression, sementara penelitian ini mencapai 84,4%. Hal ini menunjukkan bahwa yang lebih bersih dan preprocessing yang optimal berkontribusi terhadap peningkatan performa model.

Dari segi implikasi praktis, model Logistic Regression lebih direkomendasikan untuk digunakan dalam implementasi prediksi risiko stunting pada aplikasi kesehatan masyarakat, mengingat kestabilan dan kemampuannya dalam menggeneralisasi data baru. Model ini dapat dengan mudah diintegrasikan dalam dashboard kesehatan atau

aplikasi mobile yang digunakan oleh tenaga kesehatan di lapangan. Namun demikian, model Decision Tree tetap memiliki keunggulan dari sisi interpretabilitas, terutama untuk edukasi dan pengambilan keputusan berbasis logika sederhana.

Penelitian ini memiliki beberapa keterbatasan. Dataset yang digunakan masih terbatas pada kesehatan fisik tanpa mempertimbangkan faktor lingkungan, sosial ekonomi, maupun pola asuh, yang juga menjadi determinan penting dalam kejadian stunting. Oleh karena itu, penelitian lanjutan diharapkan dapat mengintegrasikan data yang lebih komprehensif dan memperluas cakupan prediktor, sehingga model yang dihasilkan lebih akurat dan holistik.

Kedua model ini menunjukkan bahwa prediksi risiko stunting dapat dilakukan secara cukup akurat dengan menggunakan data terbuka dan teknik machine learning yang tepat. Temuan ini diharapkan dapat memberikan kontribusi dalam upaya mitigasi stunting, khususnya dalam mendukung pengambilan keputusan berbasis data di bidang kesehatan anak.

SIMPULAN DAN SARAN

Berdasarkan hasil penelitian yang telah dilakukan, dapat disimpulkan bahwa penggunaan model algoritma Logistic Regression dan Decision Tree mampu memprediksi risiko stunting pada balita dengan tingkat akurasi yang cukup baik. Model Logistic Regression menunjukkan performa yang lebih stabil pada data pelatihan maupun pengujian, dengan akurasi sebesar 84,4% di kedua kondisi tersebut. Hal ini menunjukkan kemampuan model dalam melakukan generalisasi terhadap data baru. Sementara itu, model Decision Tree menghasilkan akurasi yang lebih tinggi pada data pelatihan, yaitu 96,5%, namun mengalami penurunan performa pada data pengujian menjadi 78,7%. Penurunan ini mengindikasikan adanya kecenderungan overfitting, dimana model terlalu menyesuaikan diri dengan data pelatihan dan kurang efektif dalam mengolah data yang belum pernah dilihat sebelumnya.

Dalam visualisasi pohon keputusan, model Decision Tree mengungkapkan bahwa fitur usia dan berat badan menjadi atribut

paling dominan dalam menentukan klasifikasi risiko stunting. Meskipun Decision Tree memiliki interpretabilitas yang baik, namun dalam konteks prediksi stunting berbasis data terbuka, Logistic Regression direkomendasikan sebagai model yang lebih efektif dan efisien untuk diterapkan dalam sistem pendukung keputusan kesehatan masyarakat.

Berdasarkan hasil dan keterbatasan penelitian ini, beberapa saran yang dapat diajukan antara lain: pertama, pengembangan penelitian selanjutnya sebaiknya melibatkan variabel-variabel lain yang berkaitan dengan aspek sosial, ekonomi, pola asuh, dan lingkungan yang lebih luas, sehingga mampu meningkatkan akurasi prediksi dan memberikan gambaran yang lebih holistik mengenai faktor risiko stunting; kedua, penggunaan model Logistic Regression dapat diintegrasikan ke dalam aplikasi berbasis web atau mobile yang dapat diakses oleh tenaga kesehatan, guna mendukung skrining awal risiko stunting secara lebih cepat dan mudah. Ketiga, model Decision Tree dapat tetap dimanfaatkan sebagai alat edukasi untuk masyarakat atau tenaga kesehatan yang membutuhkan penjelasan logika sederhana dalam memahami risiko stunting, mengingat keunggulannya dari sisi interpretabilitas; keempat, diharapkan adanya kolaborasi antara peneliti data, tenaga kesehatan, dan pemerintah dalam mengembangkan sistem prediksi risiko stunting yang berbasis data real-time dan terintegrasi dengan program pencegahan stunting yang sedang berjalan.

Dengan demikian, hasil penelitian ini diharapkan dapat menjadi salah satu alternatif solusi berbasis data dalam upaya mendukung mitigasi stunting secara lebih efektif, efisien, dan berkelanjutan di masyarakat.

DAFTAR PUSTAKA

- Angriani, S., Jalil, N., Aminah, S., & Salim, N. A. (2023). Classification of Big Data Stunting Using Support Vector Regression and Voting Classifier. *Indonesian Journal of Artificial Intelligence and Data Mining*, 6(1), 1-10.
- Aini, N., Hastomo, W., & Yulika Go, R. (2023). Prediction of Anthropogenic Greenhouse Gas Emissions via Manure Management in Indonesia and Alternative Policies for Indonesian Livestock Development. *Journal of Renewable Energy and Environment*, 10(3), 99-106.
<https://doi.org/10.30501/jree.2022.35479>
6.1423
- Fitriani, D., Sutrisno, R., & Lestari, S. (2023). Prediksi Stunting Menggunakan Algoritma Logistic Regression Berbasis Data Kesehatan Balita di Sumatera Barat. *Jurnal Kesehatan Masyarakat Nasional*, 18(1), 87-95.
- Han, J., Kamber, M., & Pei, J. (2022). Data Mining: Concepts and Techniques. Morgan Kaufmann.
- Hastomo, W., Aini, N., Karno, A. S. B., & Rere, L. M. R. (2022). Machine Learning Methods for Predicting Manure Management Emissions. *Jurnal Nasional Teknik Elektro Dan Teknologi Informasi*, 11(2 SE-Articles).
<https://doi.org/10.22146/jnteti.v11i2.2586>
- Hastomo, W., Bayangkari Karno, A. S., Kalbuana, N., Meiriki, A., & Sutarno. (2021). Characteristic Parameters of Epoch Deep Learning to Predict Covid-19 Data in Indonesia. *Journal of Physics: Conference Series*, 1933(1), 012050.
<https://doi.org/10.1088/1742-6596/1933/1/012050>
- Kemenkes RI. (2021). Profil Kesehatan Indonesia 2021. Kementerian Kesehatan Republik Indonesia.
- Kumaidi. 2005. Pengukuran Bekal Awal Belajar dan Pengembangan Tesnya. *Jurnal Ilmu Pendidikan*. Jilid 5, No. 4,
- Kusnadi, B., Setiawan, R., & Dewi, M. (2022). Implementasi Logistic Regression untuk Prediksi Risiko Stunting pada Balita. *Jurnal Teknologi Informasi Kesehatan*, 10(2), 145-153.
- Nurhayati, N., & Rahmawati, E. (2022). Pemodelan Prediksi Stunting Menggunakan Pendekatan Machine Learning Berbasis Data Terbuka. *Jurnal Data Science Indonesia*, 5(2), 112-120.

- Purwati, N., & Sulisty, G. B. (2023). Stunting Early Warning Application Using KNN Machine Learning Method. *Jurnal Riset Informatika*, 5(3), 373-380
- Rao, B., Rashid, M., Hasan, M. G., & Thunga, G. (2025). Machine Learning in Predicting Child Malnutrition: A Meta-Analysis of Demographic and Health Surveys Data. *International Journal of Environmental Research and Public Health*, 22(3), 449.
- Wulandari, S., & Putra, Y. (2023). Decision Tree untuk Identifikasi Faktor Risiko Stunting pada Balita. *Jurnal Informatika dan Sistem Kesehatan*, 7(3), 221-230.
- Yusuf, D., Setri, E. (2020). Metode Decision Tree Dalam Klasifikasi Kredit Pada Nasabah PT Bank Perkreditan Rakyat (Studi Kasus : PT BPR Lubuk Raya Mandiri). *Jurnal Sistem Informasi (JUSIN)*, 1(1), 21-28.
- Yulianto, R., Faqihudin, Rusli, M. S., Karno, A. S. B., Hastomo, W., Kardian, A. R., Terisia, V., & Surawan, T. (2023). Innovative UNET-Based Steel Defect Detection Using 5 Pretrained Models. *Evergreen*, 10(4), 2365–2378.
<https://doi.org/10.5109/7160923>
- Zemariam, A. B., Yimer, A., Abebe, G. K., Wondie, W. T., Abate, B. B., & Alamaw, A. W. (2025). Prediction of Stunting and Its Socioeconomic Determinants Among Adolescent Girls in Ethiopia Using Machine Learning Algorithms. *PLOS ONE*, 20(1), e0316452.