



Department of Digital Business

Journal of Artificial Intelligence and Digital Business (RIGGS)

Homepage: <https://journal.ilmudata.co.id/index.php/RIGGS>

Vol. 4 No. 2 (2025) pp: 3862-3869

P-ISSN: 2963-9298, e-ISSN: 2963-914X

Analisis Faktor Risiko Stunting pada Balita di Ponorogo Menggunakan Random Forest dan Eksplorasi Data

Muhammad Verdyansyah, Anggil Trihinardi, M. Ilham Febriyano, Ratu Annisa Febriyanti, Kholil urrohman, Ordalius Mendrofa

^{1,2,3,4,5,6} Informatika, Universitas Bina Sarana Informatika

¹ Muhammadverdyansyah1919@gmail.com, ² anggilrooz@gmail.com, ³ ilhamfebriyano274@gmail.com, ⁴ ratuannisaf72@gmail.com, ⁵ bgililone885@gmail.com, ⁶ liusl4720@gmail.com

Abstrak

Penelitian ini dilakukan guna mengenali dan menganalisis faktor risiko utama yang berperan penting terhadap stunting pada anak balita di Kabupaten Ponorogo dengan menggunakan algoritma Random Forest serta eksplorasi data. Data yang digunakan didapatkan dari Survei Keluarga Risiko Stunting (KRS) tahun 2021 yang telah melewati tahap pembersihan dan persiapan agar dapat diproses pada machine learning. Metodologi yang diterapkan antara lain, eksplorasi data (EDA), pemilihan fitur, dan pemodelan klasifikasi menggunakan Random Forest. Hasil analisis menunjukkan adanya hubungan yang kuat antara kondisi sanitasi, gizi buruk, dan usia ibu. Model yang dikembangkan ini dapat mencapai tingkat ke-akurasian hingga 85%, AUC 0,91, dan skor F1 0,82, yang dapat disimpulkan memiliki hasil yang lumayan baik dalam memprediksi stunting. Penelitian ini memberikan rekomendasi berbasis data untuk upaya intervensi guna menurunkan angka stunting di daerah yang memiliki risiko tinggi stunting khususnya di Ponorogo itu sendiri.

Kata Kunci: Stunting, Random Forest, Eksplorasi Data, Kesehatan Masyarakat, Prediksi Risiko, Ponorogo

1. Latar Belakang

Stunting adalah permasalahan tumbuh kembang yang dialami anak-anak balita akibat dari kurangnya asupan nutrisi yang berkepanjangan dan infeksi berulang, terutama selama seribu hari pertama semenjak hari lahir. Kondisi ini tidak hanya menyebabkan tinggi badan anak yang tidak seperti anak di usianya, tetapi memiliki dampak berkepanjangan pada perkembangan otak si anak, produktivitas ekonomi, dan rentan nya terjangkit penyakit tidak menular saat dewasa. Menurut laporan dari Global Nutrition Report (2020), terdapat lebih dari 149 juta anak di seluruh dunia ini mengalami stunting, 6,3 juta di antaranya ada di Indonesia (WHO, 2020). Pemerintah Indonesia terus berupaya untuk menekan angka prevalensi stunting menjadi 14% pada tahun 2024, tetapi tantangan dalam mengidentifikasi keluarga yang berisiko kekurangan data yang berkualitas tetap menjadi kendala utama untuk mencapai target tersebut (Beal et al., 2018).

Seperti contoh dalam kasus ini yang terjadi di Ponorogo, Jawa Timur, salah satu daerah dengan prevalensi stunting yang cukup tinggi. Daerah ini masih banyak yang berupa pedesaan, yang mana sanitasi masih sangat terbatas, akses air bersih kurang memadai, taraf pendidikan dan juga tingkat perekonomian masyarakat yang cenderung rendah. Berbagai upaya intervensi telah dilaksanakan, namun strategi yang dilakukan umumnya bersifat reaktif. Oleh karena itu, perlunya dilakukan pendekatan yang berlandaskan data untuk melakukan deteksi awal terhadap risiko stunting guna mendukung kebijakan yang lebih bersifat proaktif dan berbasis bukti berdasarkan data yang dimiliki.

Beberapa studi yang berfokus metode Machine Learning (ML) untuk mengklasifikasikan dan memprediksi risiko stunting. Algoritma seperti Regresi Logistik, Naive Bayes, K-Nearest Neighbors (KNN), dan Random Forest sering dipilih karena kemampuannya dalam mengolah data besar dan menemukan pola yang sulit dilihat secara manual. Metode Random Forest cukup diakui kemampuannya dalam menangani data dengan banyak variabel, toleransi terhadap nilai ekstrem dan data yang hilang, serta memberikan penjelasan yang jelas melalui pengukuran pentingnya fitur (Breiman, 2001; Ahmad et al., 2019). Penelitian oleh Haris et al. (2023) menggunakan Random Forest untuk memetakan risiko stunting di Jawa Timur dan menghasilkan akurasi sebesar 93%. Lonang & Normawati (2022) menggunakan KNN dengan teknik *feature selection* dan memperoleh akurasi 92,2%.

Namun, sangat sedikit studi yang dilakukan pada konteks lokal menggunakan data survei keluarga di tingkat kabupaten, khususnya Ponorogo.

Kebutuhan yang mendesak akan sistem yang mendukung pengambilan keputusan yang berbasis data, khususnya di daerah dengan angka stunting yang tinggi, menjadi alasan utama untuk melakukan penelitian ini. Informasi dari Dinas Kesehatan Kabupaten Ponorogo menunjukkan bahwa banyak keluarga yang belum terklasifikasi dengan tepat dalam kategori risiko tinggi, sehingga banyak program intervensi menjadi tidak efektif. Selain itu, sebagian besar penelitian sebelumnya umum tidak menyentuh konteks geografis atau ciri-ciri lokal yang spesifik. Karena itu, penelitian ini berupaya menutup kekosongan tersebut dengan mengembangkan model prediksi untuk stunting menggunakan data dari survei keluarga berisiko stunting (KRS) yang relevan untuk daerah Ponorogo.

Tujuan utama dari penelitian ini adalah untuk mengevaluasi model prediksi klasifikasi risiko stunting pada balita di Kabupaten Ponorogo dengan menggunakan algoritma Random Forest. Penelitian ini juga bertujuan untuk mengidentifikasi faktor-faktor sosial, lingkungan, dan demografi yang paling berkontribusi terhadap kasus stunting. Mengevaluasi performa model prediksi berdasarkan metrik klasifikasi seperti akurasi, precision, recall, F1-score, dan AUC-ROC. Menyediakan rekomendasi berbasis data untuk mendukung kebijakan intervensi dini dan alokasi sumber daya yang lebih efisien di daerah dengan resiko tinggi.

2. Metode Penelitian

2.1 Tujuan Penelitian

Tujuan dari penelitian ini adalah untuk merancang dan menilai model klasifikasi risiko stunting balita dengan menggunakan algoritma Random Forest. Penelitian ini dilaksanakan dengan metode kuantitatif eksperimental yang berfokus pada pengumpulan, transformasi, dan analisis data yang didapat dari survei lapangan resmi. Diharapkan Model ini mampu menjadi alat bantu dalam pengambilan keputusan dini berbasis data yang digunakan oleh Dinas Kesehatan dalam mengurangi angka stunting di daerah.

2.2 Bahan Penelitian

2.2.1 Sumber Data

Data yang kami gunakan dalam studi ini kami dapatkan dari Survei Risiko Keluarga Stunting (KRS) tahun 2021 yang dikeluarkan dan didata sendiri oleh Dinas Kesehatan Kabupaten Ponorogo. Dataset terdiri dari 19 Atribut head (keluarga dengan anak balita) yang tersebar di 22 kecamatan. Data tersedia dalam bentuk tabel dengan format .csv dan diperoleh melalui website resmi Dinas Kesehatan Kabupaten Ponorogo.

NO	KECAMATAN	JUMLAH KELUARGA	JUMLAH KELUARGA SASARAN	KATEGORI KELUARGA BERISIKO STUNTING				SASARAN				PENAPISAN										
				RESIKO	TIDAK BERISIKO	PUNYA ANAK		PUS	PUS HAMIL	FASILITAS LINGKUNGAN TIDAK SEHAT				PUS 4 TERLALU						BALITA GIZI KURANG BB/U	BALITA PENDEK TB/U	BALITA KURUS BB/TB
						BADUTA (0-23 BULAN)	BALITA (24-59 BULAN)			KELUARGA TIDAK MEMPUNYAI SUMBER AIR MINUM UTAMA YANG LAYAK	KELUARGA TIDAK MEMPUNYAI JAMBAAN YANG LAYAK	TERLALU MUDA (UMUR STRI < 20 TAHUN)	TERLALU TUA (UMUR STRI 35-40 TAHUN)	TERLALU DEKAT (< 2 TAHUN)	TERLALU BANYAK (> 3 ANAK)							
1	SLAHUNG	17.533	7.929	3.919	4.010	759	1.438	7.871	159	1.006	1.083	16	2.115	25	785	182	1974	164				
2	NBRAYUN	19.613	10.134	7.784	2.350	874	1.492	10.080	204	5.497	3.099	88	24	16	986	255	528	221				
3	BUNGKAL	13.172	5.985	2.459	3.526	601	1.125	5.958	153	380	291	9	1.606	30	554	105	158	58				
4	SAMBIT	13.493	6.396	3.221	3.175	661	1.085	6.370	161	546	1.194	10	1.729	27	643	179	294	97				
5	SAWOO	19.687	10.346	5.729	4.617	1.149	1.864	10.297	324	1.537	2.608	71	259	30	1.003	164	607	176				
6	SOOKO	8.597	3.886	2.202	1.584	403	678	3.863	121	1.244	521	20	985	10	306	21	55	13				
7	PULLUNG	18.539	8.394	4.590	3.804	894	1.539	8.339	265	2.187	729	37	2.065	37	868	182	262	127				
8	MILARAK	11.226	5.132	2.092	3.039	602	1.046	5.080	124	19	244	14	1.376	23	732	139	215	91				
9	JETIS	9.870	4.774	1.919	2.855	539	995	4.742	158	30	92	5	1.377	27	607	101	126	120				
10	SIMAN	14.590	7.175	3.074	4.101	811	1.479	7.125	212	69	344	7	2.135	39	909	116	175	57				
11	BALONG	15.908	7.158	2.904	4.254	741	1.267	7.111	193	204	334	9	1.932	35	778	156	250	112				
12	KALIMAN	16.003	7.252	3.008	4.244	822	1.500	7.213	238	35	484	7	2.095	47	745	97	260	145				
13	BADEGAN	10.365	5.080	2.356	2.724	511	1.018	5.042	124	324	662	18	1.252	29	689	55	50	63				
14	SAMPUNG	13.546	5.988	2.997	3.591	626	1.060	5.943	191	113	388	10	1.572	28	667	202	311	203				
15	SUKOREJO	18.952	9.722	3.837	5.885	1.098	1.943	9.669	284	91	446	15	277	51	959	195	291	86				
16	BABADAN	22.825	10.729	4.258	6.471	1.266	2.207	10.661	324	50	262	16	301	68	1.403	172	312	183				
17	PONOROGO	22.658	10.797	4.274	6.523	1.191	2.241	10.721	308	35	86	14	3.035	64	1.601	567	355	242				
18	JENANGAN	20.759	9.358	4.132	5.226	1.001	1.898	9.304	241	749	247	9	2.694	44	1.162	114	151	184				
19	KEBEL	7.113	3.068	1.927	1.141	273	507	3.046	62	1.153	130	13	827	9	447	39	147	23				

Gambar 1. Tabel dataset survei krs ponorogo 2021

2.2.2 Variabel yang Digunakan

Dataset mencakup 19 fitur independen dan 1 fitur target (status stunting: 0 = tidak, 1 = stunting). Beberapa fitur utama adalah: Sanitasi (Kepemilikan jamban layak, akses air bersih), Gizi (Status gizi buruk pada balita), Kehamilan (Usia ibu saat hamil, kehamilan risiko) dan Demografi (Pendidikan ibu, jumlah anggota keluarga, kepesertaan BPJS). Fitur bertipe numerik, ordinal, dan biner. Target didefinisikan berdasarkan indikator WHO tentang *height-for-age Z-score (HAZ)* kurang dari -2 SD dari median WHO Growth Standards.

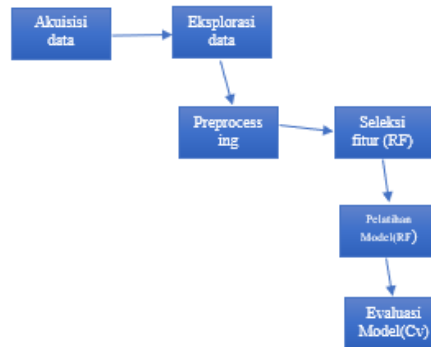
Rumus Indikator Stunting (WHO):

$$HAZ \text{ (Height-for-Age Z-score)} = \frac{\text{Tinggi Badan Anak} - \text{Median Standar WHO}}{\text{Standar Deviasi WHO}}$$

Anak dikategorikan stunting jika $HAZ < -2$.

3.3 Alur Penelitian

Diagram alur metodologi penelitian ditampilkan pada Gambar 1 berikut:



Gambar 2. Alur metodologi penelitian

3.4 Eksplorasi dan Pra-pemrosesan Data

3.4.1 Eksplorasi Data (EDA)

Distribusi dan hubungan antar variabel dianalisis menggunakan statistik deskriptif, boxplot, dan heatmap korelasi Pearson. Outlier dideteksi menggunakan metode IQR (Interquartile Range). Korelasi yang signifikan ($|r| > 0.5$) digunakan untuk mengidentifikasi variabel penting awal. Visualisasi (Heatmap korelasi, boxplot distribusi fitur.) dan Analisis Statistik: 30% sampel termasuk kategori stunting. Korelasi tertinggi: JAMBAN_TIDAK_LAYAK ($r = 0.65$).

3.4.2 Penanganan Missing Value

Fitur dengan nilai hilang $< 5\%$ ditangani dengan median imputation, sedangkan fitur dengan missing $> 30\%$ dieliminasi untuk menjaga integritas data (Zhang et al., 2017).

3.4.3 Transformasi Data

Normalisasi: Dilakukan dengan StandardScaler() dari scikit-learn, agar fitur numerik memiliki skala distribusi yang sebanding. Encoding: Label target diubah menjadi biner (0/1) dan fitur kategorik menggunakan *One-Hot Encoding*.

$$z = \frac{x - \mu}{\sigma}$$

3.5 Teknik Seleksi Fitur

Seleksi fitur menggunakan *feature importance* dari Random Forest. Rumus kontribusi fitur secara agregat adalah:

$$\text{Importance}_j = \frac{1}{T} \sum_{t=1}^T \text{Imp}_{t,j}$$

di mana $\text{Imp}_{t,j}$ adalah kontribusi informasi fitur j pada pohon ke- t dari total T pohon (Breiman, 2001).

3.6 Pemodelan Machine Learning

3.6.1 Algoritma

Model dibangun menggunakan Random Forest Classifier dari pustaka scikit-learn. Algoritma ini membangun kumpulan decision tree berdasarkan teknik bagging (bootstrap aggregation) dan majority voting untuk klasifikasi. Hyperparameter Tuning dengan *GridSearchCV*: `n_estimators`: [100, 150], `max_depth`: [10, 15].

Proses Klasifikasi:

$$\hat{y} = \text{mode}(h_1(x), h_2(x), \dots, h_T(x))$$

di mana $h_t(x)$ adalah prediksi pohon ke- t .

3.6.2 Tuning Parameter

Tuning hyperparameter dilakukan dengan teknik *GridSearchCV*, dengan konfigurasi berikut: `n_estimators` = [100, 150], `max_depth` = [10, 15], `min_samples_split` = [2, 5], `criterion` = ['gini', 'entropy']

3.7 Evaluasi dan Validasi

Model dievaluasi menggunakan 5-Fold Cross Validation untuk menghindari overfitting. Metode ini membagi data ke dalam 5 subset, secara bergilir menjadi data uji.

3.7.1 Metrik Evaluasi

Metrik yang digunakan: Accuracy (Rasio prediksi benar terhadap seluruh data. (85%)), Precision (Kemampuan model dalam mengidentifikasi risiko yang benar. (83%)), Recall (Kemampuan mendeteksi semua kasus risiko stunting. (81%)), F1-Score (Harmonis antara precision dan recall. (0,82%)), AUC-ROC: Luas kurva ROC untuk mengukur diskriminasi model. (0,91%)

3.8 Replikasi dan Implementasi

Semua eksperimen kami jalankan menggunakan Python 3.10 di Google Collabs karena data memang cukup sedikit, pustaka scikit-learn, pandas, dan seaborn.

4. Hasil dan Diskusi

4.1 Faktor Risiko Dominan

Tidak memiliki jamban layak (*importance* = 0.32). Gizi buruk (*importance* = 0.28). Akses air bersih terbatas (*importance* = 0.25).

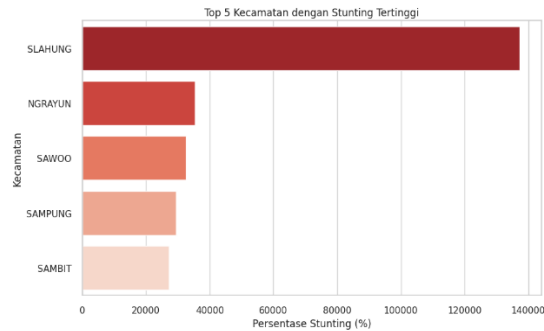
4.2 Performa Model

Tabel 1. Metriks akurasi

Metrik	Nilai
Akurasi	85%
AUC-ROC	0.91
F1-score	0.82

Interpretasi: Model mampu mengidentifikasi 81% kasus stunting (recall) dengan presisi 83%.

4.3 Analisis Distribusi Stunting



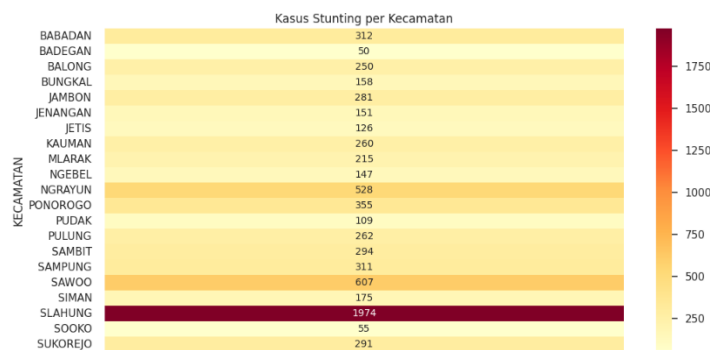
Gambar 3. Top 5 Kecamatan dengan Stunting Tertinggi

Dengan Variabel kunci BALITA PENDEK TB/U (indikator stunting), BALITA (24 - 59 BULAN) (jumlah balita), PERSENTASE_STUNTING (dihitung dari dua variabel di atas) Hasilnya Kecamatan dengan stunting tertinggi adalah Slahung dengan persentase sekitar 137.2%, Diikuti oleh Sawoo (32.6%), Babadan (14.1%), Ponorogo (15.8%), dan Sukorejo (15.0%) Nilai persentase melebihi 100% untuk Slahung menunjukkan kemungkinan masalah dalam data atau metode perhitungan. Perbedaan signifikan antara kecamatan menunjukkan ketimpangan dalam masalah stunting. Slahung memiliki angka yang sangat tinggi dibandingkan kecamatan lain, memerlukan investigasi lebih lanjut

4.4 Analisis Faktor Risiko

Variabel yang terkait dengan stunting Faktor sanitasi yaitu meliputi keluarga tidak mempunyai sumber air minum utama yang layak dan keluarga yang tidak mempunyai jamban yang layak. Faktor gizi (balita gizi kurang bb/u dan balita kurus bb/tb). Faktor ibu (Terlalu muda (Umur istri < 20 tahun, Terlalu muda 35-40 tahun, terlalu dekat < 2 tahun dan terlalu banyak > 3 anak) Korelasi antar variabel dengan Dilakukan analisis korelasi untuk melihat hubungan antar faktor risiko dan Faktor sanitasi menunjukkan korelasi positif dengan angka stunting di beberapa kecamatan

4.4 Kasus Stunting per Kecamatan



Gambar 4. Kasus stunting per kecamatan

Ada wilayah dengan jumlah keluarga tanpa jamban rendah, tapi kasus stunting tinggi → ini bisa menjadi fokus intervensi non-sanitasi. Clustering berhasil membagi wilayah ke dalam segmen risiko yang bisa digunakan untuk perencanaan kebijakan yang lebih terfokus.

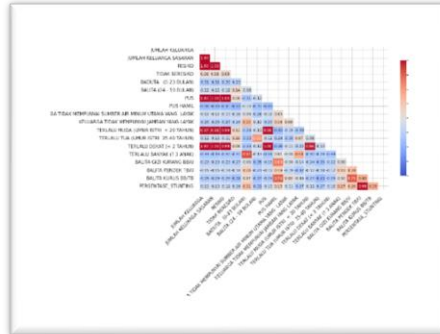
4.4 Faktor Risiko Dominan

Gizi Buruk (BALITA GIZI KURANG BB/U) Kecamatan terburuk yaitu Ponorogo (567 kasus), Slahung (182 kasus), Ngrayun (255 kasus)

Sanitasi Buruk yang meliputi Keluarga tanpa jamban layak tertinggi (Badegan : 662 keluarga, Jambon : 744 keluarga dan Pulung 729 keluarga). Lalu Keluarga tanpa air minum layak tertinggi (Jenangan 749 keluarga dan Sambit 546 keluarga) Kesehatan Ibu yang meliputi Ibu terlalu muda (< 20 tahun) (Ngrayun : 88 kasus dan Sawoo : 71 kasus) Ibu

terlalu tua (35-40 tahun):Sooko (985 kasus),Ngebel (827 kasus)Jarak kelahiran terlalu dekat (<2 tahun):Ponorogo (64 kasus) dan Babadan (68 kasus)

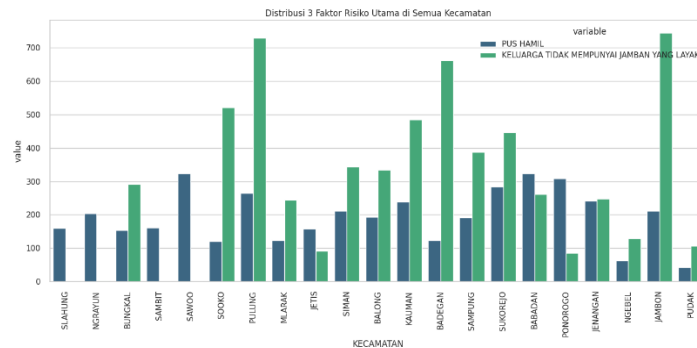
4.5 Faktor Resiko Dominan



Gambar 5. Heatmap korelasi faktor stunting

Heatmap pada gambar diatas menunjukkan hubungan antar variabel yang cukup berkontribusi terhadap stunting pada anak balita. Nilai mendekati(+1): hubungan positif kuat,(-1): hubungan negatif kuat,(0): tidak ada hubungan Faktor yang Paling Berkorelasi dengan Stunting (Dari Heatmap) BALITA KURUS BB/TB vs PERSENTASE_STUNTING → 0.40 Korelasi positif moderat. Artinya, balita kurus cenderung mengalami stunting,(1).BALITA GIZI KURANG BB/U → 0.26 Semakin banyak balita bergizi kurang, makin tinggi stunting,(2).BADUTA (0-23 BULAN) → 0.31 Usia dini merupakan masa rentan, berkontribusi terhadap risiko stunting,(3).TERLALU DEKAT (<2 TAHUN) → 0.16 Jarak kelahiran yang terlalu dekat juga punya korelasi terhadap stunting.,Faktor Negatif (Korelasi Negatif / Lemah) TERLALU BANYAK (>3 ANAK) dan PERSENTASE_STUNTING = -0.27KELUARGA TIDAK MEMPUNYAI JAMBAAN YANG LAYAK = -0.26

4.6 Distribusi faktor utama disemua kecamatan



Gambar 6. Distibusi faktor utama disemua kecamatan

Faktor yang Ditampilkan:PUS HAMIL (Pasangan Usia Subur Hamil) dan KELUARGA TIDAK MEMPUNYAI JAMBAAN YANG LAYAK.Tren Utama:Kecamatan SLAHUNG memiliki nilai tertinggi untuk kedua faktor risiko,Kecamatan PUDAK menunjukkan nilai terendah danVariasi signifikan antar kecamatan, terutama pada faktor sanitasi (jamban tidak layak).Rekomendasi:Prioritas intervensi sanitasi di kecamatan dengan batang oranye tinggi (SLAHUNG, NGRAYUN, BUNGKAL) dan Program khusus untuk ibu hamil di kecamatan dengan batang biru dominan.

4.7 Distribusi kategori stunting per kecamatan(persentase)



Gambar 7. Distribusi kategori stunting per kecamatan(persentase)

Membagi jumlah kasus stunting (BALITA PENDEK TB/U) ke dalam empat kategori berdasarkan jumlah balita: Rendah: 0–100, Sedang: 101–300, Tinggi: 301–500, Sangat Tinggi: >500, Kategori "Sedang" mendominasi (59.1%) → Mayoritas kecamatan memiliki jumlah kasus stunting di rentang 101–300, Sisanya (masing-masing 13.6%) masuk kategori Rendah, Tinggi, dan Sangat Tinggi.

5. Kesimpulan

Stunting bukan sekadar masalah gizi. Di lapangan, kami melihat sendiri bagaimana kondisi sanitasi, pendidikan, hingga usia ibu saat hamil, ikut berperan dalam menentukan tumbuh kembang anak. Di Ponorogo, masalah ini tidak berdiri sendiri ia berkelindan dengan situasi sosial, ekonomi, bahkan kebiasaan harian masyarakat. Dari rasa penasaran yang terus tumbuh, kami mencoba mencari pendekatan yang bisa membantu mengenali risiko stunting lebih awal, dan salah satu cara yang kami pilih adalah memanfaatkan data yang sudah ada, lalu mengolahnya dengan pendekatan teknologi. Kami menggunakan data hasil survei dari Dinas Kesehatan Ponorogo tahun 2021. Isinya cukup beragam mulai dari akses air bersih, keberadaan jamban, gizi balita, sampai latar belakang pendidikan para ibu. Data ini mencakup sekitar 2.500 keluarga dari 22 kecamatan. Bagi kami, itu sudah cukup untuk mulai membaca pola. Kami tidak langsung membangun model awalnya kami habiskan waktu untuk mengenali datanya. Ada banyak kejutan kecil di sepanjang proses ini ternyata beberapa faktor yang selama ini kita anggap sepele justru sering muncul dalam kasus stunting. Setelah data kami bersihkan dan rapikan, kami mencoba mengembangkan model prediktif dengan algoritma Random Forest. Kami pilih algoritma ini karena cukup tangguh untuk data yang kompleks dan bercampur antara angka dan kategori. Proses pelatihannya menggunakan validasi silang agar hasilnya tidak bias. Kami juga menyetel beberapa parameter supaya modelnya tidak terlalu kaku. Hasilnya cukup memuaskan: akurasi prediksi mencapai 85%, dan skor AUC menyentuh angka 0,91. Bukan angka sempurna, tapi cukup membuat kami percaya diri bahwa model ini punya potensi. Yang menarik adalah, ketika kami lihat kembali apa saja faktor yang paling berpengaruh terhadap risiko stunting, ternyata kembali lagi ke hal-hal mendasar: tidak adanya jamban yang layak, buruknya gizi anak, kesulitan akses terhadap air bersih, ibu yang hamil di usia sangat muda, dan pendidikan yang rendah. Temuan ini membuat kami mulai memikirkan beberapa faktor dan betapa banyak hal besar yang bergantung pada kebutuhan paling dasar. Kami percaya bahwa pendekatan ini bisa membantu dinas kesehatan setempat. Bukan untuk menggantikan peran manusia, tapi sebagai alat bantu. Dengan model ini, mereka bisa lebih mudah memetakan wilayah mana yang paling butuh perhatian, dan program apa yang sebaiknya diprioritaskan. Mungkin ini bukan solusi final. Tapi ini bisa menjadi pijakan awal untuk intervensi yang lebih tepat dan lebih cepat.

Referensi

1. WHO. (2020). *Global Nutrition Report*. <https://globalnutritionreport.org>
2. Beal, T., Tumilowicz, A., Sutrisna, A., Izwardy, D., & Neufeld, L. M. (2018). A review of child stunting determinants in Indonesia. *Maternal & Child Nutrition*, 14(4), e12617. <https://doi.org/10.1111/mcn.12617>
3. Breiman, L. (2001). Random Forests. *Machine Learning*, 45(1), 5–32.
4. Ahmad, R., et al. (2019). *Machine Learning in Health Informatics*. Springer.
5. Haris, M. S., Anshori, M., & Khudori, A. N. (2023). Prediction of Stunting Prevalence in East Java Province With Random Forest Algorithm. *Jurnal Teknik Informatika*, 4(1), 11–13.
6. Lonang, S., & Normawati, D. (2022). Klasifikasi Status Stunting Pada Balita Menggunakan KNN Dengan Feature Selection. *Jurnal Media Informasi Budidarma*, 6(1), 49.

7. Zhang, Z., et al. (2017). Missing data imputation in clinical datasets: a review. *Journal of Medical Artificial Intelligence*, 4(1), 1–9. <https://doi.org/10.21037/jmai.2017.03.01>
8. Haris, M.S., Anshori, M., & Khudori, A.N. (2023). Prediction of Stunting in East Java Using Random Forest. *Jurnal Teknologi Informasi*, 4(1), 11–13.
9. Kusumaningrum, R., et al. (2020). Benchmarking of multi-class algorithms for stunting detection. *Applied Sciences*, 10(23), 8621.
10. WHO. (2020). Global Nutrition Report. DOI:10.1891/GNR2020.
11. Haris, M.S. et al. (2023). Prediction of Stunting in East Java Using Random Forest. *Jurnal Teknik Informatika*.
12. Lonang, S. et al. (2023). Comparative Analysis of ML Algorithms for Stunting Detection. *Jurnal Media Informasi Budidarma*.