

# Sistem Pendukung Keputusan Probabilistik untuk Prediksi Rawat Inap Pasien Pneumonia Anak dengan NGBoost dan SHAP

Adam Rizky Ramadhan<sup>1</sup>, Amri Muhaimin<sup>2</sup>, Sugiarto<sup>3</sup>

<sup>1,3</sup> *Sains Data, Universitas Pembangunan "Veteran" Jawa Timur*

<sup>1</sup> [22083010064@student.upnjatim.ac.id](mailto:22083010064@student.upnjatim.ac.id)

<sup>3</sup> [sugiarto.if@upnjatim.ac.id](mailto:sugiarto.if@upnjatim.ac.id)

<sup>2</sup> *Sains Data, Universitas Pembangunan "Veteran" Jawa Timur*

Corresponding author email: [Amri.muhaimin.stat@upnjatim.ac.id](mailto:Amri.muhaimin.stat@upnjatim.ac.id)

## ABSTRAK

Abstrak— Pneumonia masih menjadi penyebab utama morbiditas dan mortalitas pada anak secara global, menuntut keputusan Rawat Inap yang cepat dan akurat. Penelitian ini bertujuan untuk mengembangkan Sistem Pendukung Keputusan (SPK) Probabilistik untuk memprediksi risiko Rawat Inap pasien pneumonia anak menggunakan model Natural Gradient Boosting (NGBoost). Metodologi penelitian mencakup praproses data klinis dan hematologi, pelatihan model NGBoost menggunakan distribusi Bernoulli untuk klasifikasi biner, dan evaluasi komprehensif. Hasil pengujian menunjukkan kinerja model yang superior dengan akurasi 0.95 (95%) dan yang paling krusial, Sensitivitas (Recall) sempurna 1.00 (100%) untuk kelas Rawat Inap. Model juga mencapai kemampuan diskriminasi yang baik (AUROC = 0.87). Kontribusi utama penelitian ini adalah integrasi prediktor probabilistik dengan SHapley Additive exPlanations (SHAP) untuk memastikan transparansi. Analisis SHAP mengidentifikasi tanggal pemeriksaan, kategori pneumonia, dan hematokrit sebagai fitur paling dominan. Penelitian ini memperlihatkan potensi NGBoost sebagai alat bantu klinis yang sangat sensitif dan dapat dijelaskan untuk meningkatkan objektivitas dan keselamatan dalam keputusan Rawat Inap pediatrik.

Keywords: Pneumonia, Anak, NGBoost, SHAP, Probabilistik

## I. PENDAHULUAN

Pneumonia adalah salah satu penyebab utama penyakit dan kematian pada anak-anak di seluruh dunia. Berdasarkan data dari Organisasi Kesehatan Dunia (WHO), pneumonia bertanggung jawab atas sekitar 14% dari total kematian anak balita di seluruh dunia, menjadikannya penyakit infeksi paling fatal di kelompok usia ini [1]. Di Indonesia, kasus pneumonia pada anak masih dianggap tinggi, seperti yang diungkapkan oleh Kementerian Kesehatan bahwa prevalensi dan jumlah rawat inap akibat pneumonia pada anak cenderung mengalami peningkatan setiap tahunnya [2]. Selain tingginya prevalensi, tantangan utama dalam penanganan pneumonia pada anak adalah menentukan kebutuhan rawat inap dengan cepat, tepat, dan objektif.

Dalam praktik klinis, keputusan untuk melakukan rawat inap biasanya ditentukan oleh kombinasi gejala klinis, hasil pemeriksaan pendukung, faktor komorbid, serta pengalaman subjektif dari tenaga kesehatan. Namun, variasi antar klinik dan keterbatasan sumber daya sering menimbulkan dua masalah utama: over-admission, yaitu merawat pasien yang seharusnya dapat dirawat jalan sehingga membebani kapasitas rumah sakit; dan under-admission, yaitu mengizinkan pasien pulang meskipun berisiko tinggi mengalami komplikasi [3]. Untuk mengurangi ketidakpastian itu,



sejumlah penelitian mulai menggunakan pendekatan machine learning sebagai alat bantu dalam pengambilan keputusan klinis, terutama dalam memprediksi tingkat keparahan pneumonia dan juga kebutuhan perawatan inap [4].

Penggunaan Machine Learning (ML) telah menjadi pendekatan yang efisien dalam menciptakan Sistem Pendukung Keputusan Klinis (SPK) untuk pneumonia pada anak. Tinjauan sistematis oleh Silva et al. (2020) menekankan penerapan berbagai algoritma dalam pemetaan dan peningkatan proses pengambilan keputusan medis yang berkaitan dengan penyakit pernapasan ini. Studi ini menekankan potensi besar ML, tetapi juga mencatat perlunya usaha tambahan untuk mengimplementasikan model-model tersebut secara praktis ke dalam proses kerja klinis [4]. Sesuai dengan tren ini, Wang et al. (2020) berhasil menciptakan model berbasis Gradient Boosting untuk mendeteksi sepsis pada anak-anak, suatu komplikasi serius pneumonia, yang menunjukkan akurasi tinggi dalam memprediksi kondisi kritis menggunakan data genetik [5].

Mempertimbangkan pentingnya model prediktif yang tidak hanya tepat tetapi juga jelas, penelitian ini mengutamakan interpretasi model. Dalam konteks yang sangat penting, Li et al. (2025) menerapkan metode SHAP (Shapley Additive exPlanation) untuk mengidentifikasi faktor-faktor klinis, seperti usia dan pola kejadian, yang paling berpengaruh dalam memprediksi patogen pneumonia pada anak yang dirawat di rumah sakit [6]. Keberhasilan studi-studi sebelumnya yang menerapkan model berbasis tree dan penggabungan interpretasi model menjadi pendorong utama. Oleh sebab itu, studi ini bertujuan untuk mengatasi kekurangan tersebut dengan menerapkan NGBoost, metode probabilistik, untuk menghasilkan prediksi rawat inap pneumonia pada anak yang disertai dengan estimasi ketidakpastian dan penjelasan yang solid menggunakan SHAP.

Di sisi lain, kemajuan dalam metode prediksi modern seperti Natural Gradient Boosting (NGBoost) menawarkan keuntungan berupa kemampuan untuk menghasilkan prediksi probabilistik yang terkalibrasi, sehingga memberikan estimasi risiko yang lebih tepat untuk pengambilan keputusan klinis [7]. Tidak seperti model boosting tradisional yang hanya memberikan prediksi kelas atau nilai tunggal, NGBoost memperkirakan distribusi probabilitas, sehingga memungkinkan dokter memahami sejauh mana ketidakpastian dalam prediksi. Namun, prediksi yang tepat saja tidak memadai—model harus bisa dijelaskan. Di sinilah SHapley Additive exPlanations (SHAP) menjadi sangat krusial. SHAP menawarkan suatu kerangka interpretasi yang didasarkan pada teori Shapley yang dapat menguraikan kontribusi setiap fitur terhadap hasil prediksi, baik secara global maupun individual [8]. Pendekatan ini semakin penting dalam konteks klinis yang memerlukan transparansi, akuntabilitas, dan keputusan yang berdasarkan bukti

Menyadari adanya peluang itu, penelitian ini menawarkan Sistem Pendukung Keputusan Probabilistik untuk Memperkirakan Rawat Inap Anak dengan Pneumonia menggunakan NGBoost dan SHAP. Sistem ini dibuat untuk menawarkan prediksi risiko rawat inap yang bukan hanya tepat dan terkalibrasi, tetapi juga dapat dijelaskan dengan jelas kepada tenaga medis. Dengan menggabungkan prediksi probabilistik dan pemahaman yang mendalam, studi ini diharapkan dapat meningkatkan objektivitas serta konsistensi dalam pengambilan keputusan klinis dan menjadi sumbangan signifikan dalam penggunaan machine learning di layanan kesehatan anak

## II. METODOLOGI PENELITIAN

### 2.1. Dataset

Dataset dalam penelitian ini adalah data sekunder yang diambil dari rekam medis anak yang didiagnosis pneumonia di RSUD Haji Surabaya pada rentang waktu Januari tahun 2024 – Oktober tahun 2025. Dataset ini dibagi menjadi variabel prediktor klinis dan laboratoriumnya (X) serta variabel keluaran biner (Y) yang merupakan target untuk prediksi. Variabel prediktor (X) terdiri dari tujuh penanda kuantitatif dalam bidang hematologi dan inflamasi, yaitu: hb, leukosit, hematokrit, trombosit, neutrofil, limfosit, crp, serta gda. Di samping itu, dua variabel kategorikal (gda\_stik dan kategori\_pneumonia) juga ditambahkan, di mana gda\_stik membedakan cara pemeriksaan glukosa darah dan kategori\_pneumonia membagi jenis pneumonia menjadi tipikal dan atipikal. Semua variabel ini ditujukan untuk meramalkan variabel target hospitalized (status Rawat Inap). Rincian dan penjelasan statistik setiap variabel yang diterapkan dalam pemodelan tertera pada Tabel 1 dibawah ini.

Table 1 Variable Dataset

Variable	Tipe	Keterangan
Hb	Numerik	Menggambarkan kadar hemoglobin dalam darah (g/dL).
Leukosit	Numerik	Jumlah total leukosit (sel darah putih) dalam darah ( $\times 10^3/\mu\text{L}$ ).
Hematokrit	Numerik	Persentase volume sel darah merah terhadap total volume darah (%).
Trombosit	Numerik	Jumlah trombosit dalam darah ( $\times 10^3/\mu\text{L}$ ).
Neutrofil	Numerik	Presentase neutrofil (%)
Limfosit	Numerik	Presentase limfosit (%)
GDA	Numerik	Kadar gula darah yang diukur secara acak (mg/dL).
CRP	Numerik	Penanda inflamasi (mg/L).
GDA_STIK	Kategorik	Variabel indikator yang menunjukkan apakah nilai GDA diperoleh melalui pemeriksaan

		cepat menggunakan metode stik (1 = GDA diukur dengan stik; 0 = GDA dari pemeriksaan laboratorium).
Usia	Numerik	Usia Pasien (dalam tahun)
Kategori_pneumonia	Kategorik	Variabel hasil pengerucutan klinis dari beberapa jenis pneumonia menjadi dua tipe utama: pneumonia typical dan pneumonia atypical
Hospitalized	Kategorik	Status apakah pasien dirawat inap (1) atau tidak (0).

## 2.2. Pre-processing Data

Proses prapemrosesan data yang penting dilakukan untuk menjamin kualitas serta konsistensi data masukan untuk model Machine Learning [9]. Pertama, variabel usia yang sebelumnya terpisah menjadi hari, bulan, dan tahun, disatukan menjadi satu variabel numerik tunggal yang merepresentasikan usia dalam tahun, dihitung dengan rumus  $\{Usia\ (Tahun)\} = \{Tahun\} + (\{Bulan\} / 12) + (\{Hari\} / 365)$  [10]. Kedua, variabel jenis pneumonia dikategorikan menjadi dua klasifikasi biner (kategori\_pneumonia), yaitu pneumonia tipikal dan pneumonia atipikal, guna mempermudah klasifikasi etiologi [11]. Terakhir, variabel yang memiliki nilai yang hilang diatasi melalui imputasi menggunakan median untuk fitur numerik dan modus untuk fitur kategorikal. Setelah proses pembersihan data, variabel prediktor (X) dipisahkan dari variabel target (hospitalized), yang berbentuk biner. Semua variabel kategorikal (termasuk gda\_stik dan kategori\_pneumonia) kemudian dikonversi menjadi format numerik dengan teknik Encoding agar bisa diproses oleh model NGBoost [12].

## 2.3. Pemodelan

Prediksi Rawat Inap pasien pneumonia anak dimodelkan menggunakan algoritma Natural Gradient Boosting (NGBoost). NGBoost dipilih secara khusus karena kemampuannya menghasilkan prediksi probabilistik yang terkalibrasi, yang sangat penting untuk memberikan estimasi risiko yang tepat dalam sistem pendukung keputusan klinis. Dalam konteks isu klasifikasi biner, NGBoost diset untuk memodelkan distribusi Bernoulli bersyarat dan memanfaatkan Log Score sebagai fungsi kerugian untuk melatih distribusi itu [13]. Proses pelatihan model dilakukan menggunakan data pelatihan dengan serangkaian hyperparameter yang dioptimasi;

informasi mendetail mengenai konfigurasi dan nilai hyperparameter yang dipakai disajikan dalam Tabel 2 berikut ini.

Table 2 Hyperparameter NGBost

Hyperparameter	Nilai	Keterangan
Distribusi (Dist)	Bernoulli	Digunakan untuk prediksi luaran biner
Fungsi Kerugian (Score)	LogScore	Fungsi yang dioptimalkan selama pelatihan.
Jumlah Estimator (n_estimators)	500	Jumlah pohon yang dibangun.
Laju Pembelajaran (learning_rate)	0.05	Ukuran langkah penyesuaian bobot.
Random State	42	Memastikan hasil dapat direproduksi.

#### 2.4. Evaluasi Model

Model NGBost dievaluasi menggunakan set data pengujian ( $X_{test}$ ) untuk mengukur kinerja prediksi dan kalibrasi probabilitasnya. Langkah awal mencakup peramalan kelas ( $y_{pred}$ ) dan pengambilan probabilitas Rawat Inap ( $y_{proba}$ ) untuk kelas positif (kelas 1). Kinerja klasifikasi dinilai menggunakan metrik standar seperti Akurasi, Presisi, Sensitivitas, dan Spesifisitas, yang disajikan dalam Laporan Klasifikasi dan Confusion Matrix. Metrik ini memberikan gambaran mengenai kemampuan model dalam mengklasifikasikan kasus Rawat Inap (True Positive) dan Rawat Jalan (True Negative) dengan tepat [14].

Selain evaluasi klasifikasi, kualitas prediksi probabilistik model NGBost diukur dengan Area Under the ROC Curve (AUROC). Nilai AUROC mencerminkan kemampuan model dalam membedakan pasien yang benar-benar dirawat inap dari yang tidak. Visualisasi kurva ROC (Receiver Operating Characteristic) juga ditampilkan untuk menunjukkan trade-off antara True Positive Rate dan False Positive Rate di berbagai ambang batas. Sebagai penunjang utama, diterapkan analisis SHAP untuk menginterpretasikan model, yang bertujuan untuk menjelaskan kontribusi khusus setiap fitur klinis terhadap prediksi risiko Rawat Inap, baik secara individu maupun agregat [15].

### III. HASIL DAN ANALISIS

Table 3 Classification Report

	Precision	Recall	F1-Score	Support
0	1.00	0.83	0.91	6

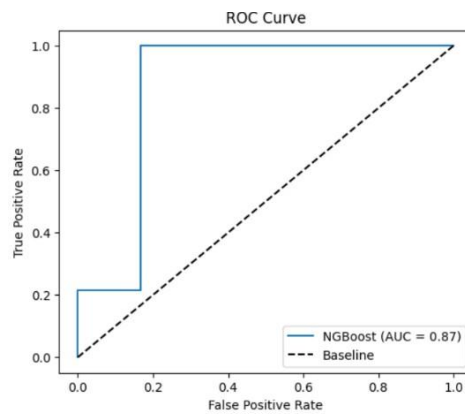
1	0.93	1.00	0.97	14
Accuracy			0.95	20
Macro avg	0.97	0.92	0.94	20
Weighted avg	0.95	0.95	0.95	20

Tabel 3 menyajikan ringkasan *Classification Report*, yang menunjukkan akurasi keseluruhan (accuracy) model sebesar 0.95 (95%) dan nilai F1-Score tertimbang (*weighted avg*) yang tinggi, yaitu 0.95. Kinerja model secara spesifik terlihat pada Recall (Sensitivitas) untuk Kelas 1 (Rawat Inap) yang mencapai sempurna 1.00 (100%), mengindikasikan bahwa model tidak pernah gagal mengidentifikasi pasien yang benar-benar memerlukan rawat inap. Kualitas klasifikasi ini diperkuat oleh Tabel 4 (*Confusion Matrix*), yang menunjukkan bahwa tidak ada kasus False Negative (FN = 0), sementara hanya terdapat satu kasus *False Positive* (FP = 1), menegaskan prioritas model pada keselamatan pasien dengan meminimalkan risiko terlewatnya penanganan kritis.

Table 4 Confusion Matrix

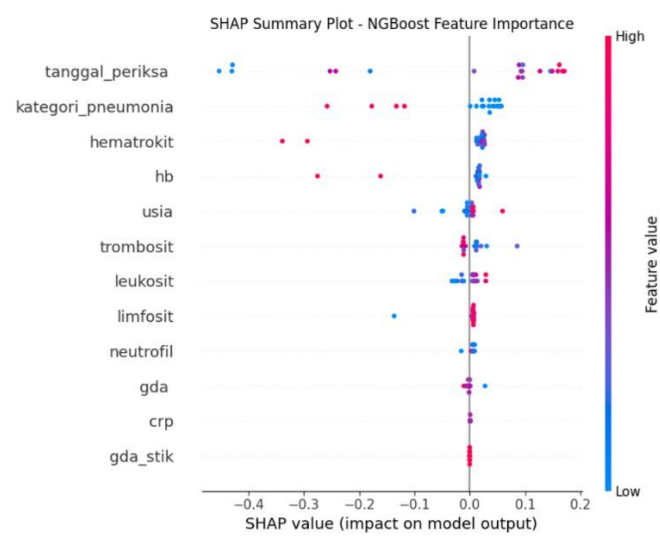
	Predictive Positive	Predictive Negative
Actual Positive	5	1
Actual Negative	0	14

Selain kinerja klasifikasi, model NGBost juga dinilai berdasarkan kemampuan diskriminasi probabilistiknya. Gambar 4.1 menampilkan ROC Curve (Receiver Operating Characteristic), yang mencerminkan trade-off antara True Positive Rate dan False Positive Rate pada berbagai ambang batas. Model menunjukkan kemampuan diskriminasi yang baik dengan nilai AUROC (Area Under the Curve) sebesar 0.87. Nilai AUROC ini mengkonfirmasi bahwa model sangat baik dalam membedakan probabilitas antara pasien Rawat Inap dan Rawat Jalan, menjadikannya andal sebagai dasar untuk prediksi risiko yang terkalibrasi.



Gambar 4. 1 ROC Curve

Untuk menjamin transparansi, model diinterpretasikan menggunakan kerangka SHapley Additive exPlanations (SHAP).



Gambar 4. 2 SHAP Summary Plot

Gambar 4.2 menampilkan SHAP Summary Plot , yang mengidentifikasi fitur klinis dengan dampak terbesar pada prediksi Rawat Inap. Fitur paling dominan yang memengaruhi luaran model adalah tanggal pemeriksaan, diikuti oleh kategori pneumonia (typical/atypical), hematokrit, dan hb. Secara umum, nilai fitur yang tinggi (warna merah) pada tanggal\_periksa dan kategori\_pneumonia cenderung meningkatkan nilai SHAP positif (mendukung prediksi Rawat Inap), sedangkan nilai hematokrit dan hb yang lebih rendah cenderung mendukung prediksi Rawat Inap (ditunjukkan oleh warna merah di sisi positif dan biru di sisi negatif). Analisis SHAP ini secara efektif menjelaskan mengapa model mencapai akurasi tinggi dan mengidentifikasi penanda klinis mana yang paling relevan untuk dipertimbangkan oleh dokter dalam keputusan rawat inap.

#### IV. KESIMPULAN

Penelitian ini berhasil merancang dan mengembangkan Sistem Pendukung Keputusan Probabilistik yang efisien untuk memprediksi risiko Rawat Inap pada anak dengan pneumonia menggunakan model Natural Gradient Boosting (NGBoost). Model yang diterapkan menunjukkan performa klasifikasi yang sangat baik, dengan akurasi keseluruhan 0.95 dan yang terpenting, sensitivitas (Recall) sempurna 1.00 (100%) untuk kategori Rawat Inap. Hasil ini menunjukkan bahwa model NGBoost dapat mengidentifikasi semua kasus pasien yang memerlukan penanganan Rawat Inap (False Negative = 0), mengutamakan keselamatan pasien di atas segalanya. Kualitas probabilistik dari model juga divalidasi dengan nilai AUROC sebesar 0.87, menegaskan kemampuan diskriminasi yang baik.

Kontribusi utama dari studi ini adalah penggabungan prediktor probabilistik dengan SHapley Additive exPlanations (SHAP), menjamin bahwa sistem tidak hanya tepat, tetapi juga sepenuhnya jelas. Analisis SHAP mengungkapkan bahwa elemen-elemen seperti tanggal pemeriksaan, kategori pneumonia, hematokrit, dan hb adalah faktor klinis terkuat yang mempengaruhi keputusan model. Transparansi ini krusial untuk membangun kepercayaan dari klinisi dan memvalidasi keputusan model yang didasarkan pada data medis yang relevan.

REFERENSI

- [1] Global Initiative for Asthma (GINA), “Pneumonia remains leading cause of child mortality,” ginasthma.org.
- [2] “PROFIL KESEHATAN INDONESIA 2022.”
- [3] P. J. Marcos *et al.*, “Discordance of physician clinical judgment vs. pneumonia severity index (PSI) score to admit patients with low risk community-acquired pneumonia: a prospective multicenter study,” *J Thorac Dis*, vol. 9, no. 6, pp. 1538–1546, Jun. 2017, doi: 10.21037/jtd.2017.05.44.
- [4] V. Silva, A. D. Ramos Novo, D. Souza, and A. Rêgo, “Machine Learning to Assist in Pneumonia Decision Making: A Systematic Review of the Literature,” in *Anais do Symposium on Knowledge Discovery, Mining and Learning (KDMiLe 2020)*, Sociedade Brasileira de Computação, Oct. 2020, pp. 201–208. doi: 10.5753/kdmile.2020.11977.
- [5] J. Ying, Q. Wang, T. Xu, and Z. Lu, “Diagnostic potential of a gradient boosting-based model for detecting pediatric sepsis,” *Genomics*, vol. 113, no. 1, pp. 874–883, Jan. 2021, doi: 10.1016/j.ygeno.2020.10.018.
- [6] T.-H. Chang *et al.*, “Clinical characteristics of hospitalized children with community-acquired pneumonia and respiratory infections: Using machine learning approaches to support pathogen prediction at admission,” *Journal of Microbiology, Immunology and Infection*, vol. 56, no. 4, pp. 772–781, Aug. 2023, doi: 10.1016/j.jmii.2023.04.011.
- [7] T. Duan *et al.*, “NGBoost: Natural Gradient Boosting for Probabilistic Prediction,” 2020.
- [8] G. Feretzakis *et al.*, “Integrating Shapley Values into Machine Learning Techniques for Enhanced Predictions of Hospital Admissions,” *Applied Sciences*, vol. 14, no. 13, p. 5925, Jul. 2024, doi: 10.3390/app14135925.
- [9] S.-A. N. Alexandropoulos, S. B. Kotsiantis, and M. N. Vrahatis, “Data preprocessing in predictive data mining,” *Knowl Eng Rev*, vol. 34, p. e1, Jan. 2019, doi: 10.1017/S026988891800036X.
- [10] R. M. Dreger, “A Basic Program for Converting Chronological Ages into Decimal for Averaging a Series of C As,” *Educ Psychol Meas*, vol. 50, no. 4, pp. 835–836, Dec. 1990, doi: 10.1177/0013164490504011.

- [11] G. I. Paterson, “Rough Classification of Pneumonia Patients using a Clinical Database,” 1994, pp. 412–419. doi: 10.1007/978-1-4471-3238-7\_48.
- [12] E. C. Gök and M. O. Olgun, “SMOTE-NC and gradient boosting imputation based random forest classifier for predicting severity level of covid-19 patients with blood samples,” *Neural Comput Appl*, vol. 33, no. 22, pp. 15693–15707, Nov. 2021, doi: 10.1007/s00521-021-06189-y.
- [13] Y. Raita *et al.*, “Machine learning-based prediction of acute severity in infants hospitalized for bronchiolitis: a multicenter prospective study,” *Sci Rep*, vol. 10, no. 1, p. 10979, Jul. 2020, doi: 10.1038/s41598-020-67629-8.
- [14] S. Melhem, A. Al-Aiad, and M. S. Al-Ayyad, “Patient care classification using machine learning techniques,” in *2021 12th International Conference on Information and Communication Systems (ICICS)*, IEEE, May 2021, pp. 57–62. doi: 10.1109/ICICS52457.2021.9464582.
- [15] J.-B. Excoffier, N. Salaün-Penquer, M. Ortala, M. Raphaël-Rousseau, C. Chouaid, and C. Jung, “Analysis of COVID-19 inpatients in France during first lockdown of 2020 using explainability methods,” *Med Biol Eng Comput*, vol. 60, no. 6, pp. 1647–1658, Jun. 2022, doi: 10.1007/s11517-022-02540-0.