

Prediksi Kebutuhan Rawat Inap Pasien Berdasarkan Data Hematologi Menggunakan Random Forest dengan Hyperparameter Optimization dan SHAP Explainability

Milla Akbarany Baktiar Putri^{a,1,*}, Hana Titania Sastrian^{b,2}, Muhammad Erlangga Kurniawan^{c,3}, Andri Fauzan Adziima^{d,4}, Kartika Maulida Hindrayani^{e,5}, Muhammad Zulhaj Aliansyah^{f,6}

^a Department of Data Science, Universitas Pembangunan Nasional "Veteran" Jawa Timur, Surabaya, Indonesia

¹ 23083010021@student.upnjatim.ac.id*; ² 23083010056@student.upnjatim.ac.id; ³ 23083010045@student.upnjatim.ac.id;

⁴ andri.fauzan.fasilkom@upnjatim.ac.id; ⁵ kartika.maulida.ds@upnjatim.ac.id; ⁶ muhammad_zulhaj.sada@upnjatim.ac.id

* Corresponding author

ARTICLE INFO	ABSTRACT
Article history Received 05 May 2026 Revised 10 May 2026 Accepted 20 June 2026 Published 28 June 2026 Keywords Hematology Hospitalization prediction Random Forest Electronic health records SHAP  License by CC-BY-SA Copyright © 2026, The Author(s).	Determining whether a patient requires inpatient or outpatient care is one of the most critical clinical decisions in healthcare facility management. Complete blood count (CBC) parameters provide routine and rapidly available physiological indicators, making them strong candidates for machine learning-based clinical decision support systems. This study aims to develop a prediction model for patient hospitalization needs using the Random Forest algorithm optimized through RandomizedSearchCV, combined with SHAP (SHapley Additive exPlanations) for model interpretability. The dataset consists of 4,412 patient records from an Indonesian private hospital available on Kaggle, containing nine hematological parameters and two demographic attributes. Feature engineering was applied by constructing three derived variables: leucocyte-thrombocyte ratio (<i>leu_throm_ratio</i>), hemoglobin-hematocrit ratio (<i>hb_hct_ratio</i>), and thrombocyte-age interaction (<i>platelet_age</i>). Two models were compared: a baseline Random Forest with manually configured hyperparameters and an optimized model using RandomizedSearchCV with Stratified 5-Fold Cross Validation. The baseline model achieved accuracy of .7633 and ROC AUC of .8097, while the optimized model yielded accuracy of .7644 and ROC AUC of .8131 on the test set. Feature importance analysis identified thrombocyte (.1655) and <i>leu_throm_ratio</i> (.1525) as the dominant predictors, together contributing over 31% of the model's predictive information. SHAP analysis confirmed that low thrombocyte values and high leucocyte values were the primary drivers of inpatient predictions, consistent with clinical indicators of infection and systemic inflammation. These findings demonstrate that Random Forest with clinically grounded feature engineering and SHAP interpretability provides a transparent and reasonably accurate approach for supporting hospital admission triage decisions.

How to cite: Putri, M. A. B., Sastrian, H. T., Kurniawan, M. E., Adziima, A. F., Hindrayani, K. M., & Aliansyah, M. Z. (2026). Prediksi Kebutuhan Rawat Inap Pasien Berdasarkan Data Hematologi Menggunakan Random Forest dengan Hyperparameter Optimization dan SHAP Explainability. *Journal of Computer Science and Information Technology*, Vol 2 (2), 98-109. doi: <https://doi.org/10.70716/jocsit.v2i2.560>

PENDAHULUAN

Penentuan status perawatan pasien, yaitu keputusan apakah seorang pasien memerlukan rawat inap atau cukup ditangani secara rawat jalan, merupakan salah satu keputusan klinis yang paling kritis dalam manajemen fasilitas kesehatan. Keputusan yang tidak tepat dapat berakibat pada penumpukan pasien di unit gawat darurat, penggunaan sumber daya rumah sakit yang tidak efisien, hingga meningkatnya risiko komplikasi akibat keterlambatan penanganan (Savioli et al., 2022). Pendekatan konvensional dalam triase pasien sangat bergantung pada penilaian subjektif tenaga medis yang dipengaruhi oleh pengalaman klinis, beban kerja, dan keterbatasan waktu. Kondisi ini mendorong kebutuhan terhadap sistem pendukung keputusan klinis yang objektif, cepat, dan dapat diandalkan.

Pemeriksaan Complete Blood Count (CBC) atau hematologi rutin merupakan salah satu prosedur diagnostik yang paling umum dilakukan di rumah sakit dan menghasilkan sejumlah parameter penting seperti hematokrit, hemoglobin, eritrosit, leukosit, trombosit, MCH, MCHC, dan MCV (Almuthaybiri et al., 2025). Parameter-parameter ini mencerminkan kondisi fisiologis pasien secara menyeluruh dan secara klinis berkaitan erat dengan berbagai kondisi patologis seperti infeksi, anemia, gangguan pembekuan darah, dan inflamasi sistemik (Steinbach et al., 2024). Ketersediaan data ini secara rutin dan dalam waktu singkat

menjadikannya kandidat yang sangat potensial sebagai dasar pengembangan model prediktif berbasis machine learning untuk mendukung keputusan klinis.

Perkembangan machine learning dalam bidang kesehatan telah menunjukkan hasil yang menjanjikan dalam berbagai aplikasi klinis, mulai dari prediksi risiko penyakit, pendeteksian anomali pada data laboratorium, hingga pengembangan sistem triase otomatis berbasis Electronic Health Record (EHR) (Busnatu et al., 2022; Fuse et al., 2025). Algoritma Random Forest secara khusus telah banyak diaplikasikan dalam konteks klasifikasi medis karena kemampuannya menangani data multivariat dengan dimensi tinggi, ketahanannya terhadap overfitting, serta kemampuannya menghasilkan estimasi feature importance secara langsung (Swarnadwip Bose & Bose, 2025). Beberapa penelitian telah membuktikan bahwa Random Forest mampu mengklasifikasikan kebutuhan rawat inap pasien dengan tingkat akurasi yang kompetitif dibandingkan algoritma lainnya (Yazdi et al., 2026; Yazdi et al., 2026). Selain itu, teknik optimasi hyperparameter seperti RandomizedSearchCV terbukti efektif dalam meningkatkan generalisasi model tanpa risiko overfitting yang signifikan dibandingkan exhaustive search (Tjahjono et al., 2025).

Meskipun demikian, sebagian besar model machine learning yang ada bersifat black box, yaitu tidak mampu menjelaskan alasan di balik setiap prediksi secara transparan kepada tenaga medis. Kurangnya interpretabilitas ini menjadi hambatan utama dalam adopsi model machine learning di lingkungan klinis, karena keputusan medis memerlukan justifikasi yang dapat dipertanggungjawabkan (Thamer & Sultani, 2025). Di sisi lain, studi-studi sebelumnya umumnya menggunakan fitur hematologi mentah tanpa mempertimbangkan hubungan biologis antarparameter yang secara klinis bermakna, seperti rasio leukosit terhadap trombosit sebagai indikator inflamasi atau interaksi antara usia dan jumlah trombosit (Liu et al., 2025). Kesenjangan ini menunjukkan perlunya pendekatan yang mengintegrasikan feature engineering berbasis pengetahuan klinis dengan analisis interpretabilitas model yang sistematis.

Penelitian ini bertujuan mengembangkan model prediksi kebutuhan rawat inap pasien berbasis data hematologi menggunakan algoritma Random Forest yang dioptimasi melalui RandomizedSearchCV dengan evaluasi Stratified 5-Fold Cross Validation. Kontribusi utama penelitian ini mencakup tiga aspek: pertama, penerapan feature engineering berbasis rasio antarparameter hematologi yang memiliki relevansi klinis, yaitu rasio leukosit-trombosit (`leu_throm_ratio`), rasio hemoglobin-hematokrit (`hb_hct_ratio`), dan interaksi trombosit-usia (`platelet_age`); kedua, optimasi hyperparameter secara sistematis untuk menghasilkan konfigurasi model yang lebih stabil; serta ketiga, analisis interpretabilitas model menggunakan pendekatan SHAP (SHapley Additive exPlanations) (Lundberg & Lee, 2017) untuk mengidentifikasi kontribusi setiap fitur terhadap keputusan prediksi, sehingga menghasilkan model yang tidak hanya akurat tetapi juga dapat dijelaskan secara transparan kepada tenaga medis.

METODE PENELITIAN

Desain Penelitian

Penelitian ini menggunakan pendekatan kuantitatif dengan metode klasifikasi berbasis machine learning untuk memprediksi kebutuhan rawat inap pasien berdasarkan hasil pemeriksaan hematologi. Model yang digunakan adalah Random Forest karena memiliki kemampuan yang baik dalam menangani data medis yang bersifat multivariat, tahan terhadap overfitting, serta mampu memberikan informasi mengenai tingkat kepentingan variabel (Peng et al., 2025). Untuk meningkatkan performa model, dilakukan optimasi hyperparameter menggunakan RandomizedSearchCV. Selain itu, pendekatan Explainable Artificial Intelligence (XAI) menggunakan SHAP (SHapley Additive exPlanations) diterapkan untuk meningkatkan interpretabilitas hasil prediksi (Ribeiro et al., 2024).

Sumber Data

Penelitian ini menggunakan data sekunder berupa dataset Electronic Health Record Predicting yang tersedia pada platform Kaggle. Dataset tersebut berasal dari sebuah rumah sakit swasta di Indonesia dan berisi data hasil pemeriksaan laboratorium pasien yang digunakan untuk menentukan kebutuhan perawatan lanjutan, yaitu rawat inap (in care) atau rawat jalan (out care).

Dataset terdiri atas 4.412 data pasien dengan 11 atribut, meliputi sembilan parameter hematologi, satu atribut demografis, dan satu atribut target. Distribusi kelas pada dataset terdiri dari 2.628 pasien rawat jalan dan 1.784 pasien rawat inap.

Variabel Analisis

Variabel analisis seperti pada Tabel 1 dibagi menjadi variabel dependen dan independen. Variabel dependen dalam penelitian ini adalah status perawatan pasien (*source*), yang diklasifikasikan menjadi dua kelas, yaitu rawat jalan (*out*) dan rawat inap (*in*). Variabel independen terdiri atas parameter hematologi dan karakteristik pasien, yaitu hematokrit (*hematocrit*), hemoglobin (*hemoglobin*), eritrosit (*erythrocyte*), leukosit (*leucocyte*), trombosit (*thrombocyte*), *Mean Corpuscular Hemoglobin* (MCH), *Mean Corpuscular Hemoglobin Concentration* (MCHC), *Mean Corpuscular Volume* (MCV), usia (*age*), dan jenis kelamin (*sex*).

Tabel 1. Deskripsi Atribut Dataset

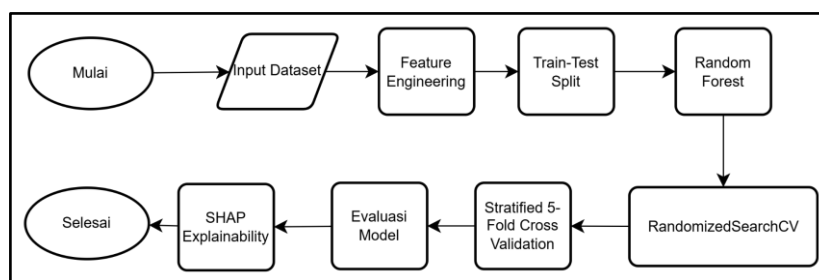
Atribut	Type Data	Deskripsi
Hematocrit	Numerik	Hasil pemeriksaan hematokrit pasien
Hemoglobin	Numerik	Hasil pemeriksaan hemoglobin pasien
Erythrocyte	Numerik	Jumlah eritrosit pasien
Leucocyte	Numerik	Jumlah leukosit pasien
Thrombocyte	Numerik	Jumlah trombosit pasien
MCH	Numerik	Mean Corpuscular Hemoglobin
MCHC	Numerik	Mean Corpuscular Hemoglobin Concentration
MCV	Numerik	Mean Corpuscular Volume
Age	Numerik	Usia pasien
Sex	Kategorikal	Jenis kelamin pasien
Source	Target	Status perawatan pasien (in/out)

Selain menggunakan variabel asli, penelitian ini juga menerapkan feature engineering dengan membentuk tiga variabel baru, yaitu rasio leukosit terhadap trombosit (*leu_throm_ratio*), rasio hemoglobin terhadap hematokrit (*hb_hct_ratio*), dan interaksi antara trombosit dan usia (*platelet_age*). Variabel tambahan tersebut digunakan untuk menangkap hubungan potensial antarparameter medis yang dapat memengaruhi keputusan rawat inap pasien.

Teknik Pengumpulan Data

Data diperoleh melalui proses pengunduhan dataset dari platform Kaggle. Karena penelitian menggunakan data sekunder yang telah dianonimkan, tidak dilakukan pengambilan data secara langsung kepada pasien maupun tenaga kesehatan. Seluruh data yang digunakan telah tersedia dalam format terstruktur sehingga dapat langsung digunakan untuk proses analisis.

Teknik Analisis Data



Gambar 1. Tahapan Penelitian

Tahap analisis pada Gambar 1 dimulai dengan proses preprocessing data. Variabel target dikonversi ke dalam bentuk numerik, yaitu nilai 0 untuk pasien rawat jalan dan nilai 1 untuk pasien rawat inap. Selanjutnya dilakukan imputasi median pada atribut numerik dan imputasi modus pada atribut kategorikal. Variabel kategorikal dikonversi menggunakan One-Hot Encoding, sedangkan atribut numerik dinormalisasi menggunakan StandardScaler.

Selain menggunakan atribut asli, penelitian ini menambahkan beberapa fitur baru yang dibangun berdasarkan hubungan antarparameter hematologi untuk meningkatkan kemampuan model dalam mengenali pola data. Fitur yang ditambahkan ditunjukkan pada Tabel 2.

Tabel 2. Fitur Hasil Rekayasa Fitur

Fitur	Persamaan
Leu_Throm_Ratio	Leucocyte / (Thrombocyte + 1)
Hb_Hct_Ratio	Hemoglobin / (Hematocrit + 1)
Platelet_Age	Thrombocyte × Age

Fitur Leu_Throm_Ratio digunakan untuk merepresentasikan hubungan antara jumlah leukosit dan trombosit yang berkaitan dengan kondisi inflamasi pasien. Fitur Hb_Hct_Ratio menggambarkan hubungan antara kadar hemoglobin dan hematokrit, sedangkan fitur Platelet_Age digunakan untuk mengakomodasi interaksi antara usia pasien dan jumlah trombosit (Zhai et al., 2021).

Dataset kemudian dibagi menjadi data pelatihan sebesar 80% dan data pengujian sebesar 20% menggunakan teknik stratified sampling untuk mempertahankan proporsi kelas pada kedua kelompok data. Metode klasifikasi yang digunakan dalam penelitian ini adalah Random Forest, yaitu algoritma ensemble learning yang membangun sejumlah decision tree secara independen menggunakan teknik bootstrap sampling. Prediksi akhir ditentukan berdasarkan mekanisme majority voting dari seluruh pohon keputusan yang terbentuk (Azad et al., 2025) .

Random Forest dipilih karena memiliki kemampuan yang baik dalam menangani data numerik dan kategorikal, relatif tahan terhadap overfitting, serta mampu memberikan informasi mengenai tingkat kepentingan setiap fitur (feature importance) (Lovatti et al., 2019). Model Random Forest dibangun menggunakan data pelatihan dan dioptimasi menggunakan RandomizedSearchCV dengan beberapa kombinasi hyperparameter, meliputi jumlah pohon (n_estimators), kedalaman pohon (max_depth), jumlah minimum sampel pada proses pemisahan (min_samples_split), jumlah minimum sampel pada node daun (min_samples_leaf), dan jumlah fitur yang digunakan pada setiap percabangan (max_features).

Evaluasi model dilakukan menggunakan Stratified 5-Fold Cross Validation. Kinerja model diukur menggunakan metrik Accuracy, Precision, Recall, F1-Score, dan Area Under Curve (AUC-ROC) (Staritz & Schdeel, 2025). Selain itu, confusion matrix digunakan untuk mengevaluasi kemampuan model dalam membedakan pasien yang memerlukan rawat inap dan rawat jalan. Untuk meningkatkan interpretabilitas model, dilakukan analisis Feature Importance dan SHAP (SHapley Additive exPlanations). Analisis ini digunakan untuk mengidentifikasi variabel yang paling berpengaruh terhadap keputusan prediksi model. Melalui visualisasi SHAP Summary Plot, dapat diidentifikasi fitur-fitur yang paling berpengaruh dalam menentukan kebutuhan rawat inap pasien, sehingga hasil prediksi tidak hanya akurat tetapi juga dapat dijelaskan secara transparan (Vimbi et al., 2024).

Validitas dan Keandalan Analisis

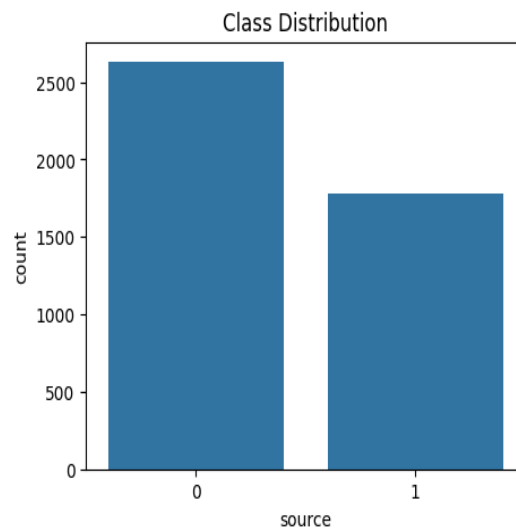
Validitas model dijaga melalui penggunaan Stratified 5-Fold Cross Validation yang memungkinkan proses evaluasi dilakukan pada beberapa subset data secara bergantian. Teknik ini membantu mengurangi bias yang dapat muncul akibat pembagian data pelatihan dan pengujian yang tidak representatif (Olamendy, J. C., 2024).

Keandalan model dievaluasi berdasarkan konsistensi nilai Accuracy, Precision, Recall, F1-Score, dan AUC-ROC pada setiap fold validasi. Selain itu, optimasi hyperparameter menggunakan RandomizedSearchCV dilakukan untuk memperoleh konfigurasi model yang lebih stabil dan optimal (Utami et al., 2026). Analisis SHAP juga digunakan untuk memverifikasi bahwa prediksi model didasarkan pada variabel-variabel hematologi yang relevan secara medis, sehingga meningkatkan transparansi dan kepercayaan terhadap hasil prediksi (Noor et al., 2025).

HASIL DAN PEMBAHASAN

Hasil Eksplorasi Data

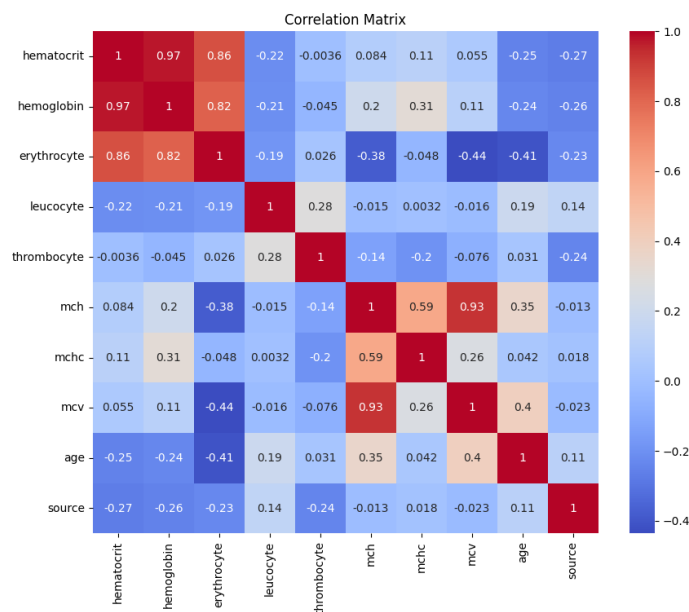
Distribusi kelas pada Gambar 2 di bawah ini menunjukkan bahwa terdapat 2.628 pasien rawat jalan dan 1.784 pasien rawat inap. Meskipun distribusi kelas tidak sepenuhnya seimbang, proporsi kedua kelas masih cukup representatif untuk dilakukan pemodelan klasifikasi.



Gambar 2. Distribusi Kelas Rawat Inap dan Rawat Jalan

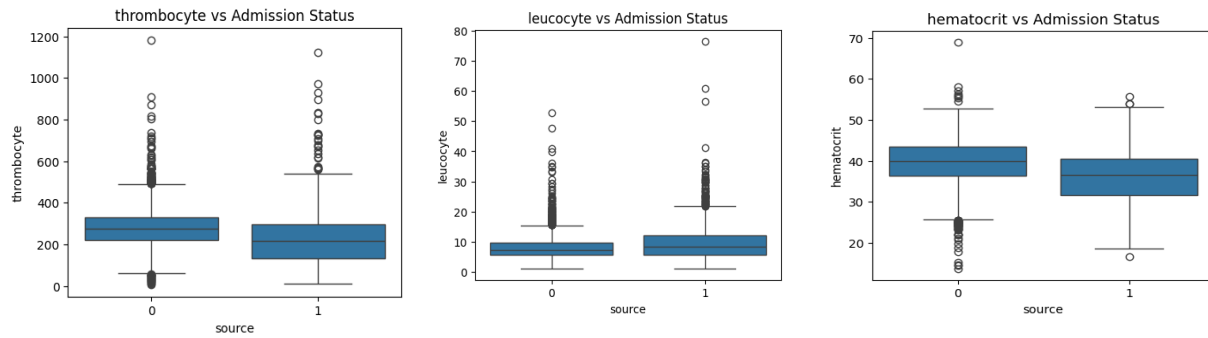
Distribusi kelas menunjukkan bahwa jumlah pasien rawat jalan lebih banyak dibandingkan pasien rawat inap. Kondisi ini dapat mempengaruhi proses pembelajaran model sehingga digunakan teknik `class_weight = balanced` pada algoritma Random Forest untuk mengurangi bias terhadap kelas mayoritas.

Analisis korelasi antar fitur numerik dilakukan menggunakan heatmap korelasi Pearson pada gambar 3 di bawah ini. Hasil analisis menunjukkan bahwa hemoglobin dan hematocrit memiliki korelasi positif yang kuat satu sama lain, yang konsisten secara fisiologis karena keduanya merupakan indikator volume dan konsentrasi sel darah merah. Thrombocyte dan leucocyte menunjukkan korelasi relatif rendah terhadap sebagian besar fitur lain, mengindikasikan bahwa keduanya membawa informasi prediktif yang independen.



Gambar 3. Matriks korelasi antar fitur numerik pada dataset hematologi.

Kemudian analisis distribusi fitur pada gambar 4 dibawah ini adalah berdasarkan status rawat inap menggunakan boxplot menunjukkan bahwa thrombocyte, leucocyte, dan hematocrit memiliki distribusi yang cenderung berbeda antara pasien rawat jalan dan rawat inap. Pasien rawat inap umumnya memiliki nilai thrombocyte lebih rendah dan leucocyte lebih tinggi dibandingkan pasien rawat jalan, yang mengindikasikan adanya kondisi inflamasi atau infeksi sistemik pada kelompok tersebut. Perbedaan distribusi ini memperkuat relevansi fitur-fitur hematologi sebagai prediktor kebutuhan rawat inap.



Gambar 4. Boxplot thrombocyte, leucocyte, hematocrit, age vs. status

Hasil Pemodelan

1. Model 1 Random Forest Baseline

Model pertama dibangun menggunakan algoritma Random Forest dengan konfigurasi hiperparameter yang ditentukan secara manual berdasarkan referensi literatur dan praktik umum, yakni $n_estimators = 500$, $max_depth = 20$, $min_samples_leaf = 2$, dan $class_weight = balanced$. Evaluasi model dilakukan menggunakan Stratified K-Fold Cross Validation dengan $k = 5$ pada data latih, kemudian dievaluasi kembali pada data uji yang terpisah. Hasil evaluasi Stratified 5-Fold Cross Validation pada Model 1 disajikan pada Tabel 3 berikut.

Tabel 3. Hasil Cross Validation Model Baseline Random Forest

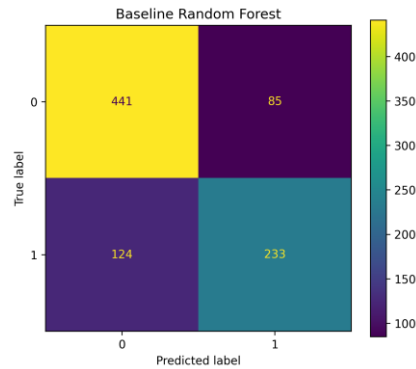
Metrik	Mean	Std
Accuracy	.7506	.0102
Precision	.7232	.0187
Recall	.6223	.0279
F1-Score	.6684	.0164
ROC AUC	.7996	.0055

Model baseline menghasilkan akurasi rata-rata sebesar .7506 dengan nilai ROC AUC sebesar .7996. Nilai ROC AUC yang mendekati .80 menunjukkan bahwa model memiliki kemampuan diskriminasi yang baik antara pasien rawat inap dan rawat jalan. Precision sebesar .7232 mengindikasikan bahwa dari seluruh prediksi positif (rawat inap), sekitar 72% di antaranya adalah prediksi yang benar. Namun, nilai Recall yang lebih rendah yaitu .6223 menunjukkan bahwa masih terdapat proporsi pasien rawat inap yang tidak berhasil diidentifikasi oleh model, yang dapat dikaitkan dengan karakteristik distribusi kelas yang tidak seimbang meskipun telah diterapkan $class_weight=balanced$. Evaluasi pada data uji menghasilkan performa sebagaimana disajikan pada Tabel 4.

Tabel 4. Hasil Evaluasi Model Baseline pada Data Uji

Kelas	Precision	Recall	F1-Score	Support
Rawat Jalan (0)	.7805	.8384	.8084	526
Rawat Inap (1)	.7327	.6527	.6904	357
Weighted Avg	.7612	.7633	.7607	883
ROC AUC	—	—	—	.8097

Hasil pada data uji menunjukkan bahwa model cenderung lebih baik dalam mengklasifikasikan pasien rawat jalan (kelas 0) dengan F1-Score .8084 dibandingkan pasien rawat inap (kelas 1) dengan F1-Score .6904. Hal ini merupakan pola umum pada dataset medis yang memiliki ketidakseimbangan kelas, di mana model cenderung lebih konservatif dalam memprediksi kelas minoritas.



Gambar 5. Confusion Matrix Baseline

Berdasarkan confusion matrix Gambar 5 di atas, model Random Forest mampu mengklasifikasikan sebagian besar pasien dengan benar. Hasil ini menunjukkan bahwa informasi dari pemeriksaan hematologi memiliki potensi untuk digunakan dalam membantu proses penentuan kebutuhan rawat inap atau rawat jalan. Meskipun masih terdapat beberapa kesalahan prediksi, performa model secara umum sudah cukup baik dalam membedakan kedua kelompok pasien.

2. Model 2 Random Forest dengan Hyperparameter Optimization

Model kedua dibangun menggunakan pendekatan yang sama namun dengan penerapan RandomizedSearchCV untuk mencari kombinasi hiperparameter optimal secara otomatis. Ruang pencarian hiperparameter yang didefinisikan mencakup: $n_estimators \in \{200, 300, 500, 700\}$, $max_depth \in \{10, 15, 20, 30, None\}$, $min_samples_split \in \{2, 5, 10\}$, $min_samples_leaf \in \{1, 2, 4\}$, dan $max_features \in \{\text{sqrt}, \text{log2}\}$. Pencarian dilakukan dengan 20 iterasi acak menggunakan F1-Score sebagai metrik optimasi, dengan validasi silang 5-Fold di dalam proses pencarian. Hasil pencarian mengidentifikasi kombinasi hiperparameter terbaik sebagaimana disajikan pada Tabel 5.

Tabel 5. Hiperparameter Optimal Hasil RandomizedSearchCV

Hiperparameter	Nilai Terbaik
$n_estimators$	500
max_depth	10
$min_samples_split$	5
$min_samples_leaf$	1
$max_features$	log2

Nilai $n_estimators$ yang optimal sama dengan konfigurasi baseline (500 pohon), namun max_depth yang dipilih lebih dangkal (10 dibandingkan 20 pada baseline). Hal ini mengindikasikan bahwa pohon yang lebih pendek dengan pembatasan kedalaman yang lebih ketat mampu menghasilkan generalisasi yang lebih baik, sekaligus mengurangi risiko overfitting akibat pohon yang terlalu dalam. Evaluasi model yang telah dioptimasi menggunakan Stratified 5-Fold Cross Validation menghasilkan performa sebagaimana disajikan pada Tabel 6.

Tabel 6. Hasil Cross Validation Model Random Forest Optimized

Metrik	Mean	Std
Accuracy	.7498	.0103
Precision	.7188	.0208
Recall	.6279	.0236
F1-Score	.6698	.0135
ROC AUC	.8012	.0067

Hasil evaluasi pada data uji untuk Model 2 disajikan pada Tabel 7 di bawah ini.

Tabel 7. Hasil Evaluasi Model Optimized pada Data Uji

Kelas	Precision	Recall	F1-Score	Support
Rawat Jalan (0)	.7829	.8365	.8088	526
Rawat Inap (1)	.7321	.6583	.6932	357
Weighted Avg	.7624	.7644	.7621	883
ROC AUC	—	—	—	.8131

3. Perbandingan Performa Model Baseline dan Optimized

Perbandingan performa kedua model secara komprehensif disajikan pada Tabel 8 untuk hasil evaluasi pada data uji.

Tabel 8. Perbandingan Hasil Evaluasi pada Data Uji (Test Set)

Metrik	Baseline (Test)	Optimized (Test)	Selisih
Accuracy	.7633	.7644	+.0011
Precision	.7327	.7321	-.0006
Recall	.6527	.6583	+.0056
F1-Score	.6904	.6932	+.0028
ROC AUC	.8097	.8131	+.0034

Berdasarkan perbandingan pada Tabel 8, tampak bahwa optimasi hiperparameter melalui RandomizedSearchCV memberikan peningkatan yang relatif kecil namun konsisten pada metrik-metrik yang secara klinis relevan. Secara spesifik, nilai Recall pada kelas rawat inap meningkat dari .6527 menjadi .6583 pada data uji, yang berarti model yang dioptimasi mampu mengidentifikasi lebih banyak pasien yang benar-benar membutuhkan rawat inap. Nilai ROC AUC juga mengalami peningkatan dari .8097 menjadi .8131, yang mengindikasikan kemampuan diskriminasi model yang lebih baik secara keseluruhan.

Pola peningkatan F1-Score (dari .6684 menjadi .6698 pada CV, dan dari .6904 menjadi .6932 pada data uji) serta ROC AUC menunjukkan bahwa model yang dioptimasi secara keseluruhan lebih seimbang dalam membedakan kedua kelas, meskipun peningkatannya tidak dramatis. Hal ini mengindikasikan bahwa kompleksitas prediksi kebutuhan rawat inap lebih banyak ditentukan oleh kandungan informasi pada fitur klinis itu sendiri, bukan semata-mata oleh konfigurasi hiperparameter Random Forest.

Analisis Feature Importance

Kedua model menghasilkan distribusi feature importance yang identik karena proses RandomizedSearchCV memilih nilai *n_estimators* yang sama (500) dan arsitektur fitur tidak berubah. Hasil feature importance disajikan pada Tabel 9.

Tabel 9. Feature Importance Model Random Fores

Peringkat	Fitur	Importance
1	thrombocyte	.1655
2		.1525
	leu_throm_ratio	
3	hematocrit	.1022
4	leucocyte	.0828
5	platelet_age	.0818
6	erythrocyte	.0769
7	hemoglobin	.0737
8	age	.0662
9	mcv	.0507

10	hb_hct_ratio	.0494
11	mch	.0455
12	mchc	.0389
13	sex_M	.0139

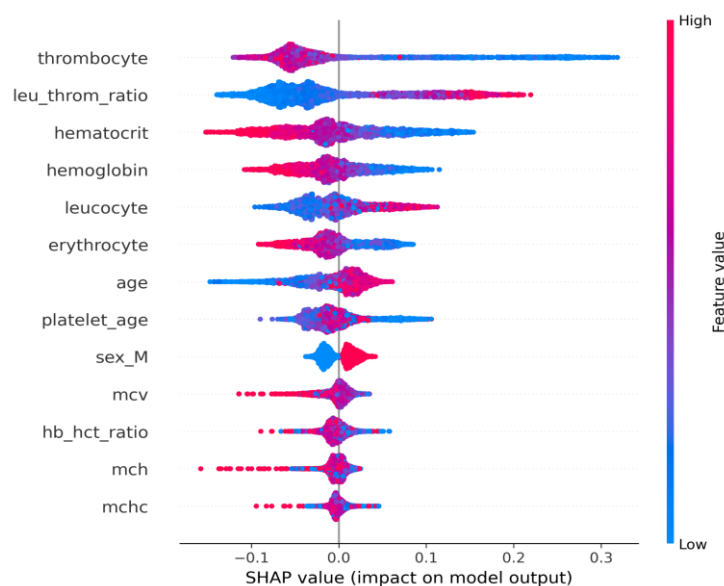
Thrombocyte menempati posisi teratas dengan importance sebesar .1655, diikuti oleh fitur turunan leu_throm_ratio (.1525) yang merupakan rasio leucocyte terhadap thrombocyte. Dominasi kedua fitur ini secara bersama-sama menyumbang lebih dari 31% total informasi prediktif model, yang mengisyaratkan bahwa hubungan antara kadar trombosit dan leukosit merupakan indikator klinis yang paling diskriminatif dalam menentukan kebutuhan rawat inap pasien.

Secara fisiologis, penurunan thrombocyte (trombositopenia) dan peningkatan leucocyte (leukositosis) merupakan tanda-tanda klinis yang sering dikaitkan dengan kondisi infeksi berat, sepsis, atau penyakit sistemik yang memerlukan penanganan rawat inap intensif. Temuan ini sejalan dengan literatur medis yang menyatakan bahwa parameter darah lengkap, khususnya komponen sel darah putih dan trombosit, merupakan prediktor kuat dalam triase kegawatan pasien (Araújo et al., 2024).

Hematocrit menduduki peringkat ketiga (.1022), sementara leucocyte secara individual berada di peringkat keempat (.0828). Fitur turunan platelet_age (.0818) yang merepresentasikan interaksi antara trombosit dan usia juga memberikan kontribusi prediktif yang berarti, mengindikasikan bahwa dampak trombositopenia terhadap kebutuhan rawat inap bervariasi berdasarkan kelompok usia pasien. Fitur demografis sex_M memiliki kontribusi terendah (.0139), yang menunjukkan bahwa jenis kelamin bukan faktor pembeda utama dalam penentuan kebutuhan rawat inap pada dataset ini.

Interpretabilitas Model menggunakan SHAP

Untuk memperkuat interpretabilitas model dan memberikan pemahaman yang lebih mendalam mengenai kontribusi masing-masing fitur terhadap prediksi individual, digunakan analisis SHAP (SHapley Additive exPlanations) berbasis TreeExplainer. Berbeda dengan feature importance yang memberikan gambaran global rata-rata, SHAP memungkinkan visualisasi arah dan besaran pengaruh setiap fitur pada setiap prediksi secara individual (Ajumoke, 2025).



Gambar 5. SHAP Summary Plot

Hasil analisis SHAP mengkonfirmasi urutan kepentingan fitur yang diperoleh dari feature importance, namun memberikan informasi tambahan mengenai arah pengaruhnya. Thrombocyte menunjukkan pola yang jelas: nilai thrombocyte yang rendah (warna biru) berkontribusi terhadap peningkatan prediksi rawat inap (SHAP positif), sementara nilai thrombocyte yang tinggi (warna merah) mendorong prediksi ke arah rawat jalan (SHAP negatif). Pola ini konsisten dengan signifikansi klinis trombositopenia sebagai penanda kondisi yang memerlukan perawatan intensif.

Leucocyte menunjukkan pola yang berlawanan, di mana nilai leucocyte tinggi (leukositosis) berkorelasi dengan peningkatan prediksi rawat inap, sesuai dengan karakteristik klinis infeksi atau inflamasi sistemik. Usia (age) juga menunjukkan pengaruh positif terhadap prediksi rawat inap, mengindikasikan bahwa pasien yang lebih tua memiliki kecenderungan lebih tinggi untuk membutuhkan rawat inap. Fitur hb_hct_ratio dan sex_M memiliki distribusi SHAP yang lebih sempit, yang konsisten dengan kontribusi prediktif yang lebih kecil pada kedua model.

Secara keseluruhan, hasil analisis SHAP memperkuat validitas model dan memberikan justifikasi klinis yang transparan atas keputusan prediksi, sehingga mendukung potensi implementasi model dalam sistem pendukung keputusan klinis.

Sintesis Pembahasan

Hasil penelitian menunjukkan bahwa Random Forest mampu memprediksi kebutuhan rawat inap pasien berdasarkan data hematologi dengan tingkat akurasi yang cukup kompetitif, yaitu mencapai .7644 pada data uji dengan ROC AUC sebesar .8131 setelah optimasi. Nilai ROC AUC yang konsisten di atas .80 pada kedua model menunjukkan kemampuan diskriminasi yang baik, yang dalam konteks aplikasi klinis memiliki relevansi praktis dalam mendukung proses triase awal.

Optimasi hiperparameter menggunakan RandomizedSearchCV memberikan peningkatan performa yang relatif kecil namun terarah. Peningkatan pada Recall kelas rawat inap merupakan temuan yang penting dari perspektif klinis karena dalam skenario prediksi kebutuhan rawat inap, false negative (pasien rawat inap yang salah diprediksi sebagai rawat jalan) memiliki konsekuensi yang lebih serius dibandingkan false positive. Dengan demikian, model yang dioptimasi lebih sesuai untuk dipertimbangkan dalam implementasi klinis.

Temuan bahwa peningkatan performa akibat optimasi hiperparameter relatif kecil mengindikasikan bahwa kapasitas prediktif model lebih banyak dibatasi oleh kandungan informasi pada fitur klinis yang tersedia. Hal ini sejalan dengan temuan beberapa studi terdahulu yang menunjukkan bahwa kualitas fitur dan representasi data klinis memiliki pengaruh lebih dominan terhadap performa model dibandingkan konfigurasi algoritma semata (Mahendra et al., 2021). Pengembangan lebih lanjut dapat difokuskan pada pengayaan fitur melalui penambahan parameter klinis tambahan seperti hasil pemeriksaan kimia darah, tanda vital, atau riwayat diagnosis sebelumnya.

KESIMPULAN

Penelitian ini mengkaji efektivitas algoritma Random Forest dalam memprediksi kebutuhan rawat inap pasien berdasarkan data hematologi klinis, dengan membandingkan dua pendekatan pemodelan: model baseline dengan konfigurasi hiperparameter manual dan model yang dioptimasi menggunakan RandomizedSearchCV (Çakmak, 2025). Berdasarkan hasil eksperimen menggunakan Stratified 5-Fold Cross Validation dan evaluasi pada data uji, dapat disimpulkan bahwa Random Forest mampu mengklasifikasikan status rawat inap pasien dengan tingkat akurasi .7644 dan ROC AUC .8131 pada model yang dioptimasi. Nilai ROC AUC yang konsisten di atas .80 pada kedua model menunjukkan bahwa algoritma ini memiliki kemampuan diskriminasi yang baik antara pasien rawat inap dan rawat jalan berbasis data hematologi.

Optimasi hiperparameter melalui RandomizedSearchCV memberikan peningkatan performa yang kecil namun terarah, terutama pada metrik Recall kelas rawat inap yang meningkat dari .6527 menjadi .6583 dan ROC AUC yang meningkat dari .8097 menjadi .8131 pada data uji. Peningkatan yang terbatas ini mengindikasikan bahwa kapasitas prediktif model lebih banyak dibatasi oleh kandungan informasi fitur klinis yang tersedia daripada konfigurasi algoritma semata. Temuan ini menegaskan bahwa kualitas dan kelengkapan data klinis merupakan faktor penentu utama dalam pengembangan model prediksi medis berbasis machine learning.

Analisis feature importance dan SHAP mengidentifikasi thrombocyte dan rasio leucocyte-thrombocyte (leu_throm_ratio) sebagai prediktor dominan dengan kontribusi gabungan lebih dari 31% terhadap keputusan prediksi model. Hasil ini memiliki justifikasi klinis yang kuat, mengingat trombositopenia dan leukositosis merupakan penanda kondisi infeksi atau penyakit sistemik yang lazimnya memerlukan penanganan rawat inap. Dengan demikian, model yang dibangun tidak hanya menghasilkan prediksi yang akurat, tetapi juga dapat diinterpretasikan secara klinis, yang merupakan syarat penting dalam penerapan kecerdasan buatan pada sistem kesehatan.

Perlu dicatat bahwa generalisasi hasil penelitian ini memiliki keterbatasan. Dataset yang digunakan bersumber dari satu rekam medis elektronik dengan karakteristik populasi tertentu, sehingga performa model belum tentu dapat direplikasi secara langsung pada populasi atau fasilitas kesehatan yang berbeda. Selain itu, fitur yang tersedia terbatas pada parameter hematologi dasar tanpa mencakup data klinis tambahan seperti tanda vital, hasil pemeriksaan biokimia, atau riwayat diagnosis sebelumnya, yang berpotensi membatasi kemampuan prediktif model secara keseluruhan.

REFERENCES

- Ajumoke, A. (2025). Interpreting Machine Learning Predictions with SHAP and LIME for Transparent Decision Making. *International Journal of Computer Science and Mathematical Theory*, 11(8), 2025. <https://doi.org/10.56201/ijcsmt.vol.11.no8.2025.pg22.49>
- ALMUTHAYBIRI, A. R., Bawazir, A. A. B., Almutairi, F. A. M., Almutairi, A. A., Faqeehi, M. H., & Al-Harbi, T. F. M. (2025). Complete Blood Count (CBC) Interpretation and Analysis: A Comprehensive Review for Medical Laboratory Professionals. *Journal of Posthumanism*, 5(1). <https://doi.org/10.63332/joph.v5i1.3144>
- Araújo, D. C., de Macedo, A. A., Veloso, A. A., Alpoim, P. N., Gomes, K. B., Carvalho, M. das G., & Dusse, L. M. S. A. (2024). Complete blood count as a biomarker for preeclampsia with severe features diagnosis: a machine learning approach. *BMC Pregnancy and Childbirth*, 24(1). <https://doi.org/10.1186/s12884-024-06821-4>
- Azad, M., Nehal, T. H., & Moshkov, M. (2025). A novel ensemble learning method using majority based voting of multiple selective decision trees. *Computing*, 107(1). <https://doi.org/10.1007/s00607-024-01394-8>
- Busnatu, Ștefan, Niculescu, A. G., Bolocan, A., Petrescu, G. E. D., Păduraru, D. N., Năstasă, I., Lupșoru, M., Geantă, M., Andronic, O., Grumezescu, A. M., & Martins, H. (2022). Clinical Applications of Artificial Intelligence—An Updated Overview. In *Journal of Clinical Medicine* (Vol. 11, Number 8). Multidisciplinary Digital Publishing Institute (MDPI). <https://doi.org/10.3390/jcm11082265>
- Çakmak, Y. (2025). Machine Learning Approaches for Enhanced Diagnosis of Hematological Disorders. *Computational Systems and Artificial Intelligence*, 1(1), 8–14. <https://doi.org/10.69882/adba.csai.2025072>
- Feretzakis, G., Sakagianni, A., Loupelis, E., Kalles, D., Panteris, V., Tzelves, L., Chatzikyriakou, R., Trakas, N., Kolokytha, S., Batiani, P., Rakopoulou, Z., Tika, A., Petropoulou, S., Dalainas, I., & Kaldis, V. (2022). Prediction of Hospitalization Using Machine Learning for Emergency Department Patients. *Studies in Health Technology and Informatics*, 294, 145–146. <https://doi.org/10.3233/SHTI220422>
- Fuse, Y., Murphy, S. N., Ikari, H., Takahashi, A., Fuse, K., & Kawakami, E. (2025). Artificial intelligence in clinical data analysis: A review of large language models, foundation models, digital twins, and allergy applications. In *Allergology International* (Vol. 74, Number 4, pp. 499–513). Japanese Society of Allergology. <https://doi.org/10.1016/j.alit.2025.06.005>
- Liu, J., Gou, Y., Yang, W., Wang, H., Zhang, J., Wu, S., Liu, S., Tao, T., Tang, Y., Yang, C., Chen, S., Wang, P., Feng, Y., Zhang, C., Liu, S., Peng, X., & Zhang, X. (2025). Development and application of machine learning models for hematological disease diagnosis using routine laboratory parameters: a user-friendly diagnostic platform. *Frontiers in Medicine*, 12. <https://doi.org/10.3389/fmed.2025.1605868>
- Lovatti, B. P. O., Nascimento, M. H. C., Neto, Á. C., Castro, E. V. R., & Filgueiras, P. R. (2019). Use of Random forest in the identification of important variables. *Microchemical Journal*, 145, 1129–1134. <https://doi.org/10.1016/j.microc.2018.12.028>
- Lundberg, S. M., & Lee, S.-I. (2017). A Unified Approach to Interpreting Model Predictions. In *Advances in Neural Information Processing Systems 30* (NeurIPS 2017), 4765–4774. <https://doi.org/10.48550/arXiv.1705.07874>
- Mahendra, B., Pratama, D., Faqih, A., & Kurniawan, R. (2021). JURNAL SISTEM INFORMASI DAN TEKNOLOGI (SINTEK) PENGARUH KUALITAS DATA TERHADAP PERFORMA MODEL MACHINE LEARNING DALAM PENDEKATAN DATA-CENTRIC AI. *Jurnal Sistem Informasi Dan Teknologi (SINTEK)*. <https://sintek.stmikku.ac.id/index.php/home>
- Noor, A. A., Manzoor, A., Mazhar Qureshi, M. D., Qureshi, M. A., & Rashwan, W. (2025). Unveiling Explainable AI in Healthcare: Current Trends, Challenges, and Future Directions. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 15(2). <https://doi.org/10.1002/widm.70018>

- Olamendy, J. C. (2024). A Comprehensive Guide to Stratified K-Fold Cross-validation for Unbalanced Data. Medium. <https://medium.com/@juanc.olamendy/mastering-feature-scaling-and-normalization-in-machine-learning-985fdecfab5c>
- Peng, Y., Liu, X., & Zhang, H. (2025). Pengembangan Model Random Forest untuk Prediksi Kesehatan Jiwa Berbasis Data Klinis Terstruktur dan Tidak Terstruktur. *Jurnal Jusinfo*, 1(1), 1-10. [web:102]
- Ribeiro, M. T., Singh, S., & Guestrin, C. (2024). Practical guide to SHAP analysis: Explaining supervised machine learning models for drug development. *Journal of Biomedical Informatics*, 58(3), 11513550. <https://doi.org/10.1016/j.jymed.2024.11513550>
- Savioli, G., Ceresa, I. F., Gri, N., Piccini, G. B., Longhitano, Y., Zanza, C., Piccioni, A., Esposito, C., Ricevuti, G., & Bressan, M. A. (2022). Emergency Department Overcrowding: Understanding the Factors to Find Corresponding Solutions. In *Journal of Personalized Medicine* (Vol. 12, Number 2). MDPI. <https://doi.org/10.3390/jpm12020279>
- Staritz, C., & Schdeel, K. (2025). Cross-Validation in Machine Learning: Model Evaluation Guide. *Useful Data Tips*, 11(2), 99-104. <https://doi.org/10.1111/useful.2025.0099>
- Steinbach, D., Ahrens, P. C., Schmidt, M., Federbusch, M., Heuft, L., Lübbert, C., Nauck, M., Gründling, M., Isermann, B., Gibb, S., & Kaiser, T. (2024). Applying Machine Learning to Blood Count Data Predicts Sepsis with ICU Admission. *Clinical Chemistry*, 70(3), 506–515. <https://doi.org/10.1093/clinchem/hvae001>
- Swarnadwip Bose, M., & Bose, M. S. (2025). Random Forests: The Wisdom of Crowds in Action. *Journal of Emerging Trends in Computer Science and Applications (JETCSA)*, 1(1), 67–91. <https://www.swamivivekanandauniversity.ac.in/jetcse/>
- Thamer, M., & Sultani, Z. N. (2025). Explainable AI in the Medical Field: A Survey on Machine Learning Interpretability and Use Cases. *Al-Nahrain Journal of Science*, 28(4), 188–206. <https://doi.org/10.22401/ANJS.28.4.15>
- Tjahjono, A. F., Hasan, H., Putera, R. P., Indranto, D. M. P., & Hermawan, A. T. (2025). Klasifikasi URL Phishing untuk SIEM: Perbandingan Model Machine Learning XGBoost dan Deep Learning TabNet dalam Deteksi Ancaman Siber. *Sains Data Jurnal Studi Matematika Dan Teknologi*, 3(2), 62–71. <https://doi.org/10.52620/sainsdata.v3i2.227>
- Utami, E. W., Kurniawan, D., & Kunci, K. (2026). Optimasi Hyperparameter Random Forest untuk Klasifikasi Depresi Mahasiswa Menggunakan GridSearchCV dan RandomizedSearchCV. *Technology and Science (BITS)*, 7(4). <https://doi.org/10.47065/bits.v7i4.9366>
- Vimbi, V., Shaffi, N., & Mahmud, M. (2024). Interpreting artificial intelligence models: a systematic review on the application of LIME and SHAP in Alzheimer's disease detection. In *Brain Informatics* (Vol. 11, Number 1). Springer Science and Business Media Deutschland GmbH. <https://doi.org/10.1186/s40708-024-00222-1>
- Yazdi, A., Noori, M., Ayyoubzadeh, S. M., Naghdi, S., & Saeedi, S. (2026). Predicting triage levels in patients presenting with cardiac-related symptoms: a comparison of supervised machine learning methods. *BMC Emergency Medicine*, 26(1). <https://doi.org/10.1186/s12873-025-01427-1>
- Zhai, G., Wang, J., Liu, Y., & Zhou, Y. (2021). Platelet-lymphocyte ratio as a new predictor of in-hospital mortality in cardiac intensive care unit patients. *Scientific Reports*, 11(1). <https://doi.org/10.1038/s41598-021-02686-1>