

KLASIFIKASI RISIKO INFEKSI SALURAN PERNAPASAN AKUT MENGUNAKAN *ENSEMBLE SOFT VOTING* BERBASIS REKAM MEDIS

**Alvian Syah Burhani¹⁾, Muhammad Faisal²⁾, Fahrir Irhamna Rachman³⁾, Darniati⁴⁾,
Titin Wahyuni⁵⁾, Muhammad Syafaat S.Kuba⁶⁾, Farida Gaffar⁷⁾**

1. Informatika, Universitas Muhammadiyah Makassar
email: 105841103522@student.unismuh.ac.id
2. Informatika, Universitas Muhammadiyah Makassar
email: muhfaisal@unismuh.ac.id
3. Informatika, Universitas Muhammadiyah Makassar
email: fachrim141020@unismuh.ac.id
4. Informatika, Universitas Muhammadiyah Makassar
email: darniati@unismuh.ac.id
5. Informatika, Universitas Muhammadiyah Makassar
email: titinwahyuni@unismuh.ac.id
6. Teknik Pengairan, Universitas Muhammadiyah Makassar
email: syafaat_skuba@unismuh.ac.id
7. Teknik Pengairan, Universitas Muhammadiyah Makassar
email: faridagaffar@unismuh.ac.id

Abstract

This study proposes a risk classification model for Acute Respiratory Infections using a Soft Voting-based Ensemble Learning approach, which integrates prediction probabilities from the Random Forest and Extreme Gradient Boosting algorithms. The research utilized 456 patient medical record data from RSUD Latemamala spanning January 2020 to December 2025. Comparative evaluation results show that the Random Forest model achieved an accuracy of 94.57%, Extreme Gradient Boosting reached 95.65%, and the Soft Voting Ensemble model delivered the best performance with an accuracy of 96.74%, precision of 96.97%, recall of 96.88%, and an F1-Score of 96.77%. Furthermore, the Soft Voting Ensemble model successfully achieved a perfect recall score for the Severe Acute Respiratory Infection category, ensuring that no high-risk patients went undetected. In conclusion, the Soft Voting Ensemble model serves as a reliable decision-support tool to assist medical professionals in triaging Acute Respiratory Infection patients quickly, objectively, and accurately.

Kata Kunci : *Infeksi Saluran Pernapasan Akut; Ensemble Learning; Soft Voting; Random Forest; Extreme Gradient Boosting*

A. PENDAHULUAN

Infeksi Saluran Pernapasan Akut masih menempati posisi strategis sebagai penyumbang masalah kesehatan utama di Indonesia dengan angka kesakitan dan kematian yang tinggi, terutama di kalangan kelompok rentan seperti balita dan lansia [1]. Tren kasus Infeksi Saluran Pernapasan Akut menunjukkan peningkatan signifikan yang diperparah oleh kondisi polusi udara dan perubahan iklim ekstrem [2], [3]. Penyakit ini menyerang saluran pernapasan dan membutuhkan penanganan cepat untuk mencegah komplikasi mematikan seperti pneumonia, sehingga prevalensinya berdampak negatif pada sistem perawatan kesehatan nasional termasuk fasilitas rujukan daerah.

Kompleksitas penanganan Infeksi Saluran Pernapasan Akut tidak hanya terletak pada tingginya jumlah kasus, tetapi juga pada variasi faktor risikonya yang beragam. Kondisi lingkungan rumah, kualitas ventilasi, kepadatan hunian, status gizi, kelengkapan imunisasi, serta perilaku kesehatan masyarakat memegang peranan vital dalam prevalensi Infeksi Saluran Pernapasan Akut [4], [5]. Banyaknya variabel yang berpengaruh menjadikan proses diagnosis dan klasifikasi risiko secara manual menjadi tidak efisien dan rentan terhadap subjektivitas tenaga medis.

Kondisi tersebut menciptakan tantangan operasional nyata bagi Rumah Sakit Umum Daerah Latemmamala di Kabupaten Soppeng sebagai pusat rujukan daerah. Volume pasien yang terus meningkat setiap tahunnya seringkali tidak sebanding dengan ketersediaan tenaga medis spesialis di unit gawat darurat maupun poliklinik. Kesalahan atau keterlambatan dalam mengklasifikasikan tingkat risiko Infeksi Saluran Pernapasan Akut, apakah pasien Tidak Infeksi Saluran

Pernafasan Akut, mengalami Infeksi Saluran Pernafasan Akut Sedang, atau terindikasi Infeksi Saluran Pernafasan Akut Berat, dapat berdampak fatal pada keselamatan pasien serta menyebabkan inefisiensi penggunaan sumber daya medis.

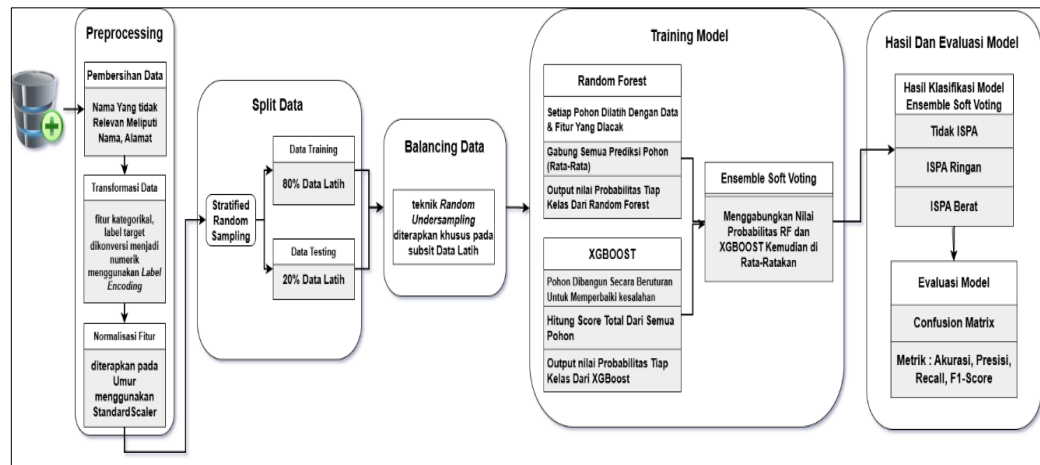
Berbagai penelitian terdahulu telah menunjukkan efektivitas algoritma *Machine Learning* dalam klasifikasi penyakit berbasis data rekam medis. Penerapan *Random Forest* pada klasifikasi penyakit saluran pernapasan berhasil mencapai akurasi sempurna pada pengujian *cross-validation*, mengungguli algoritma pembandingan lainnya [6], [7]. Namun, penggunaan model tunggal seringkali belum mampu menangani kompleksitas data klinis secara optimal karena masing-masing algoritma memiliki keterbatasan tersendiri dalam menangani variasi pola data [8].

Pendekatan *Ensemble Learning* yang menggabungkan beberapa model prediksi terbukti menghasilkan output yang lebih akurat dan stabil dibandingkan model tunggal. Teknik *Soft Voting* yang menggabungkan probabilitas prediksi dari beberapa model telah terbukti meningkatkan performa klasifikasi pada berbagai domain medis, termasuk prediksi kelangsungan hidup pasien gagal jantung dan deteksi pasien COVID-19 [9], [10]. Namun, penerapan *Ensemble Soft Voting* yang mengintegrasikan *Random Forest* dan *Extreme Gradient Boosting* secara bersamaan untuk klasifikasi risiko Infeksi Saluran Pernafasan Akut berbasis data rekam medis klinis di Indonesia belum banyak dieksplorasi. Oleh karena itu, penelitian ini berkontribusi dalam membangun model klasifikasi risiko Infeksi Saluran Pernafasan Akut berbasis *Ensemble Soft Voting* menggunakan 456 data rekam medis Rumah Sakit Umum Daerah Latemmamala periode 2020–2025, dengan target menghasilkan performa yang lebih superior

dibandingkan model *Random Forest* dan *Extreme Gradient Boosting* secara individual sebagai alat bantu pengambilan keputusan klinis dalam proses triase pasien.

B. METODE PENELITIAN

Penelitian ini dilakukan melalui pendekatan kuantitatif eksperimental dengan tahapan sistematis yang dimulai dari pengumpulan data, pra-pemrosesan, pembagian data, pelatihan model, hingga evaluasi dan perbandingan performa. Alur metodologi penelitian secara keseluruhan ditunjukkan pada Gambar 1. ilmiah.



Gambar 1. Desain Perancangan Sistem

Dataset yang digunakan merupakan data sekunder berupa rekam medis pasien Infeksi Saluran Pernafasan Akut dari Instalasi Rekam Medis Rumah Sakit Umum Daerah Latemmamala, Kabupaten Soppeng, dengan waktu periode Januari 2020 hingga Desember 2025. Dataset berjumlah 456 sampel dengan rincian 16 variabel input dan 1 variabel target. Variabel input terdiri dari 2 atribut demografi yaitu jenis kelamin dan umur, serta 14 atribut gejala klinis meliputi batuk kering, batuk berdahak, demam, pilek, hidung tersumbat, sesak napas, nyeri tenggorokan, sakit kepala, mual muntah, nyeri dada, suara serak, nafsu makan menurun, hilang penciuman, dan nyeri saat menelan. Variabel target terdiri dari tiga kelas diagnosis yaitu ISPA

Berat, ISPA Sedang, dan Tidak Infeksi Saluran Pernafasan Akut.

Pra-Pemrosesan Data

Tahapan pra-pemrosesan dilakukan untuk memastikan kualitas dan kesiapan data sebelum digunakan dalam pelatihan model. Pertama, Pembersihan data, tahap ini dilakukan untuk memastikan bahwa dataset yang digunakan dalam proses pelatihan model bebas dari data yang tidak relevan dan redundan seperti nama dan Alamat pasien dan penghapusan data duplikat.

Kedua, transformasi fitur dilakukan menggunakan *Label Encoding* pada atribut jenis kelamin memiliki nilai kategorikal berbasis teks, atribut ini perlu dikonversi terlebih dahulu ke dalam representasi numerik (Laki-laki=0, Perempuan=1) dan label target (ISPA

Berat=0, ISPA Sedang=1, Tidak ISPA=2).

Ketiga, tahap normalisasi nilai atribut numerik agar berada dalam rentang skala yang seragam *feature scaling* hanya diterapkan pada atribut Umur menggunakan metode *StandardScaler*. Atribut Umur memiliki rentang nilai antara 1 tahun hingga 80 tahun dengan rata-rata sebesar 41,12 dan standar deviasi sebesar 23,34, sehingga perlu dinormalisasi agar sebanding dengan atribut lainnya. *StandardScaler* bekerja dengan cara mentransformasi nilai atribut menggunakan rumus z-score sebagai berikut:

$$z = \frac{x - \mu}{\sigma} \quad (1)$$

Dimana x adalah nilai data, μ adalah rata-rata (41,12), dan σ adalah standar deviasi (23,34).

Atribut gejala klinis tidak memerlukan scaling karena seluruhnya berskala biner. Keempat, penyeimbangan kelas dilakukan menggunakan *Random undersampling* yang diterapkan hanya pada data latih untuk mengatasi ketidakseimbangan distribusi kelas.

Pembagian Data

Dataset dibagi dengan proporsi 80% data latih dan 20% data uji menggunakan *stratified sampling* setelah dilakukan pembagian data, selanjutnya dilakukan *Random Undersampling* diterapkan hanya pada data latih untuk menyeimbangkan distribusi kelas, menghasilkan 118 sampel per kelas (total 354 sampel). Data uji tidak disentuh untuk menjaga objektivitas evaluasi.

Arsitektur Model

Penelitian ini mengusulkan model klasifikasi berbasis *Ensemble Soft Voting* yang membangun dua algoritma sebagai

base learner dengan *base learner* pertama *Random Forest* tersusun dari himpunan pohon keputusan yang beroperasi secara simultan [11]. Pembentukan setiap pohon didasarkan pada teknik *bootstrapping* dan *random feature selection* pada setiap simpul, sehingga menciptakan keberagaman antar pohon yang menjadi kunci ketahanan model terhadap *noise* pada data klinis [12]. Kualitas pembelahan diukur menggunakan Gini Impurity [13], dengan rumus:

$$\text{Gini}(D) = 1 - \sum_{i=1}^c p_i^2 \quad (2)$$

Dimana $\text{Gini}(D)$ adalah Index Gini dari dataset D , P_i adalah probabilitas kemunculan kelas ke- i dalam dataset, dan C adalah jumlah hasil kelas target.

Hasil prediksi seluruh pohon digabungkan melalui mekanisme Majority Voting

$$H(x) = \underset{= y}{\operatorname{argmax}_y} \sum_{k=1}^K I(h_k(x)) \quad (3)$$

Dimana $H(x)$ adalah hasil prediksi akhir model, $H_k(x)$ adalah hasil prediksi dari pohon keputusan ke- k , dan K adalah jumlah total pohon dalam *Random Forest*. *Hyperparameter* terbaik diperoleh melalui *Randomized Search Cross-Validation* dengan konfigurasi: $n_estimators=100$, $max_depth = \text{None}$, $min_samples_split=2$, $min_samples_leaf=1$, $max_features=\log 2$.

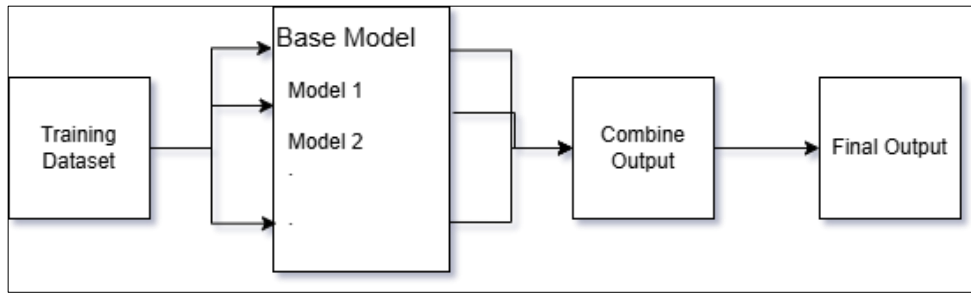
Extreme Gradient Boosting dirancang untuk mengoptimalkan performa komputasi dan akurasi prediksi melalui teknik regulasi yang ketat [14]. Fungsi objektif pada iterasi ke- t dirumuskan sebagai:

$$L^{(t)} = \sum_{i=1}^n l(y_i, \hat{y}_1^{(t-1)} + f_t(x_i) + \Omega(f_x)) \quad (4)$$

Dimana l adalah fungsi kerugian (*loss function*) yang mengukur selisih antara label aktual y_i dan prediksi $\hat{y}_i$, untuk $f_t(x_i)$ adalah fungsi pohon baru yang ditambahkan pada langkah ke- t , dan $\Omega(f_x)$ adalah suku regulasi yang membatasi pertumbuhan pohon agar tidak terlalu kompleks.

Dalam penelitian ini, XGBoost berperan menghasilkan nilai probabilitas prediksi setiap kelas yang diteruskan ke mekanisme *Ensemble Soft Voting*. Konfigurasi hyperparameter terbaik hasil tuning: $n_estimators=200$, $max_depth=7$, $learning_rate=0.05$, $subsample=1.0$, $colsample_bytree=1.0$.

Kedua model kemudian diintegrasikan menggunakan mekanisme *Ensemble Soft Voting* yang menghitung rata-rata probabilitas prediksi sebagaimana diilustrasikan pada Gambar 2:



Gambar 2. Arsitektur Ensemble Soft Voting [15]

Untuk menghasilkan keputusan klasifikasi akhir [8], dirumuskan sebagai:

$$S = \frac{1}{N} \sum_{i=1}^N P_i \quad (5)$$

Dimana S adalah nilai probabilitas akhir dari *Soft Voting Classifier*, N adalah jumlah total model dasar yang digunakan dalam *ensemble*, dan P_i adalah probabilitas prediksi yang dihasilkan oleh model ke- i .

Evaluasi Model

Kinerja model diukur menggunakan *Confusion Matrix*, yaitu alat evaluasi yang bekerja

dengan mengomparasikan *output* prediksi model terhadap nilai sebenarnya pada data uji [16]. Matriks ini dikonstruksikan berdasarkan empat parameter fundamental, yaitu *True Positive* (TP), *True Negative* (TN), *False Positive* (FP), dan *False Negative* (FN) [17]. Berdasarkan keempat parameter tersebut, metrik evaluasi dihitung sebagai berikut:

Akurasi: Mengukur ketepatan model secara keseluruhan dalam mengklasifikasikan seluruh data uji:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (6)$$

Dimana Presisi: merupakan metrik yang digunakan untuk menilai

kesesuaian antara permintaan informasi oleh pengguna dengan hasil respons yang dihasilkan oleh sistem.

$$Precision = \frac{TP}{TP+FP}$$

Dimana *Recall* mengindikasikan tingkat kebersihan model, yaitu sejauh mana model mampu menemukan kembali seluruh sampel yang berasal dari kelas positif pada dataset pengujian.

$$Recall = \frac{TP}{TP+FP} \quad (8)$$

Dimana *F1- Score*: evaluasi yang mengombinasikan *precision* dan *recall* melalui perhitungan rata-rata harmonik dalam satu nilai untuk menghasilkan evaluasi yang seimbang.

$$F1 - Score = 2 \times \frac{Presisi \times Recall}{Presisi + Recall} \quad (9)$$

C. HASIL DAN PEMBAHASAN

Pada bagian ini dijelaskan hasil penelitian dan sekaligus diberikan pembahasan yang komprehensif. Hasil dapat disajikan dalam bentuk gambar, grafik, tabel, dan lain-lain yang membuat pembaca mudah memahaminya. Pembahasan dapat dilakukan dalam beberapa sub bab. Sangat disarankan bahwa perbandingan dengan hasil dari artikel lain yang diterbitkan disediakan untuk memberikan lebih banyak konteks dan untuk memperkuat klaim kebaruan.

Hasil Preprocessing Data

Sebelum data digunakan dalam pelatihan model, dilakukan serangkaian tahapan *preprocessing* untuk memastikan kualitas dan kesiapan data. Atribut yang tidak relevan terhadap proses klasifikasi seperti nama dan alamat pasien dihapus dari dataset. Pengecekan data duplikat dilakukan menggunakan fungsi `drop_duplicates()` dan tidak ditemukan adanya duplikasi, sehingga seluruh 456 sampel dipertahankan. Seluruh atribut kategorikal dikonversi ke representasi numerik menggunakan teknik *Label Encoding*, di mana atribut jenis kelamin dikonversi menjadi nilai biner (Laki-laki=0, Perempuan=1) dan label target dikonversi menjadi nilai integer Berat=0, Sedang=1, Tidak=2). *Feature scaling* diterapkan pada atribut umur menggunakan *StandardScaler* dengan rata-rata $\mu=41,12$ dan standar deviasi $\sigma=23,34$. Dataset kemudian dibagi dengan proporsi 80% data latih dan 20% data uji menggunakan *stratified sampling* ditunjukkan pada Tabel 1:

Tabel 1. Distribusi Kelas Data Latih dan Data Uji

Kelas	Data	
	Latih	Uji
Berat	118	30
Sedang	127	32
Tidak	120	30
Total	365	92

Random *Undersampling* kemudian diterapkan hanya pada data latih untuk mengatasi ketidakseimbangan kelas, menghasilkan 118 sampel per kelas

dengan total 354 sampel latih ditunjukkan pada Tabel 2.

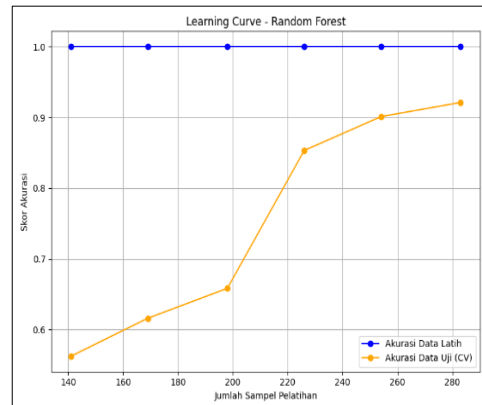
Tabel 2. Distribusi Kelas Data Latih dan Data Uji

Kelas	Data Balancing	
	Sebelum	Setelah
Berat	118	118
Sedang	127	118
Tidak	120	118
Total	365	354

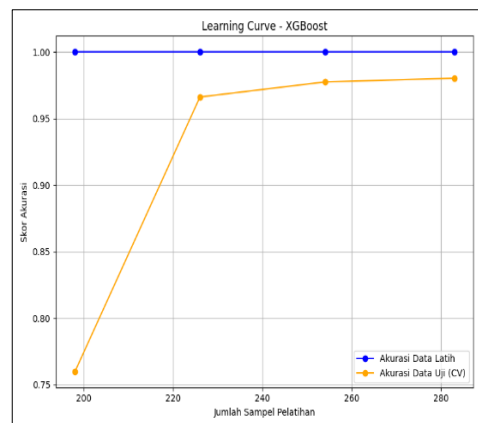
Hasil Pelatihan Model

Sebelum dilakukan evaluasi menggunakan *Confusion Matrix*, stabilitas dan kemampuan generalisasi kedua model dasar dianalisis terlebih dahulu melalui *Learning Curve*. Analisis ini bertujuan untuk memastikan bahwa model tidak mengalami *overfitting* maupun *underfitting* sebelum diintegrasikan ke dalam mekanisme *Ensemble Soft Voting*.

Berdasarkan Gambar 3, akurasi data latih (*training score*) *Random Forest* bertahan secara konsisten pada skor 1,00, menunjukkan kemampuan model dalam mempelajari pola data latih secara sempurna. Sementara itu, akurasi pada data uji (*cross-validation score*) menunjukkan tren peningkatan yang stabil, dimulai dari skor 0,56 hingga mencapai titik konvergensi pada angka 0,92. Jarak yang semakin menyempit antara kurva latih dan kurva uji mengindikasikan bahwa model *Random Forest* berhasil melakukan generalisasi dengan baik tanpa gejala *overfitting* yang signifikan.



Gambar 3. Learning Curve Random Forest



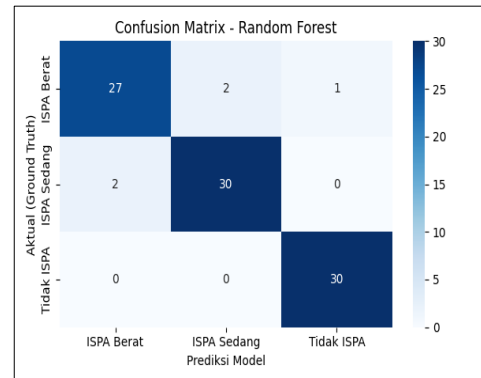
Gambar 4. Learning Curve Extreme Gradient Boosting

Pada model *Extreme Gradient Boosting*, *Learning Curve* pada Gambar 4 menunjukkan profil performa yang sangat progresif. Akurasi data latih bertahan konsisten pada skor 1.00, sementara akurasi data uji menunjukkan tren peningkatan yang jauh lebih tajam dibandingkan *Random Forest*, bergerak dari kisaran 0.76 hingga mencapai titik konvergensi pada angka 0.98. Celah yang sangat minimal antara kurva latih dan kurva uji membuktikan efektivitas mekanisme regulasi internal algoritma ini dalam memitigasi risiko *overfitting*, sehingga *Extreme*

Gradient Boosting dinyatakan sangat reliabel untuk diintegrasikan ke dalam mekanisme *Ensemble Soft Voting*.

Setelah stabilitas model terkonfirmasi, evaluasi performa dilakukan menggunakan *Confusion Matrix* untuk memetakan distribusi prediksi benar dan kesalahan klasifikasi pada setiap kategori risiko Infeksi Saluran Pernafasan Akut.

Berdasarkan Gambar 5, model *Random Forest* menunjukkan performa yang sangat kompetitif dalam mengenali pola klinis pada data uji.



Gambar 5. Confusin Matrix Random Forest

Pada kategori ISPA Berat, model mampu mengidentifikasi 27 data secara tepat dengan 3 kasus misklasifikasi sebagai *False Negative*. Pada kategori ISPA Sedang, model menunjukkan reliabilitas tinggi dengan mengidentifikasi 30 sampel secara benar dan hanya menyisakan 2 data sebagai *False Negative*. Capaian paling impresif terlihat pada kategori Tidak ISPA, di mana model berhasil mengklasifikasikan seluruh 30 data secara akurat tanpa kesalahan (TP=30, FN=0). Hasil evaluasi lengkap disajikan pada Tabel 3.

Tabel 3. Klasifikasi Report Random Forest

Kelas	Matriks			
	Presisi	Recall	F1-Score	Support
Berat	0.93	0.90	0.92	30
Sedang	0.94	0.94	0.94	32
Tidak	0.97	1.00	0.98	30
Total	0.9454	0.9458	0.9455	92

Secara akumulatif, model *Random Forest* mencatatkan akurasi sebesar 94,57%, presisi 94,54%, *recall* 94,58%, dan F1-Score 94,55%. Hasil ini membuktikan bahwa *Random Forest* memiliki kapabilitas yang solid sebagai *base learner*

dalam sistem pendukung keputusan klinis Infeksi Saluran Pernafasan Akut.

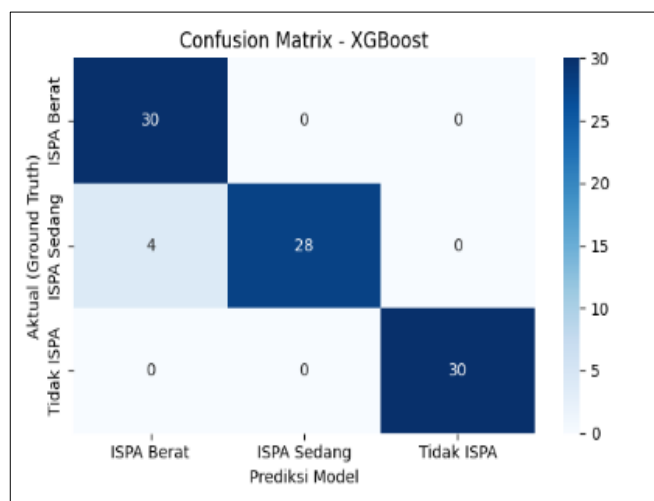
Berdasarkan Gambar 6, *Extreme Gradient Boosting* menunjukkan kapabilitas yang sangat impresif,

terutama dalam meminimalisir risiko kesalahan fatal pada diagnosis medis.

Model ini berhasil mencapai nilai Recall sempurna (1,00) pada kategori ISPA Berat, di mana seluruh 30 sampel berhasil diidentifikasi secara tepat tanpa adanya kasus *False Negative*. Keberhasilan yang identik juga tercatat pada kategori Tidak ISPA. Meskipun demikian, terdapat

tantangan spesifik pada kategori ISPA Sedang, di mana 4 sampel terprediksi secara keliru sebagai ISPA Berat sehingga nilai *recall* pada kelas tersebut terkoreksi menjadi 0,88.

Hal ini mengindikasikan adanya kemiripan pola fitur gejala antara kategori sedang dan berat. Hasil evaluasi disajikan pada Tabel 4-5.



Gambar 6. Confusion Matrix Extreme Gradient Boosting

Tabel 4. Klasifikasi Report Random Forest

Kelas	Metriks			
	Presisi	Recall	F1-Score	Support
Berat	0.99	1.00	0.99	30
Sedang	1.00	0.88	0.93	32
Tidak	1.00	1.00	1.00	30
Total	0.9608	0.9583	0.9569	92

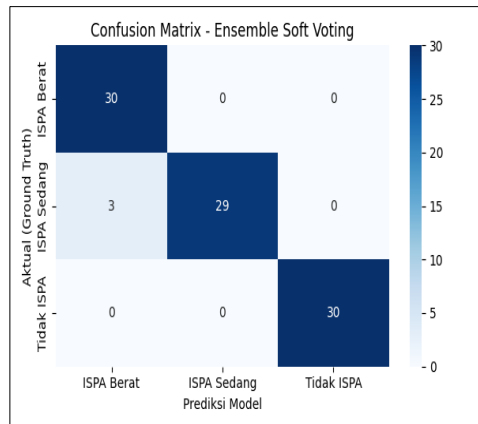
Tabel 5. Klasifikasi Report Ensemble Soft Voting

Kelas	Metriks			
	Presisi	Recall	F1-Score	Support
Berat	0.91	1.00	0.95	30
Sedang	1.00	0.91	0.95	32
Tidak	1.00	1.00	1.00	30
Total	0.9697	0.9688	0.9677	92

Secara keseluruhan, *Extreme Gradient Boosting* memberikan performa yang lebih unggul

dibandingkan *Random Forest* dengan capaian akurasi 95,65%, presisi

96,08%, *recall* 95,83%, dan F1-Score 95,69%.



Gambar 7. Confusin Matrix Ensemble Soft Voting

Berdasarkan Gambar 7, model *Ensemble Soft Voting* membuktikan keunggulannya melalui mekanisme integrasi probabilitas yang mampu menghasilkan keputusan klasifikasi yang lebih halus dan akurat dibandingkan model tunggal. Model ini berhasil mempertahankan nilai *Recall* sempurna (1,00) pada kategori ISPA Berat (TP=30, FN=0) serta kategori Tidak ISPA (TP=30, FN=0). Pencapaian ini memiliki signifikansi klinis yang sangat tinggi karena menjamin deteksi seluruh pasien

berisiko kritis tanpa adanya kasus *False Negative*. Peningkatan kualitas prediksi yang nyata juga terlihat pada kategori ISPA Sedang, di mana model berhasil mengidentifikasi 29 dari 32 data secara tepat, membuktikan bahwa sinergi antara *Random Forest* dan *Extreme Gradient Boosting* melalui teknik *Soft Voting* mampu memitigasi ambiguitas pola gejala yang sebelumnya menjadi tantangan bagi model individual. Hasil evaluasi lengkap disajikan pada Tabel 6.

Puncaknya, penerapan metode *Ensemble Soft Voting* memberikan hasil paling optimal dengan capaian akurasi tertinggi sebesar 96,74%, presisi 96,97%, dan *recall* 96,88%. Signifikansi klinis model usulan ini terlihat pada pencapaian *recall* sempurna (1,00) untuk kategori Berat dan Tidak, yang sangat krusial guna menjamin deteksi seluruh pasien berisiko tinggi tanpa ada yang terlewat. Meskipun akurasi pada kelas ISPA terkoreksi ke angka 90,63% akibat kemiripan pola gejala klinis dengan kategori berat, sinergi melalui mekanisme penggabungan probabilitas terbukti mampu menghasilkan sistem triase yang lebih *robust* dan objektif

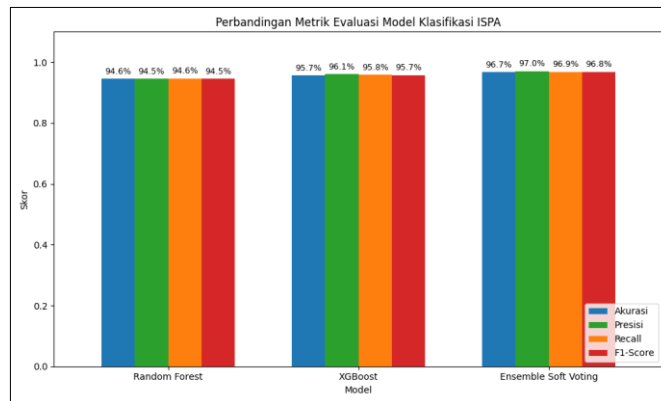
Tabel 6. Akurasi Per Kelas Ensemble Soft Voting

Kelas	Kesesuaian		
	Data Uji	Presiksi Benar	Akurasi
Berat	30	30	100.00%
Sedang	32	29	90.63%
Tidak	30	30	100.00%

Secara akumulatif, model *Ensemble Soft Voting* mencatatkan performa tertinggi dalam seluruh eksperimen dengan capaian akurasi 96,74%, presisi 96,97%, *recall*

96,88%, dan F1-Score 96,77%. Akurasi sempurna (100%) dicapai pada kelas ISPA Berat dan Tidak ISPA, yang menunjukkan bahwa model memiliki sensitivitas absolut

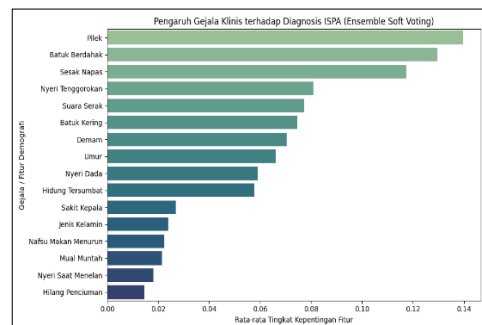
dalam mendeteksi kondisi pasien yang paling berisiko maupun yang sehat.



Gambar 8. Perbandingan Performa Ketiga Model

Visualisasi perbandingan kinerja pada Gambar 8 memberikan gambaran eksplisit mengenai keunggulan model *Ensemble Soft Voting* dibandingkan kedua model tunggal. Berdasarkan data tersebut, *Ensemble Soft Voting* secara konsisten mencatatkan skor tertinggi pada seluruh metrik evaluasi. Peningkatan akurasi sebesar 2,17% terhadap *Random Forest* dan 1,09% terhadap *Extreme Gradient Boosting* menjadi bukti empiris bahwa mekanisme penggabungan probabilitas mampu mengompensasi kelemahan spesifik dari masing-masing algoritma dasar secara efektif. Dari perspektif klinis, pencapaian paling krusial adalah nilai *recall* sempurna (100%) pada kategori Infeksi Saluran Pernafasan Akut Berat yang memberikan jaminan reliabilitas dalam proses triase pasien dengan memastikan tidak adanya kasus berisiko tinggi yang terabaikan oleh sistem. Meskipun akurasi pada kelas Infeksi Saluran Pernafasan Akut Sedang terkoreksi ke angka 90,63% akibat kemiripan pola gejala klinis

dengan kategori berat, sinergi melalui mekanisme penggabungan probabilitas terbukti mampu menghasilkan sistem triase yang lebih *robust* dan objektif.



Gambar 9. Feature Importance Ensemble Soft Voting

Analisis *Feature Importance* pada Gambar 9 menempatkan Pilek, Batuk Berdahak, dan Sesak Napas sebagai tiga indikator klinis paling dominan dalam proses klasifikasi risiko Infeksi Saluran Pernafasan Akut. Temuan ini sejalan dengan literatur medis yang mengonfirmasi bahwa ketiga gejala tersebut merupakan manifestasi klinis utama yang membedakan tingkat keparahan Infeksi Saluran Pernafasan

Akut secara signifikan. Konvergensi keunggulan prediktif dari kedua algoritma dasar pada variabel-variabel tersebut memperkuat validitas ilmiah model yang diusulkan.

Validasi Pakar

Validasi ini bertujuan untuk menilai kesesuaian variabel input, logika klasifikasi, kemampuan model dalam memprediksi tingkat risiko Infeksi Saluran Pernafasan Akut secara tepat, serta relevansi *output* yang dihasilkan terhadap standar praktik klinis nyata dalam penanganan pasien di Rumah Sakit Umum Daerah Latemmamala.

Gambar 10. Lembar Validasi Pakar I (dr. Hasnawati, Sp.PD. FINASIM)

Gambar 11. Lembar Validasi Pakar I (dr. Safriadi, Sp.PD)

Berdasarkan hasil validasi pakar pada Gambar 10 dan Gambar 11, kedua pakar menyatakan bahwa

variabel input yang digunakan telah sesuai dengan standar pemeriksaan klinis Infeksi Saluran Pernafasan Akut di rumah sakit. Selain itu, kategori klasifikasi yang dihasilkan model, yaitu Tidak Infeksi Saluran Pernafasan Akut, Infeksi Saluran Pernafasan Akut Sedang, dan Infeksi Saluran Pernafasan Akut Berat, dinilai relevan dan dapat diimplementasikan sebagai alat bantu triase pasien. Kedua dokter spesialis juga mengonfirmasi bahwa kesalahan klasifikasi yang terjadi pada sebagian kecil pasien Infeksi Saluran Pernafasan Akut Sedang yang terprediksi sebagai Infeksi Saluran Pernafasan Akut Berat merupakan kondisi yang dapat dimaklumi secara klinis, mengingat kedua kategori tersebut memiliki kemiripan manifestasi gejala yang sangat tinggi dalam praktik nyata di lapangan. Hal ini sekaligus memvalidasi bahwa pola prediksi model telah mencerminkan realitas klinis yang sesungguhnya.

Berdasarkan keseluruhan penilaian tersebut, kedua dokter spesialis menegaskan bahwa model *Ensemble Soft Voting* layak digunakan sebagai sistem pendukung keputusan klinis yang membantu tenaga medis dalam melakukan klasifikasi risiko Infeksi Saluran Pernafasan Akut secara lebih cepat, objektif, dan akurat di lingkungan Rumah Sakit Umum Daerah Latemmamala.

D. KESIMPULAN DAN SARAN

Penelitian ini berhasil membangun model klasifikasi risiko Infeksi Saluran Pernafasan Akut berbasis *Ensemble Soft Voting* menggunakan 456 data rekam medis

pasien Rumah Sakit Umum Daerah Latemmamala periode 2020–2025. Algoritma *Random Forest* sebagai model dasar pertama mencatatkan akurasi sebesar 94,57%, presisi 94,54%, recall 94,58%, dan F1-Score 94,55% dengan kemampuan generalisasi yang baik dan capaian sempurna pada kategori Tidak Infeksi Saluran Pernafasan Akut. Algoritma Extreme Gradient Boosting sebagai model dasar kedua memberikan performa yang lebih progresif dengan akurasi 95,65%, presisi 96,08%, recall 95,83%, dan F1-Score 95,69%, dengan keunggulan dalam mendeteksi kategori Infeksi Saluran Pernafasan Akut Berat meskipun masih terdapat tantangan pada klasifikasi Infeksi Saluran Pernafasan Akut Sedang akibat kemiripan manifestasi gejala klinis dengan kategori Infeksi Saluran Pernafasan Akut Berat.

Penerapan metode *Ensemble Soft Voting* yang mengintegrasikan probabilitas prediksi dari kedua algoritma tersebut terbukti menghasilkan performa terbaik dengan akurasi 96,74%, presisi 96,97%, recall 96,88%, dan F1-Score 96,77%, mengungguli kedua model tunggal secara konsisten pada seluruh metrik evaluasi. Pencapaian nilai recall sempurna (1,00) pada kategori ISPA Berat dan Tidak ISPA menjamin tidak adanya pasien berisiko tinggi yang luput dari deteksi, yang memiliki signifikansi klinis tertinggi dalam proses triase. Temuan ini diperkuat oleh validasi dua dokter spesialis penyakit dalam yang mengonfirmasi bahwa model telah mencerminkan realitas klinis dan layak diimplementasikan sebagai alat bantu pengambilan keputusan

klinis di Rumah Sakit Umum Daerah Latemmamala.

Pada penelitian selanjutnya, disarankan agar model dikembangkan dengan menambahkan variabel klinis yang lebih lengkap seperti data laboratorium dan tanda vital, memperluas cakupan dataset dari beberapa fasilitas kesehatan, serta mengeksplorasi teknik *ensemble* lainnya seperti *Stacking* atau *Blending*. Implementasi model dalam bentuk sistem berbasis web atau aplikasi *mobile* juga dapat menjadi langkah konkret untuk memperluas manfaatnya bagi tenaga medis di lapangan.

E. REFERENSI

- [1] D. Zebua, I. Alfionita, Y. Lawa, D. Siregar, and L. Harefa, “Acute respiratory infection and its associated factors among children under five years,” *Enferm. Clínica*, vol. 33, pp. S50–S54, Mar. 2023, doi: 10.1016/j.enfcli.2023.01.010.
- [2] T. B. Purnama, K. Wagatsuma, and R. Saito, “Prevalence and risk factors of acute respiratory infection and diarrhea among children under 5 years old in low-middle wealth household, Indonesia,” *Infect. Dis. Poverty*, vol. 14, no. 1, p. 13, Feb. 2025, doi: 10.1186/s40249-025-01286-9.
- [3] Trurly Santika, Salut Muhidin, Sugeng Budiharta, Kerrie A. Wilson, and Matthew J. Struebig, “Deterioration of respiratory health following changes to land cover and climate in Indonesia,” *One Earth*, vol. 6, pp. 290–302,

- Mar. 2023, doi: 10.1016/j.oneear.2023.02.012.
- [4] M. Islam, K. Islam, K. Dalal, and M. D. Hossain Hawlader, "In-house environmental factors and childhood acute respiratory infections in under-five children: a hospital-based matched case-control study in Bangladesh," *BMC Pediatr.*, vol. 24, p. 38, Jan. 2024, doi: 10.1186/s12887-024-04525-4.
- [5] R. A. Wulandari, S. Fauzia, and F. Kurniasari, "Investigations on the risk factors of Acute Respiratory Infections (ARIs) among under-five children in Depok City, Indonesia," *Ann. Ig. Med. Prev. E Comunita*, vol. 36, no. 1, pp. 15–25, 2024, doi: 10.7416/ai.2023.2580.
- [6] T. Z. Yehuala, B. M. Fente, S. M. Wubante, and N. M. Derseh, "Exploring machine learning algorithms to predict acute respiratory tract infection and identify its determinants among children under five in Sub-Saharan Africa," *Front. Pediatr.*, vol. 12, p. 1388820, Nov. 2024, doi: 10.3389/fped.2024.1388820.
- [7] X. Yang, Y. Li, L. Liu, and Z. Zang, "Prediction of respiratory diseases based on random forest model," *Front. Public Health*, vol. 13, p. 1537238, Feb. 2025, doi: 10.3389/fpubh.2025.1537238.
- [8] C. C. B. Lirna, T. Trimono, and A. T. Damaliana, "Employee Voluntary Attrition Prediction At Pt.Xyz: Ensemble Machine Learning Approach With Soft Voting Classifier," *J. Tek. Inform. Jutif*, vol. 5, no. 5, pp. 1231–1239, Oct. 2024, doi: 10.52436/1.jutif.2024.5.5.2007.
- [9] A. Munandar, W. Maulana Baihaqi, and A. Nurhopipah, "A Soft Voting Ensemble Classifier to Improve Survival Rate Predictions of Cardiovascular Heart Failure Patients," *Ilk. J. Ilm.*, vol. 15, no. 2, pp. 344–352, Aug. 2023, doi: 10.33096/ilkom.v15i2.1632.344-352.
- [10] A. Manconi, G. Armano, M. Gnocchi, and L. Milanesi, "A Soft-Voting Ensemble Classifier for Detecting Patients Affected by COVID-19," *Appl. Sci.*, vol. 12, no. 15, p. 7554, Jul. 2022, doi: 10.3390/app12157554.
- [11] M. Azad, T. Hasan, and M. Moshkov, "A novel ensemble learning method using majority based voting of multiple selective decision trees," *Computing*, vol. 107, Dec. 2024, doi: 10.1007/s00607-024-01394-8.
- [12] D. Dhiyaussalam, A. Yusuf, I. Wardiah, and N. L. Putri, "Predicting Respiratory Conditions Using Random Forest and XGBoost," *J. Inf. Syst. Inform.*, vol. 7, no. 2, pp. 1642–1657, Jun. 2025, doi: 10.51519/journalisi.v7i2.1124.
- [13] W. Hong et al., "A Comparison of XGBoost, Random Forest, and Nomograph for the Prediction of Disease Severity in Patients With COVID-19 Pneumonia: Implications of Cytokine and Immune Cell Profile," *Front. Cell. Infect. Microbiol.*, vol. 12, p. 819267, Apr. 2022, doi: 10.3389/fcimb.2022.819267.

- [14] T. Chen and C. Guestrin, "XGBoost: A Scalable Tree Boosting System," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, Aug. 2020, pp. 785–794. doi: 10.1145/2939672.2939785.
- [15] P. Mahajan, S. Uddin, F. Hajati, and M. A. Moni, "Ensemble Learning for Disease Prediction: A Review," *Healthcare*, vol. 11, p. 1808, Jun. 2023, doi: 10.3390/healthcare11121808.
- [16] I. Markoulidakis and G. Markoulidakis, "Probabilistic Confusion Matrix: A Novel Method for Machine Learning Algorithm Generalized Performance Analysis," *Technologies*, vol. 12, no. 11, 2024, doi: 10.3390/technologies12070113.
- [17] S. Sathyanarayanan and B. R. Tantri, "Confusion Matrix-Based Performance Evaluation Metrics," *Afr. J. Biomed. Res.*, vol. 27, no. 4S, pp. 4023–4031, Nov. 2024, doi: 10.53555/AJBR.v27i4S.4345.