PENERAPAN TEKNIK *STACKING ENSEMBLE LEARNING* UNTUK IDENTIFIKASI INFEKSI SALURAN KEMIH

Daffa Pratama, Sucipto, Barry Ceasar Octariadi
Teknik Informatika, Universitas Muhammadiyah Pontianak
Jl. Jend. Ahmad Yani No. 111, Kota Pontianak, Kalimantan Barat, Indonesia
daffapratama28052003@gmail.com

ABSTRAK

Infeksi Saluran Kemih (ISK) adalah infeksi pada sistem urinaria yang umum terjadi, dengan prevalensi global mencapai 404,6 juta kasus pada tahun 2019 dan berpotensi menimbulkan komplikasi serius seperti gagal ginjal dan sepsis. Diagnosis konvensional melalui kultur urin memerlukan waktu inkubasi 24–48 jam, bersifat tidak spesifik, dan berpotensi menunda terapi serta memperparah resistensi antibiotik. Penelitian ini bertujuan menerapkan dan mengevaluasi teknik *Stacking Ensemble Learning* untuk identifikasi ISK berdasarkan data urinalisis, serta membandingkan performanya dengan model tunggal. Model dibangun dengan menggabungkan *Naïve Bayes*, *Random Forest*, dan *XGBoost* sebagai *base learner* serta *Logistic Regression* sebagai *meta learner* melalui pendekatan *Out-of-Fold (OOF)* dengan *Stratified 5-Fold Cross-Validation*. Dataset *Urinalysis Test Results* yang terdiri dari 1.436 sampel dari *Kaggle* diproses melalui *encoding*, normalisasi, SMOTE, dan *hyperparameter tuning* menggunakan Optuna. Model *stacking* mencapai akurasi 96,49%, presisi 71,43%, *recall* 62,50%, *F1-Score* 66,67%, dan *ROC-AUC* 89,52%, mengungguli seluruh model tunggal dalam keseimbangan presisi dan *recall* serta konsistensi performa pada berbagai skenario pembagian data.

Kata kunci : *Stacking Ensemble Learning, Naïve Bayes, Random Forest, XGBoost, Logistic Regression, ISK*

1. PENDAHULUAN

Infeksi Saluran Kemih (ISK) merupakan salah satu infeksi bakteri terbanyak yang menyerang sistem urinaria, dengan *Escherichia coli* sebagai patogen utama penyebabnya [1]. Kondisi ini dapat memengaruhi seluruh segmen saluran kemih, mulai dari uretra hingga ginjal, sehingga berdampak nyata terhadap kualitas hidup penderita. Tingginya beban penyakit ini tercermin dari data *Scientific Reports* yang mencatat 404,6 juta kasus ISK di seluruh dunia pada tahun 2019 [2]. Tanpa penanganan yang cepat dan tepat, ISK berpotensi berkembang menjadi kondisi yang mengancam jiwa, termasuk gagal ginjal dan sepsis.

Metode diagnosis konvensional yang menjadi standar emas saat ini, yakni kultur urin, memerlukan waktu inkubasi antara 24 hingga 48 jam dan tidak selalu mampu mendeteksi seluruh jenis patogen penyebab ISK [3]. Di sisi lain, pemeriksaan urinalisis sebagai skrining awal juga memiliki keterbatasan spesifisitas yang cukup berarti [3]. Keterbatasan ini sering menunda terapi yang tepat dan berkontribusi terhadap meningkatnya resistensi antibiotik [3]. Dari sisi komputasi, pendekatan *machine learning*, khususnya *stacking ensemble learning*, telah menunjukkan efektivitas tinggi pada berbagai permasalahan klasifikasi medis [4], [5]. *Stacking* menggabungkan prediksi dari beberapa *base learner* menggunakan *meta learner* untuk menghasilkan prediksi yang lebih akurat dan stabil [6], [7]. Meskipun demikian, penerapannya secara spesifik untuk identifikasi ISK berdasarkan data urinalisis masih terbatas [8], [9]

Berdasarkan rumusan permasalahan tersebut, penelitian ini bertujuan untuk mengimplementasikan teknik *stacking ensemble learning* dalam mengidentifikasi pasien penyakit ISK berdasarkan hasil pemeriksaan urinalisis. Selain itu, penelitian ini juga bertujuan untuk mengetahui tingkat performa model tersebut serta membandingkannya secara komparatif dengan algoritma tunggal pembentuknya.

Sebagai rencana penyelesaian masalah, penelitian ini mengusulkan sebuah kerangka kerja komputasi prediktif berbasis *machine learning* yang andal. Pendekatan yang dilakukan mencakup pemanfaatan data klinis urinalisis yang diproses melalui tahapan rekayasa data, termasuk penanganan distribusi kelas yang tidak seimbang agar model terhindar dari bias prediksi. Inti dari penyelesaian masalah ini berpusat pada perancangan arsitektur *stacking ensemble* bertingkat yang dioptimasi secara otomatis. Melalui strategi ini, keputusan klasifikasi tidak hanya bergantung pada satu algoritma pembelajar, melainkan mengintegrasikan kekuatan beberapa algoritma sekaligus untuk menghasilkan sistem deteksi ISK yang lebih tangguh. Hasil akhir dari pemodelan ini kemudian dievaluasi secara komprehensif untuk menguji efektivitas dan validitasnya secara empiris..

2. TINJAUAN PUSTAKA

Penerapan teknik *stacking ensemble learning* dengan model dasar *Decision Tree*, *SVM*, dan *Neural Network* telah dilakukan untuk klasifikasi efek kesehatan akibat pencemaran udara. Model *stacking* mampu mencapai akurasi 99,9% baik pada *dataset* seimbang maupun tidak seimbang, lebih unggul dari *Decision Tree* dan *KNN*, serta terbukti konsisten meningkatkan prediksi meskipun membutuhkan waktu pelatihan lebih lama [4].

Pengembangan *framework stacking* yang menggabungkan *SVM*, *Random Forest*, *KNN*, *ANN*, dan *Logistic Regression* untuk klasifikasi penyakit Parkinson. *Framework* tersebut berhasil mencapai akurasi 96,88%, *recall* 95%, presisi 100%, dan *F1-Score* 97,44%, membuktikan bahwa pendekatan *stacking* lebih unggul dibandingkan algoritma tunggal untuk kasus klasifikasi medis [5].

Penggunaan *kombinasi SVM*, *Random Forest*, *AdaBoost*, dan *SVC* melalui metode *stacking* digunakan pula untuk prediksi penyakit ginjal kronis. Dengan penerapan SMOTE untuk penyeimbangan data, model berhasil mencapai akurasi hingga 100%, menunjukkan efektivitas kombinasi *stacking* dan SMOTE dalam menangani ketidakseimbangan kelas pada data medis [10].

Metode *Equal-Weighted Ensemble* yang menggunakan *Decision Tree*, *XGBoost*, dan *LightGBM* telah diuji untuk prediksi ISK. Model tersebut mencapai akurasi sekitar 85,7% pada tahap validasi, presisi 94,89%, *recall* 75,86%, dan *AUC* 89,53%, memberikan gambaran awal tentang potensi pendekatan *ensemble* untuk identifikasi ISK [8].

Studi perbandingan algoritma tunggal *SVM*, *XGBoost*, dan *LightGBM* pada *dataset* ISK dari klinik di Filipina, dataset yang sama dengan yang digunakan pada penelitian ini. *SVM* dengan kernel linear menghasilkan akurasi tertinggi sebesar 98,25%, diikuti *XGBoost* dan *LightGBM* pada 97,95% [9].

Berdasarkan kajian-kajian tersebut, metode *stacking ensemble learning* secara konsisten memberikan performa lebih unggul dibandingkan model tunggal [4], [5], [10]. Namun, penelitian yang secara spesifik menerapkan *stacking* untuk identifikasi ISK berbasis data urinalisis masih terbatas [8], [9]. Penelitian ini mengusulkan kombinasi *Naïve Bayes*, *Random Forest*, dan *XGBoost* sebagai *base learner* dengan *Logistic Regression* sebagai *meta learner*, dilengkapi *hyperparameter tuning* berbasis Optuna dan penanganan ketidakseimbangan kelas menggunakan SMOTE

2.1. Stacking Ensemble Learning

Ensemble learning meningkatkan kinerja klasifikasi dengan menggabungkan beberapa model dasar (*base learner*) untuk meminimalkan kelemahan individu seperti bias dan varians [6]. Salah satu tekniknya adalah *stacking*, yang mengintegrasikan berbagai algoritma melalui sebuah *meta learner* [7]. Dalam arsitektur ini, prediksi dari model-model dasar digunakan sebagai fitur masukan baru untuk dilatih kembali oleh *meta learner*, sehingga sistem dapat mempelajari kombinasi terbaik guna menghasilkan prediksi akhir yang lebih akurat dibandingkan model tunggal [5], [6].

2.2. Naïve Bayes

Naïve Bayes adalah algoritma klasifikasi probabilistik yang mengasumsikan independensi antar fitur [11]. Algoritma ini sederhana, efisien secara komputasi, dan memberikan hasil yang cukup akurat bahkan pada *dataset* kecil [11], [12]. Dalam penelitian ini digunakan varian *Gaussian Naïve Bayes* untuk data numerik hasil normalisasi.

2.3. Random Forest

Random Forest adalah algoritma *ensemble* berbasis *bagging* yang membangun banyak pohon keputusan dari sampel *bootstrap* dan menggabungkan hasilnya melalui *voting* mayoritas [6]. Algoritma ini mampu menangani dimensi tinggi dan meminimalkan *overfitting* dengan mengurangi korelasi antar pohon.

2.4. XGBoost

XGBoost (Extreme Gradient Boosting) adalah algoritma *ensemble* berbasis *gradient boosting* yang menambahkan pohon secara berurutan untuk meminimalkan residual model sebelumnya [10]. *XGBoost* memanfaatkan regularisasi L1/L2 sehingga tangguh terhadap *overfitting* dan skalabel pada *dataset* besar [13].

2.5. Logistic Regression

Logistic Regression adalah metode klasifikasi biner yang menggunakan fungsi sigmoid untuk memetakan kombinasi linear fitur ke probabilitas. Sifatnya yang sederhana dan mudah diinterpretasi menjadikannya pilihan yang efektif sebagai *meta learner* dalam kerangka *stacking* [14].

3. METODE PENELITIAN

Penelitian ini mengimplementasikan metode *stacking ensemble learning* untuk mengidentifikasi penyakit infeksi saluran kemih berdasarkan data urinalisis. Arsitektur model yang dibangun mengombinasikan *Naïve Bayes*, *Random Forest*, dan *XGBoost* sebagai *base learner* dengan *Logistic Regression* sebagai *meta learner*. Proses eksperimen dilakukan menggunakan bahasa pemrograman Python, mencakup tahapan pra-pemrosesan data dengan SMOTE untuk menangani ketidakseimbangan kelas, serta optimasi *hyperparameter* menggunakan *framework* Optuna guna mencapai performa klasifikasi yang optimal.

3.1. Pengumpulan Data

Dataset yang digunakan dalam penelitian ini adalah *Urinalysis Test Results* yang bersumber dari repositori Kaggle. Data tersebut merupakan hasil pengujian urinalisis yang dikumpulkan dari sebuah klinik di Northern Mindanao, Filipina. Dataset ini mencakup sebanyak 1.436 sampel medis yang terdiri dari 15 variabel fitur, yang kemudian digunakan sebagai dasar untuk melakukan identifikasi penyakit infeksi saluran kemih. Detail mengenai variabel-variabel tersebut disajikan pada Tabel 1.

Tabel 1. Fitur dan keterangan dataset

No	Nama Variabel	Keterangan	Tipe Data
1	*Age*	Usia pasien	Numerik
2	*Gender*	Jenis kelamin pasien	Kategorikal (non-ordinal)
3	*Color*	Warna urin	Kategorikal ordinal
4	*Transparency*	Tingkat kejernihan urin	Kategorikal ordinal
5	*Glucose*	Kadar glukosa urin	Kategorikal ordinal
6	*Protein*	Kadar protein urin	Kategorikal ordinal
7	*pH*	pH urin	Numerik
8	*Specific Gravity*	Berat jenis urin	Numerik
9	*WBC*	Jumlah leukosit urin (per lapang pandang)	Kategorikal ordinal
10	*RBC*	Jumlah eritrosit urin (per lapang pandang)	Kategorikal ordinal
11	*Epithelial Cells*	Jumlah sel epitel	Kategorikal ordinal
12	*Mucous Threads*	Jumlah mukus	Kategorikal ordinal
13	*Amorphous Urates*	Kristal urat	Kategorikal ordinal
14	*Bacteria*	Jumlah bakteri dalam urin	Kategorikal ordinal
15	*Diagnosis*	Status ISK (0 = Negatif, 1 = Positif)	Kategorikal (target)

Dari Tabel 1 dapat dilihat bahwa dataset memiliki kombinasi fitur numerik dan kategorikal, dengan variabel target Diagnosis yang bersifat biner.

3.2. Preprocessing Data

Tahapan pra-pemrosesan data dilakukan melalui beberapa langkah sistematis untuk menjamin kualitas data sebelum proses pelatihan model. Pertama, penanganan *missing value* dilakukan dengan menghapus baris yang mengandung nilai kosong pada kolom *Color* dikarenakan jumlahnya yang sangat minim, diikuti dengan penghapusan data duplikat untuk menjaga integritas data. Selanjutnya, dilakukan *encoding* variabel kategorikal yang mencakup *ordinal encoding* untuk variabel ordinal, *one-hot encoding* untuk fitur *Gender*, serta *binary encoding* untuk variabel target. Seluruh data numerik kemudian dinormalisasi menggunakan *MinMaxScaler* ke dalam rentang [0,1] guna menyelaraskan skala fitur.

Data kemudian dibagi secara *stratified* dengan rasio 80:20, menghasilkan 1.137 sampel data latih dan 285 sampel data uji. Untuk mengatasi masalah ketidakseimbangan kelas pada data latih, teknik *Synthetic Minority Over-sampling Technique* (SMOTE) diterapkan. Efektivitas penanganan ketidakseimbangan kelas ini ditunjukkan pada Tabel 2.

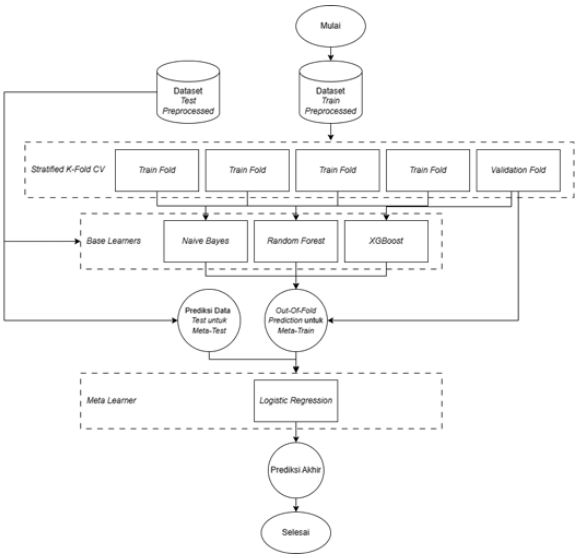
Tabel 2. Distribusi kelas diagnosis data latih sebelum dan sesudah SMOTE

No	Kondisi Data	Negatif	Positif
1	Sebelum SMOTE	1.072	65
2	Setelah SMOTE	1.072	1.072

Dari Tabel 2 terlihat bahwa kelas positif yang semula hanya berjumlah 65 sampel berhasil diseimbangkan menjadi 1.072 sampel setelah penerapan SMOTE, sehingga distribusi kelas menjadi seimbang untuk proses pelatihan model

3.3. Perancangan Model Stacking

Perancangan model *stacking* dalam penelitian ini menerapkan pendekatan *Out-of-Fold* (OOF) dengan menggunakan *Stratified 5-Fold Cross-Validation*. Pada proses ini, setiap *base learner* yang terdiri dari *Naïve Bayes*, *Random Forest*, dan *XGBoost* dilatih untuk menghasilkan prediksi OOF yang berfungsi sebagai *meta-features*. Fitur baru tersebut kemudian digunakan sebagai masukan bagi *meta learner Logistic Regression* untuk menghasilkan prediksi akhir. Guna mencapai performa optimal, *hyperparameter tuning* berbasis Optuna diterapkan pada algoritma *Random Forest*, *XGBoost*, dan *Logistic Regression*. Alur kerja perancangan dan integrasi model ini secara detail ditunjukkan pada Gambar 1.



Gambar 1. Arsitektur model stacking ensemble learning

Berdasarkan arsitektur pada Gambar 1, dapat dilihat bahwa proses dimulai dari pengolahan dataset urinalisis yang telah melalui tahap pra-pemrosesan dan optimasi menggunakan Optuna. Model *stacking* mengintegrasikan tiga algoritma *base learner* yang bekerja secara paralel untuk menghasilkan kumpulan prediksi awal. Prediksi-prediksi tersebut kemudian digabungkan menjadi fitur baru (*meta-features*) yang dipelajari kembali oleh *meta learner* untuk meminimalkan kesalahan prediksi dari model individu dan menghasilkan diagnosis akhir yang lebih akurat.

3.4. Evaluasi Model

Evaluasi dilakukan menggunakan data uji dengan metrik akurasi, presisi, *recall*, *F1-Score*, dan *ROC-AUC*. *Confusion matrix* dan *classification report* digunakan untuk analisis lebih mendalam. Pemilihan metrik ini didasarkan pada kemampuannya menggambarkan performa klasifikasi secara komprehensif, khususnya pada kondisi kelas tidak seimbang [15].

4. HASIL DAN PEMBAHASAN

4.1. Hasil Hyperparameter Tuning

Tabel 3. Hyperparameter terbaik untuk random forest

No	*Hyperparameter*	Nilai
1	*n_estimators*	187
2	*max_depth*	30
3	*min_samples_split*	7
4	*min_samples_leaf*	1
5	*max_features*	sqrt

Dari tabel 3, dapat dilihat konfigurasi yang menghasilkan model *Random Forest* yang lebih kuat dalam mengenali pola pada dataset, karena jumlah pohon yang cukup besar dan kedalaman yang luas memungkinkan model menangkap kompleksitas data dengan lebih baik

Tabel 4. Hyperparameter terbaik untuk XGBoost

No	*Hyperparameter*	Nilai
1	*n_estimators*	196
2	*learning_rate*	0.09223412739339064
3	*max_depth*	15
4	*subsample*	0.8890009848362174
5	*colsample_bytree*	0.5657753162787876
6	*gamma*	0.8007350607661885

Dari tabel 4, dapat dilihat konfigurasi yang memungkinkan *XGBoost* menghasilkan prediksi yang stabil dengan mengontrol keseimbangan antara kedalaman model, tingkat pembelajaran, dan variasi sampel pada setiap pohon

Tabel 5. Hyperparameter terbaik logistic regression

No	*Hyperparameter*	Nilai
1	C	4.519782848714315
2	solver	lbfgs
3	tol	0.0003558804589364697

Dari tabel 5, dapat dilihat dengan nilai C yang cukup besar, model cenderung memiliki regularisasi yang lebih longgar sehingga mampu beradaptasi lebih baik terhadap pola data, sementara penggunaan *solver* lbfgs mendukung proses optimisasi pada dataset berukuran besar dan multikelas

4.2. Hasil Evaluasi Model Stacking

Dari Tabel 4, model *stacking* memperoleh akurasi 96,49%, menunjukkan kemampuan yang kuat dalam mengklasifikasikan sampel. Nilai *ROC-AUC* sebesar 89,52% mengindikasikan kemampuan pemisahan kelas yang sangat baik. Nilai *recall* 62,50% menunjukkan bahwa model berhasil mengidentifikasi lebih dari setengah kasus positif ISK, sementara presisi 71,43% menunjukkan sebagian besar prediksi positif adalah benar. Kombinasi keduanya menghasilkan *F1-Score* 66,67.

Tabel 6. Hasil evaluasi model stacking ensemble learning

Metrik	Nilai
Akurasi	96,49%
Presisi	71,43%
Recall	62,50%
F1-Score	66,67%
ROC-AUC	89,52%

4.3. Perbandingan Kinerja Model

Tabel 7. Perbandingan kinerja model pada split dataset 80:20 tanpa SMOTE

Model	Akurasi (%)	Presisi (%)	*Recall* (%)	*F1-Score* (%)	*ROC-AUC* (%)
Stacking	95.09	62.50	31.25	41.67	81.53
Naïve Bayes	89.82	32.43	75.00	45.28	77.42
Random Forest	96.14	72.73	50.00	59.26	83.50
XGBoost	96.14	85.71	37.50	52.17	80.79
Logistic Regression	95.79	100.00	25.00	40.00	78.07

Berdasarkan data pada Tabel 7, terlihat bahwa pada kondisi dataset tanpa penanganan ketidakseimbangan kelas, algoritma berbasis pohon seperti *Random Forest* dan *XGBoost* menghasilkan akurasi dan *F1-Score* yang lebih tinggi dibandingkan model lainnya. Namun, hampir seluruh model pada skenario ini menunjukkan nilai *recall* yang sangat rendah (di bawah 50%, kecuali *Naïve Bayes*), yang mengindikasikan bahwa model kesulitan dalam mengidentifikasi pasien positif infeksi saluran kemih karena minimnya sampel kelas minoritas dalam proses pelatihan.

Tabel 8. Perbandingan kinerja model pada split dataset 80:20 dengan SMOTE

Model	Akurasi (%)	Presisi (%)	*Recall* (%)	*F1-Score* (%)	*ROC-AUC* (%)
Stacking	96,49	71,43	62,50	66,67	89,52
Naïve Bayes	84,56	19,57	56,25	29,03	67,77
Random Forest	96,14	72,73	50,00	59,26	86,11
XGBoost	95,79	66,67	50,00	57,14	87,27
Logistic *Regression*	85,61	23,40	68,75	34,92	78,32

Berdasarkan hasil pengujian pada skenario pembagian data 80:20 di tabel 7 dan tabel 8, model *Stacking Ensemble Learning* menunjukkan performa

yang paling unggul dan konsisten, khususnya setelah dilakukan penanganan ketidakseimbangan kelas menggunakan SMOTE. Pada Tabel 8, model *stacking* berhasil mencapai akurasi tertinggi sebesar 96,49% dengan nilai *F1-Score* 66,67% dan *ROC-AUC* 89,52%. Pencapaian ini mengungguli model tunggal lainnya karena *stacking* mampu meningkatkan nilai *recall* secara signifikan menjadi 62,50% dibandingkan saat tanpa SMOTE yang hanya sebesar 31,25%. Peningkatan ini memiliki makna klinis yang penting dalam diagnosis medis karena mampu mengurangi potensi terjadinya *false negative*, sehingga lebih banyak pasien ISK yang dapat teridentifikasi secara akurat.

Jika ditinjau dari performa masing-masing algoritma tunggal, terlihat adanya perbedaan karakteristik yang mencolok. *Naïve Bayes* dan *Logistic Regression* cenderung menghasilkan presisi yang rendah pada dataset dengan SMOTE yang mengindikasikan tingginya angka *false positive*. Di sisi lain, *Random Forest* dan *XGBoost* tampil cukup stabil dengan akurasi di atas 95%, namun keduanya belum mampu menyamai keseimbangan antara presisi dan *recall* yang dicapai oleh model *stacking*. Fenomena ini menunjukkan bahwa penggunaan SMOTE sangat krusial; tanpa SMOTE, model cenderung hanya memprediksi kelas mayoritas (negatif), namun dengan SMOTE yang dikombinasikan dengan arsitektur *stacking*, model dapat mempelajari batasan keputusan yang lebih baik antar kelas.

Keunggulan model *stacking* ini disebabkan oleh kemampuannya dalam mengintegrasikan kekuatan berbagai algoritma: kemampuan probabilistik *Naïve Bayes*, kekokohan *Random Forest* terhadap *overfitting*, serta efisiensi *XGBoost* dalam menangani pola data yang kompleks. *Meta learner Logistic Regression* kemudian mempelajari cara optimal untuk menggabungkan prediksi dari ketiga *base learner* tersebut. Penggunaan pendekatan *Out-of-Fold* (OOF) dalam proses ini juga memastikan bahwa fitur yang dipelajari oleh *meta learner* bersifat tidak bias. Secara keseluruhan, integrasi teknik *stacking*, SMOTE, dan optimasi *hyperparameter* terbukti memberikan performa klasifikasi yang lebih andal dan stabil untuk identifikasi ISK dibandingkan hanya mengandalkan satu model tunggal [6], [7].

5. KESIMPULAN DAN SARAN

Penelitian ini berhasil menerapkan teknik *stacking ensemble learning* untuk identifikasi ISK berdasarkan data urinalisis dengan mengombinasikan *Naïve Bayes*, *Random Forest*, dan *XGBoost* sebagai *base learner* serta *Logistic Regression* sebagai *meta learner* melalui pendekatan *Out-of-Fold* dengan *Stratified 5-Fold Cross-Validation*. Model *stacking* mencapai performa terbaik dibandingkan seluruh model tunggal dengan akurasi 96,49%, presisi 71,43%, *recall* 62,50%, *F1-Score* 66,67%, dan *ROC-AUC* 89,52%, serta konsisten unggul pada berbagai skenario pembagian data. Temuan ini menunjukkan bahwa pendekatan *stacking* mampu menggabungkan kelebihan berbagai algoritma dasar sehingga menghasilkan prediksi yang lebih akurat dan seimbang, menjadikannya pendekatan efektif untuk mendukung identifikasi awal ISK berbasis data urinalisis. Untuk penelitian selanjutnya, disarankan menggunakan *dataset* dengan distribusi kelas yang lebih seimbang, mengeksplorasi metode penanganan ketidakseimbangan lain seperti SMOTE-ENN atau ADASYN, serta melakukan validasi eksternal dengan data dari fasilitas kesehatan yang berbeda.

DAFTAR PUSTAKA

[1] G. Mancuso, A. Midiri, E. Gerace, M. Marra, S. Zummo, and C. Biondo, "Urinary Tract Infections: The Current Scenario and Future Prospects," Apr. 01, 2023, *MDPI*. doi: 10.3390/pathogens12040623.

[2] Y. He *et al.*, "Epidemiological trends and predictions of urinary tract infections in the global burden of disease study 2021," *Sci. Rep.*, vol. 15, no. 1, Dec. 2025, doi: 10.1038/s41598-025-89240-5.

[3] R. Patel *et al.*, "Envisioning Future Urinary Tract Infection Diagnostics," *Clinical Infectious Diseases*, vol. 74, no. 7, pp. 1284–1292, Apr. 2022, doi: 10.1093/cid/ciab749.

[4] B. Sunarko *et al.*, "Penerapan Stacking Ensemble Learning untuk Klasifikasi Efek Kesehatan Akibat Pencemaran Udara," *Edu Komputika*, vol. 10, no. 1, 2023, [Online]. Available: http://journal.unnes.ac.id/sju/index.php/edukom

[5] Y. Yang, L. Wei, Y. Hu, Y. Wu, L. Hu, and S. Nie, "Classification of Parkinson's disease based on multi-modal features and stacking ensemble learning," *J. Neurosci. Methods*, vol. 350, Feb. 2021, doi: 10.1016/j.jneumeth.2020.109019.

[6] I. D. Mienye and Y. Sun, "A Survey of Ensemble Learning: Concepts, Algorithms, Applications, and Prospects," 2022, *Institute of Electrical and Electronics Engineers Inc.* doi: 10.1109/ACCESS.2022.3207287.

[7] A. Alazba and H. Aljamaan, "Software Defect Prediction Using Stacking Generalization of Optimized Tree-Based Ensembles," *Applied Sciences (Switzerland)*, vol. 12, no. 9, May 2022, doi: 10.3390/app12094577.

[8] S. Farashi and H. E. Momtaz, "Prediction of urinary tract infection using machine learning methods: a study for finding the most-informative variables," *BMC Med. Inform. Decis. Mak.*, vol. 25, no. 1, Dec. 2025, doi: 10.1186/s12911-024-02819-2.

[9] G. Airlangga, "Efficacy of Machine Learning Techniques in Diagnosing Urinary Tract Infections: A Study Utilizing a Philippine Clinical Dataset," *Jurnal Informatika Ekonomi Bisnis*, pp. 251–255, Mar. 2024, doi: 10.37034/infeb.v6i1.855.

[10] D. Chhabra, M. Juneja, and G. Chutani, "An Efficient Ensemble-based Machine Learning approach for Predicting Chronic Kidney Disease," *Curr. Med. Imaging Rev.*, vol. 20, May 2023, doi: 10.2174/1573405620666230508104538.

[11] O. Peretz, M. Koren, and O. Koren, "Naive Bayes classifier – An ensemble procedure for recall and precision enrichment," *Eng. Appl. Artif. Intell.*, vol. 136, Oct. 2024, doi: 10.1016/j.engappai.2024.108972.

[12] Q. An, S. Rahman, J. Zhou, and J. J. Kang, "A Comprehensive Review on Machine Learning in Healthcare Industry: Classification, Restrictions, Opportunities and Challenges," May 01, 2023, *MDPI*. doi: 10.3390/s23094178.

[13] M. Wiens, A. Verone-Boyle, N. Henscheid, J. T. Podichetty, and J. Burton, "A Tutorial and Use Case Example of the eXtreme Gradient Boosting (XGBoost) Artificial Intelligence Algorithm for Drug Development Applications," *Clin. Transl. Sci.*, vol. 18, no. 3, Mar. 2025, doi: 10.1111/cts.70172.

[14] N. Afiqah, M. Zaini, and M. K. Awang, "Performance Comparison between Meta-classifier Algorithms for Heart Disease Classification," *IJACSA) International Journal of Advanced Computer Science and Applications*, vol. 13, no. 10, pp. 323–328, 2022, [Online]. Available: www.ijacsa.thesai.org

[15] S. A. Hicks *et al.*, "On evaluation metrics for medical applications of artificial intelligence," *Sci. Rep.*, vol. 12, no. 1, Dec. 2022, doi: 10.1038/s41598-022-09954-8.