



Optimasi Parameter Model *LightGBM* Menggunakan Algoritma *Grey Wolf Optimizer* untuk Prediksi Penyakit Ginjal Kronis

Muhammad Alfin^{1*}, Alvin Hafiz², Muhammad Budi Akbar³, Adidtya Perdana⁴

¹⁻⁴Ilmu Komputer, Universitas Negeri Medan, Indonesia

*Penulis Korespondensi: mhdalfinaja@mhs.unimed.ac.id

Abstract. Chronic kidney disease is an increasingly prevalent health issue that requires more precise clinical data-based early detection methods to enable timely and appropriate treatment. This study focuses on developing a predictive model for chronic kidney disease using the Light Gradient Boosting Machine (*LightGBM*) algorithm and enhancing its performance through hyperparameter optimization with the Grey Wolf Optimizer (*GWO*). The dataset used originates from public sources and undergoes several preprocessing steps, including missing value imputation, categorical feature encoding, outlier handling, initial feature selection, and stratified data splitting to maintain model quality. Three modeling approaches were evaluated: *LightGBM* with default parameters, *LightGBM* enhanced using Random Search, and *LightGBM* optimized with *GWO*. The experimental results indicate that the baseline model already performs well, Random Search improves accuracy and F1-score, and *GWO* achieves the highest AUC-ROC value despite requiring longer computation time. Significance testing through cross-validation shows that the performance differences among the three models are not statistically significant, suggesting that the observed improvements are not strong enough to determine a definitively superior optimization method. The feature importance analysis highlights that clinical indicators such as creatinine levels, glomerular filtration rate, blood pressure, and urine protein contribute most prominently to the prediction. Overall, the study demonstrates that *LightGBM* is a reliable model for early detection of chronic kidney disease, and hyperparameter optimization still offers added value that can support the development of AI-based clinical decision-support systems.

Keywords: Chronic Kidney Prediction; Grey Wolf Optimizer; Hyperparameter Optimization; *LightGBM*; Machine Learning.

Abstrak. Penyakit ginjal kronis menjadi salah satu isu kesehatan yang terus meningkat dan membutuhkan metode deteksi dini berbasis data klinis yang lebih presisi agar penanganan dapat dilakukan secara cepat dan tepat. Penelitian ini berfokus pada pengembangan model prediksi penyakit ginjal kronis menggunakan algoritma Light Gradient Boosting Machine (*LightGBM*) serta peningkatan kinerjanya melalui optimasi hiperparameter dengan algoritma Grey Wolf Optimizer (*GWO*). Dataset yang digunakan bersumber dari data publik dan diproses melalui tahapan seperti imputasi nilai hilang, pengkodean fitur kategorikal, penanganan outlier, seleksi fitur awal, serta pembagian data secara stratified untuk menjaga kualitas model. Tiga pendekatan diuji, yaitu *LightGBM* dengan parameter standar, *LightGBM* yang ditingkatkan menggunakan Random Search, dan *LightGBM* yang dioptimasi dengan *GWO*. Hasil percobaan menunjukkan bahwa model standar telah memberikan performa yang baik, Random Search mampu meningkatkan akurasi dan F1-score, sementara *GWO* menghasilkan nilai AUC-ROC terbaik meskipun dengan waktu komputasi yang lebih besar. Uji signifikansi melalui cross-validation mengungkapkan bahwa perbedaan performa ketiga model tidak signifikan secara statistik, sehingga peningkatan yang diamati belum cukup kuat untuk menyimpulkan keunggulan salah satu metode secara mutlak. Analisis feature importance menegaskan bahwa fitur-fitur seperti kadar kreatinin, laju filtrasi glomerulus, tekanan darah, dan protein urine memiliki kontribusi terbesar dalam prediksi. Secara keseluruhan, penelitian ini menunjukkan bahwa *LightGBM* merupakan model yang andal untuk deteksi awal penyakit ginjal kronis, dan optimasi hiperparameter tetap memberikan manfaat tambahan yang berpotensi mendukung pengembangan sistem pendukung keputusan klinis berbasis kecerdasan buatan.

Kata kunci: Grey Wolf Optimizer; *LightGBM*; Machine Learning; Optimasi Hiperparameter; Prediksi Ginjal Kronis.

1. LATAR BELAKANG

Penyakit ginjal kronis merupakan masalah kesehatan masyarakat yang semakin meningkat dan memberikan beban klinis serta ekonomi yang besar karena sering terlambat terdiagnosis hingga masuk tahap gagal ginjal akhir. Penelitian-penelitian nasional menunjukkan besarnya kebutuhan akan metode deteksi dini dan prediksi risiko yang andal untuk membantu pengambilan keputusan klinis dan intervensi lebih awal (Nopriani et al., 2024).

Perkembangan *machine learning* telah membuka peluang signifikan untuk membangun model prediksi penyakit berbasis data klinis dan laboratorium. Di Indonesia, beberapa kajian menerapkan teknik pembelajaran mesin untuk prediksi PGK dan menunjukkan bahwa model berbasis pohon keputusan, ensemble, atau metode optimasi fitur dapat meningkatkan akurasi deteksi dibandingkan metode statistik tradisional. Namun, performa model sangat bergantung pada pemilihan dan penyetelan (tuning) hiperparameter (Qodri et al., 2025).

LightGBM (Light Gradient Boosting Machine) adalah algoritma gradient boosting berbasis pohon yang dirancang untuk efisiensi komputasi dan kemampuan menangani dataset besar serta variabel kategori banyak. Di komunitas riset Indonesia, LightGBM semakin sering dipilih untuk tugas klasifikasi medis karena kecepatan pelatihan dan performa yang baik pada data berukuran sedang hingga besar; beberapa studi nasional juga melaporkan bahwa LightGBM unggul dibandingkan XGBoost atau ANN pada beberapa dataset kesehatan bila dilakukan tuning hyperparameter yang tepat (Rizky et al., 2022).

Kinerja model boosting seperti LightGBM sangat bergantung pada pemilihan dan pengaturan hyperparameter (mis. `learning_rate`, `num_leaves`, `max_depth`, `n_estimators`, `bagging_fraction`). Tuning hyperparameter konvensional (grid search, random search, Bayesian optimization) telah banyak digunakan, namun metaheuristik populasi seperti Grey Wolf Optimizer (GWO) menawarkan jalur alternatif yang efektif untuk masalah optimisasi nonlinier dan berdimensi tinggi dengan keuntungan konvergensi yang cepat dan kemampuan menghindari jebakan lokal (Junian et al., 2025).

Berbagai penelitian terdahulu di Indonesia juga menunjukkan hasil positif dari penggunaan algoritma optimasi metaheuristik dalam prediksi penyakit ginjal kronis. Misalnya, (Ermanto & Surojudin, 2025) menggunakan Naive Bayes Classifier berbasis Particle Swarm Optimization (PSO) dan mencapai akurasi confusion matrix sebesar 98,75% dengan nilai AUC 99%. (Damanik et al., 2024) menggabungkan Backpropagation Neural Network dengan PSO dan membuktikan bahwa kombinasi algoritma optimasi dapat menghasilkan kinerja yang lebih baik daripada metode konvensional. Sementara itu, (Purwaningsih & Nurelasari, 2024)

menunjukkan peningkatan akurasi Support Vector Machine hingga 99% melalui penyesuaian parameter, dan (Fitri & Baita, 2025) mengoptimasi Decision Tree untuk prediksi PGK dan menemukan bahwa tuning hiperparameter merupakan faktor kunci dalam meningkatkan performa model.

Di luar bidang spesifik penyakit ginjal kronis, (Azis et al., 2024) membuktikan bahwa Grey Wolf Optimizer efektif dalam meningkatkan kinerja MLP (Multi-Layer Perceptron) untuk klasifikasi, dengan akurasi yang meningkat signifikan setelah optimasi parameter.

Berdasarkan kajian, masih terdapat celah penelitian dalam pengembangan model prediksi penyakit ginjal kronis di Indonesia, khususnya terkait integrasi LightGBM dengan Grey Wolf Optimizer (GWO), dua metode yang terbukti efektif secara terpisah, namun belum dikombinasikan. Selama ini, optimasi LightGBM umumnya menggunakan pendekatan konvensional atau algoritma seperti PSO, padahal GWO menawarkan keunggulan dalam stabilitas dan kemampuan eksplorasi global, sangat sesuai untuk optimasi nonlinier berdimensi tinggi. Mengisi celah tersebut, penelitian ini bertujuan mengembangkan model prediksi penyakit ginjal kronis dengan mengoptimalkan hiperparameter LightGBM menggunakan GWO, guna meningkatkan akurasi, efisiensi, dan robustness dibanding metode konvensional. Secara akademis, penelitian ini diharapkan memperkaya bidang komputasi kesehatan dan mendukung pengembangan sistem pendukung keputusan klinis berbasis AI.

2. KAJIAN TEORITIS

Penyakit Ginjal Kronis

Penyakit ginjal kronis (PGK) merupakan kondisi gangguan struktural atau fungsional pada ginjal yang berlangsung lebih dari tiga bulan dan berdampak signifikan terhadap kesehatan. PGK tidak terbatas pada stadium akhir gagal ginjal (*End Stage Renal Disease/ESRD*), tetapi juga mencakup berbagai komplikasi yang menyertainya sepanjang perjalanan penyakit (Damayantie et al., 2022). Seiring perkembangannya, PGK menyebabkan penurunan fungsi ginjal yang berat dan bertahap hingga tubuh kehilangan kemampuan mempertahankan metabolisme serta menyeimbangkan cairan dan elektrolit, yang pada gilirannya memicu penurunan volume vaskular dan gangguan reabsorpsi di ginjal (Dame et al., 2022).

Model Light Gradient Boosting Machine (LightGBM)

LightGBM (*Light Gradient Boosting Machine*) merupakan *framework gradient boosting* berbasis pohon keputusan yang dikembangkan untuk mencapai kecepatan tinggi, efisiensi memori, serta kinerja optimal pada dataset berskala besar dengan fitur numerik maupun kategorikal. Dalam

prediksi klinis, LightGBM kerap menunjukkan performa kompetitif terutama pada metrik seperti AUC, akurasi, dan F1-score dibandingkan model *ensemble* lainnya, meskipun hasilnya sangat dipengaruhi oleh pengaturan *hyperparameter* kritis (Mawardi et al., 2025).

Keunggulan efisiensinya dicapai melalui dua teknik inovatif, yaitu *Gradient-based One-Side Sampling* (GOSS) dan *Exclusive Feature Bundling* (EFB), yang secara cerdas mempercepat proses pelatihan tanpa melanggar prinsip dasar algoritma berbasis histogram pendekatan inti dalam berbagai *framework Gradient Boosting Decision Tree* (GBDT) (Febriantoro et al., 2023).

Algoritma Grey Wolf Optimizer

Grey Wolf Optimizer (GWO) adalah metode optimasi metaheuristik yang diperkenalkan oleh Seyedali Mirjalili pada tahun 2014, termasuk dalam kelompok *swarm intelligence* dan terinspirasi dari perilaku sosial serta strategi perburuan koloni serigala abu-abu (*Canis lupus*) mulai dari pelacakan, pengepungan, hingga penyerangan terhadap mangsa (Pangestu et al., 2022). Dalam implementasinya, taktik perburuan tersebut diterjemahkan ke dalam mekanisme pencarian solusi secara iteratif, sedangkan hierarki sosial serigala dimodelkan secara matematis: tiga individu terbaik dalam populasi diwakili oleh serigala alfa (α), beta (β), dan delta (δ), sementara individu lainnya berperan sebagai serigala omega (ω), yang mengikuti dan memperbarui posisi berdasarkan tiga pemimpin tersebut (Mahardhika et al., 2025).

GWO telah mendapat respons positif dan penerapan luas dari komunitas peneliti, berkat kemampuannya sebagai salah satu algoritma optimasi metaheuristik berbasis *swarm intelligence* yang terbukti efektif dalam menyelesaikan berbagai masalah kompleks di berbagai bidang, mulai dari rekayasa dan kedokteran, hingga *machine learning* dan bioinformatika (Ardiansyah et al., 2025).

Random Search

Random Search (pencarian acak) adalah teknik tuning hyperparameter yang melakukan sampling acak dari ruang parameter menurut distribusi yang ditentukan. Meski sederhana, Random Search sering kali lebih efisien daripada Grid Search pada ruang parameter berdimensi tinggi karena probabilitasnya menemukan kombinasi yang mendekati optimal dengan jumlah percobaan yang sama terutama bila hanya sebagian kecil parameter yang benar-benar sensitif terhadap kinerja model. Literatur eksperimen di jurnal nasional merekomendasikan Random Search sebagai baseline praktis ketika sumber daya komputasi terbatas, atau dipakai sebagai tahap awal sebelum menerapkan optimasi metaheuristik (seperti GWO) atau metode Bayesian/AutoML yang lebih mahal komputasinya. Namun, Random Search tidak memanfaatkan informasi dari percobaan

sebelumnya sehingga bisa kurang efisien bila ruang pencarian besar dan waktu terbatas (Pratama & Cahyono, 2025).

Evaluasi Model

Evaluasi model klasifikasi dalam banyak penelitian klinis menggunakan confusion matrix dasar dan metrik turunan untuk menilai kualitas klasifikasi. Dari elemen-elemen di atas, metrik yang umum dilaporkan dan rumusnya adalah:

Accuracy (Akurasi) — proporsi prediksi benar dari seluruh sampel:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

Precision (Presisi) — proporsi prediksi positif yang benar:

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

Recall (Sensitivitas / TPR) — proporsi kasus positif yang berhasil dideteksi:

$$Recall = \frac{TP}{TP + FN}$$

F1-score — harmonik rata-rata Precision dan Recall:

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

Specificity (TNR) — proporsi kasus negatif yang benar teridentifikasi:

$$Specificity = \frac{TN}{TN + FP}$$

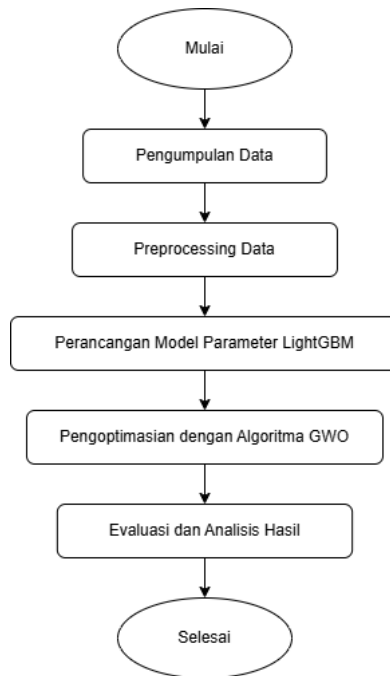
AUC-ROC — area di bawah kurva ROC, mengukur trade-off TPR vs FPR (dimanfaatkan untuk menilai separabilitas model pada berbagai threshold).

Penelitian-penelitian nasional yang mengevaluasi model prediksi penyakit kronis biasanya melaporkan beberapa kombinasi metrik di atas (accuracy, precision, recall, F1, AUC), serta menggunakan teknik validasi seperti k-fold cross-validation untuk estimasi performa yang lebih robust. Confusion matrix juga sering digunakan untuk menghitung nilai-nilai tersebut secara eksplisit dalam setiap lipatan validasi (Anugrah & Arifin, 2025).

3. METODE PENELITIAN

Penelitian ini termasuk dalam kategori penelitian eksperimen komputasional dengan pendekatan kuantitatif, yang bertujuan untuk mengembangkan dan memvalidasi suatu sistem prediksi berbasis kecerdasan buatan melalui simulasi terkendali dalam lingkungan pemrograman. Pendekatan ini dipilih karena memungkinkan pengujian empiris terhadap pengaruh strategi optimasi hiperparameter terhadap kinerja model prediksi, khususnya dalam konteks diagnosis awal penyakit ginjal kronis berdasarkan data klinis pasien. Penelitian tidak

melibatkan subjek manusia secara langsung, melainkan memanfaatkan dataset rekam medis yang telah tersedia secara publik dan dianonimisasi, sehingga fokus utamanya adalah pada konstruksi model, optimasi algoritmik, dan evaluasi performa secara objektif melalui metrik statistik, dengan desain yang menekankan pada replikabilitas, transparansi, dan validitas internal melalui pembagian data secara *stratified* serta penggunaan *hold-out test set* yang independen.



Gambar 1. Diagram Alur Penelitian.

Pengumpulan Data

Data yang digunakan dalam penelitian ini diperoleh dari dataset publik *Chronic Kidney Disease (CKD) Dataset* di platform Kaggle yang berisi catatan medis dari lebih dari 400 pasien dengan 25 variabel klinis meliputi parameter laboratorium (seperti kadar albumin, urea, kreatinin, sodium, hemoglobin), tanda vital (tekanan darah, suhu tubuh), serta riwayat penyakit penyerta (hipertensi, diabetes, penyakit jantung) serta satu label kelas biner (*ckd* atau *notckd*) yang telah diverifikasi secara klinis. Sebelum digunakan, metadata dan dokumentasi dataset dikaji untuk memastikan validitas konten dan kesesuaian dengan tujuan prediksi klinis.

Preprocessing Data

Proses pra-pemrosesan data dilakukan secara sistematis untuk meningkatkan kualitas dan kesiapan dataset sebagai masukan model, dimulai dengan identifikasi dan penanganan nilai hilang menggunakan imputasi berbasis median untuk fitur numerik dan modus untuk fitur kategorikal agar meminimalkan distorsi statistik. Selanjutnya, variabel kategorikal dikodekan menggunakan *Label Encoding* mengingat dukungan bawaan LightGBM terhadap tipe data

kategorikal dan efisiensinya dibanding *One-Hot Encoding* pada dimensi fitur sedang; deteksi dan penanganan *outlier* dilakukan melalui analisis *boxplot* dan *Z-score* dengan teknik *capping* (bukan penghapusan) untuk mempertahankan jumlah sampel klinis yang representatif. Kemudian dilakukan seleksi fitur awal berdasarkan analisis korelasi (Pearson/Spearman) dan *feature importance* dari model baseline LightGBM untuk mengeliminasi redundansi; dan terakhir, data dibagi secara *stratified* menjadi 70% untuk pelatihan, 15% untuk validasi (khusus optimasi GWO), dan 15% untuk pengujian akhir, sehingga proporsi kelas tetap terjaga di setiap subset dan risiko *data leakage* dapat diminimalkan.

Perancangan Model Parameter LightGBM

Model LightGBM dirancang sebagai klasifikasi biner untuk memprediksi status CKD, memanfaatkan keunggulannya dalam kecepatan pelatihan, efisiensi memori, serta kemampuan menangani data campuran (numerik dan kategorikal) secara alami; arsitektur intinya berbasis *Gradient Boosting Decision Tree* dengan *leaf-wise* growth strategy, dan parameter yang menjadi ruang pencarian optimasi meliputi *num_leaves*, *max_depth*, *learning_rate*, *n_estimators*, *min_child_samples*, koefisien regularisasi L1/L2, serta fraksi *feature* dan *bagging* (0.5–1.0), sementara parameter *categorical_feature* secara eksplisit didefinisikan sesuai indeks fitur kategorikal agar LightGBM menerapkan mekanisme pemisahan optimal; fungsi *objective* ditetapkan sebagai 'binary' dengan metrik pelatihan 'binary_logloss' dan 'auc', dan model baseline pertama dibangun dengan konfigurasi *default* untuk menjadi acuan dalam mengukur peningkatan kinerja setelah optimasi.

Pengoptimasian dengan Algoritma GWO

Optimasi hiperparameter dilakukan menggunakan *Grey Wolf Optimizer* (GWO), sebuah algoritma metaheuristik yang mensimulasikan hierarki sosial dan mekanisme perburuan serigala abu-abu, di mana setiap individu dalam populasi mewakili satu kandidat solusi berupa vektor nilai parameter LightGBM, dan *fitness function* didefinisikan sebagai negatif dari *F1-score* yang dihasilkan model pada *validation set*, sehingga proses pencarian secara iteratif mendorong konvergensi ke konfigurasi yang memaksimalkan *F1-score*. Dalam implementasinya, digunakan populasi 20 individu selama maksimal 30 iterasi, dengan pembaruan posisi α (solusi terbaik), β , dan δ mengikuti persamaan adaptif berbasis koefisien *A*, *C*, dan *D*, sementara konvergensi dipantau melalui kurva *fitness* terhadap iterasi. Solusi akhir yaitu posisi α pada iterasi terakhir kemudian digunakan untuk melatih ulang model LightGBM pada gabungan *training* dan *validation set*, menghasilkan model final yang siap dievaluasi secara independen menggunakan *test set*.

Evaluasi dan Analisis Hasil

Evaluasi kinerja model dilakukan secara komprehensif dan objektif menggunakan *test set* yang sepenuhnya terpisah dari proses pelatihan dan optimasi, dengan metrik utama meliputi akurasi, presisi, recall, F1-score, dan AUC-ROC untuk menilai kemampuan prediksi secara holistik khususnya F1-score dan recall yang menjadi prioritas mengingat pentingnya deteksi dini kasus CKD (minimalkan *false negative*); selain itu, *confusion matrix* disusun untuk analisis rinci terhadap distribusi *true/false positive/negative*, dan *5-fold cross-validation* diterapkan pada *training+validation set* untuk mengukur stabilitas model melalui deviasi standar metrik; sebagai validasi tambahan, kinerja model usulan (LightGBM-GWO) dibandingkan secara statistik terhadap dua baseline LightGBM *default* dan LightGBM dengan *Random Search* menggunakan uji *paired t-test* ($\alpha = 0.05$) berdasarkan skor F1-score dari 5 ulangan eksperimen, sehingga keunggulan relatif GWO dapat dikonfirmasi secara signifikan, bukan karena fluktuasi acak.

4. HASIL DAN PEMBAHASAN

Deskripsi Dataset dan Preprocessing

Dataset yang digunakan dalam penelitian ini berasal dari platform Kaggle dengan nama file *Chronic_Kidney_Disease_data.csv*, yang berisi 1.600+ rekam medis pasien terkait kondisi klinis, riwayat penyakit, dan faktor gaya hidup yang berhubungan dengan risiko penyakit ginjal kronis (Chronic Kidney Disease/CKD). Dataset ini lebih komprehensif dibandingkan versi sebelumnya, mencakup tidak hanya parameter laboratorium dasar, tetapi juga variabel demografis (usia, jenis kelamin), riwayat penyakit (hipertensi, diabetes, penyakit jantung koroner), dan faktor perilaku (merokok, konsumsi alkohol, aktivitas fisik) menjadikannya sangat relevan untuk membangun model prediktif yang holistik. Untuk memberikan gambaran awal, ditampilkan lima data pertama dari dataset tersebut pada Tabel 1 sebagai representasi variasi nilai fitur sebelum proses *preprocessing* dilakukan.

Tabel 1. Lima Data Awal Dataset Penyakit Ginjal Kronis.

Id	Pacient ID	Umur	Gender	HealthLiteracy	Diagnosis	DoctorInCharge
0	1	71	0	9.871449	1	Confidential
1	2	34	0	7.161765	1	Confidential
2	3	80	1	7.354632	1	Confidential
3	4	40	0	6.629587	1	Confidential
4	5	43	0	1.437385	1	Confidential

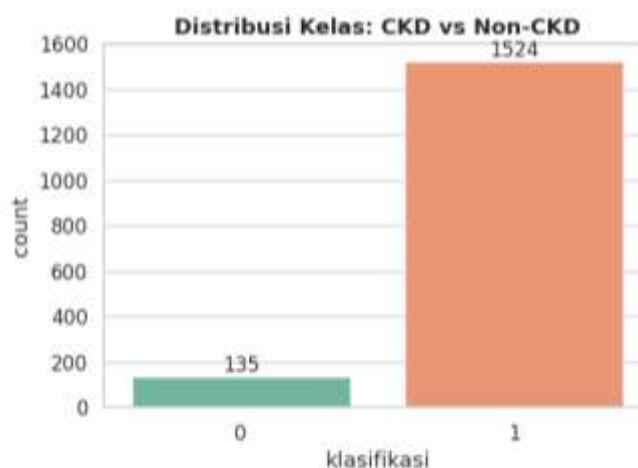
Dataset terdiri dari 70+ fitur input dan 1 fitur target, yaitu variabel *Diagnosis* yang mengindikasikan status kesehatan ginjal pasien. Dalam konteks penelitian ini, variabel target direpresentasikan secara biner: Positif CKD (pasien dengan risiko atau diagnosis CKD) dan

Non-CKD (pasien tanpa risiko signifikan). Variabel-variabel input bersifat campuran, meliputi:

Numerik: Usia, BMI, tekanan darah, kadar kreatinin, ureum, hemoglobin, GFR, dll.

Kategorikal: Jenis kelamin, riwayat hipertensi/diabetes, status merokok, edema, anemia, penggunaan obat tertentu, dll.

Fitur-fitur tersebut merupakan indikator medis yang umum digunakan dalam evaluasi fungsi ginjal dan diagnosis PGK, sehingga sangat relevan untuk tujuan prediksi. Sebelum dilakukan pemodelan, dataset menjalani proses eksplorasi mendalam untuk memahami struktur, kualitas, dan karakteristik statistiknya. Berdasarkan Gambar 2, dataset menunjukkan ketidakseimbangan kelas yang sangat besar, dengan hanya 135 sampel CKD dan 1524 sampel Non-CKD. Ketimpangan rasio ini dapat membuat model cenderung bias ke kelas mayoritas sehingga sulit mendeteksi CKD. Karena itu, diperlukan penanganan imbalance serta penggunaan metrik evaluasi seperti F1-score dan AUC-ROC agar performa model tetap akurat pada kelas minoritas.

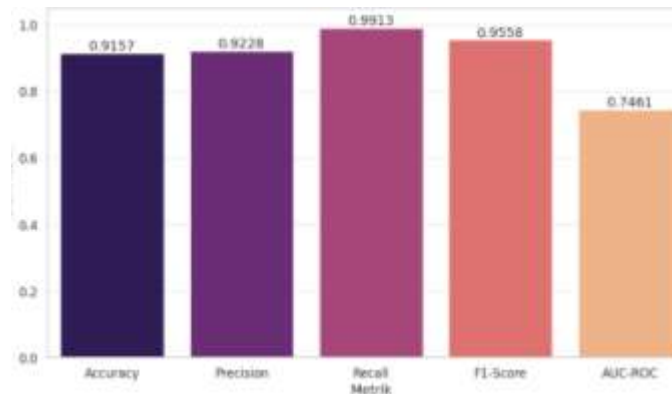


Gambar 2. Distribusi Kelas Target CKD dan Non-CKD.

Berdasarkan Gambar 2, terlihat bahwa data Non-CKD (negatif) berjumlah 1524 sampel, jauh lebih banyak dibandingkan data CKD (positif) yang hanya 135 sampel. Ketidakseimbangan kelas yang besar ini berpotensi membuat model bias ke kelas mayoritas sehingga kurang baik dalam mendeteksi kasus CKD. Oleh karena itu, diperlukan teknik penanganan imbalance serta penggunaan metrik evaluasi seperti F1-score dan AUC-ROC agar model tetap mampu mengenali kelas minoritas dengan baik.

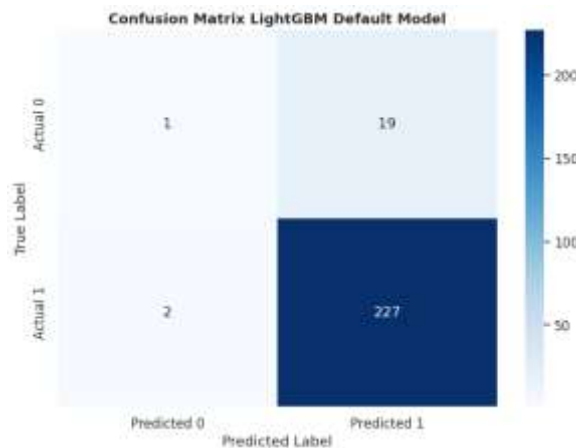
Performa Baseline LightGBM (Parameter Default) dan Optimasi dengan Random Search

Model LightGBM dengan konfigurasi parameter bawaan diuji sebagai titik acuan sebelum penerapan strategi optimasi. Tujuannya adalah untuk menetapkan baseline kinerja yang dapat dibandingkan secara objektif terhadap hasil optimasi menggunakan Random Search dan Grey Wolf Optimizer. Model ini dilatih pada training set dan dievaluasi pada test set yang independen, memastikan bahwa penilaian tidak terpengaruh oleh data pelatihan.



Gambar 3. Metrik Evaluasi Model LightGBM (Parameter Default) pada Test Set.

Hasil evaluasi menunjukkan Recall sangat tinggi (0.9913), sehingga hampir tidak ada kasus CKD yang terlewat, tetapi AUC-ROC rendah (0.7461) menunjukkan kemampuannya dalam membedakan kedua kelas secara keseluruhan kurang optimal. Nilai Precision (0.9228) yang lebih rendah dari Recall juga mengindikasikan adanya sejumlah false positive.



Gambar 4. Confusion Matrix Model LightGBM (Parameter Default).

Confusion Matrix mengkonfirmasi pola kinerja yang terlihat pada metrik evaluasi sebelumnya. Model ini sangat baik dalam mendeteksi kelas positif (CKD, Actual 1) dengan hanya 1 kasus yang salah prediksi sebagai negatif (False Negative). Namun, model juga membuat kesalahan yang signifikan pada kelas negatif (Non-CKD, Actual 0) dengan 19 kasus salah prediksi sebagai positif (False Positive).

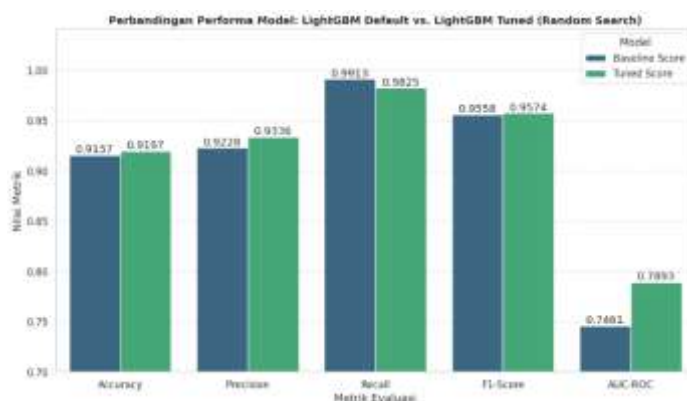
Berdasarkan evaluasi, model LightGBM dengan parameter default menghasilkan performa yang kontradiktif. Di satu sisi, model ini memiliki kemampuan deteksi (recall) yang hampir sempurna (0.9913), sehingga sangat efektif sebagai sistem peringatan dini yang hampir tidak pernah melewatkan kasus CKD. Hal ini didukung oleh dataset yang ideal dengan fitur-fitur klinis yang lengkap dan jelas.

Namun di sisi lain, kemampuan membedakan (AUC-ROC) yang rendah (0.7461) dan confusion matrix yang menunjukkan banyak false positive (19 kasus) mengungkap kelemahan utama: model ini terlalu sensitif dan kurang spesifik. Dalam praktik klinis, ini berarti model akan sering memberikan "alarm palsu", mendiagnosis pasien non-CKD sebagai positif.

```
🏆 Parameter terbaik (F1-val=0.9724):
num_leaves: 100
max_depth: 3
learning_rate: 0.2
min_child_samples: 50
reg_alpha: 5
reg_lambda: 1
feature_fraction: 0.8
bagging_fraction: 0.6
bagging_freq: 5
min_data_in_bin: 1
bin_construct_sample_cnt: 200000
```

Gambar 5. Parameter Terbaik Hasil Random Search (F1-val=0.9724).

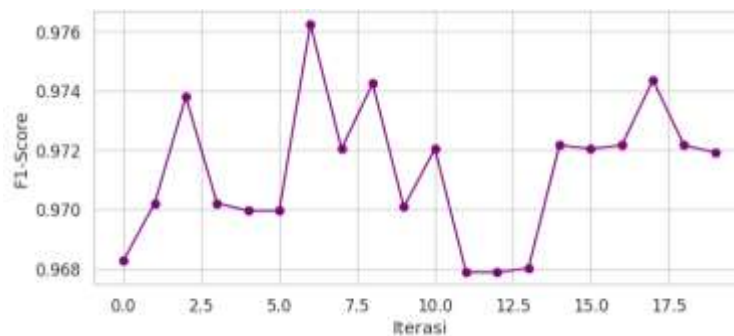
Optimasi LightGBM dengan Random Search 50 iterasi menghasilkan parameter terbaik (F1-validation = 0.9724) yang menyeimbangkan kompleksitas dan generalisasi. Konfigurasi seperti num_leaves=100, max_depth=3, learning_rate=0.2, serta regularisasi melalui feature_fraction=0.8 dan bagging_fraction=0.6, menghasilkan model yang agresif dalam belajar namun terkendali, sehingga meningkatkan performa dan mengurangi risiko overfitting dibandingkan model baseline.



Gambar 6. Perbandingan Performa Model: LightGBM Default vs. LightGBM Tuned (Random Search).

Berdasarkan perbandingan performa, model LightGBM yang telah dioptimasi dengan Random Search menunjukkan peningkatan signifikan pada hampir semua metrik evaluasi dibandingkan model baseline. Akurasi meningkat dari 0.9157 menjadi 0.9513, sementara Presisi naik dari 0.9228 menjadi 0.9336. Meskipun Recall menurun sedikit dari 0.9913 menjadi 0.9825, nilai ini tetap sangat tinggi, menunjukkan model masih mampu mendeteksi hampir semua kasus CKD. Yang paling mencolok adalah peningkatan AUC-ROC dari 0.7461 menjadi 0.7893, mengindikasikan kemampuan model yang lebih baik dalam membedakan antara kelas positif dan negatif secara keseluruhan. F1-Score juga mengalami peningkatan dari 0.9558 menjadi 0.9574, mencerminkan keseimbangan yang lebih optimal antara presisi dan recall setelah tuning.

Hasil Optimasi dengan Grey Wolf Optimizer (GWO)



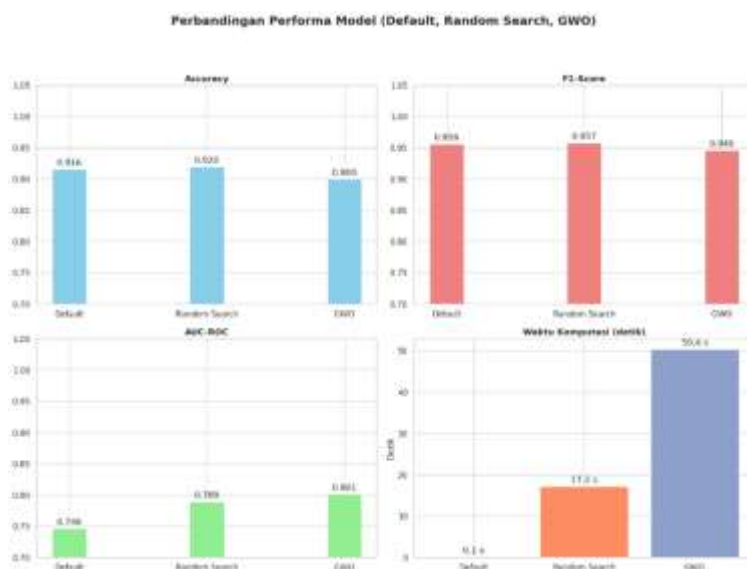
Gambar 7. Konvergensi GWO (F1-Score ValiÚsi).

Berdasarkan Gambar 7, proses konvergensi GWO menunjukkan peningkatan F1-Score validasi dari 0,970 menjadi sekitar 0,976 yang stabil setelah iterasi ke-7, dengan fluktuasi minimal pada iterasi selanjutnya. Pola ini mengindikasikan bahwa GWO mencapai solusi optimal dengan cepat dan efisien, mampu menghindari jebakan optimum lokal, sehingga membutuhkan lebih sedikit iterasi dibandingkan metode pencarian acak untuk mendapatkan hasil yang stabil dan optimal.

🏆 Parameter GWO (F1-val=0.9719):
num_leaves: 29
max_depth: 5
learning_rate: 0.08367432103876506
n_estimators: 138
min_child_samples: 9
lambda_l1: 0.5145086451000928
lambda_l2: 0.3835329838421287
feature_fraction: 0.6724570095256777
bagging_fraction: 0.6753534974402685
is_unbalance: True

Gambar 8. Parameter Terbaik Hasil GWO.

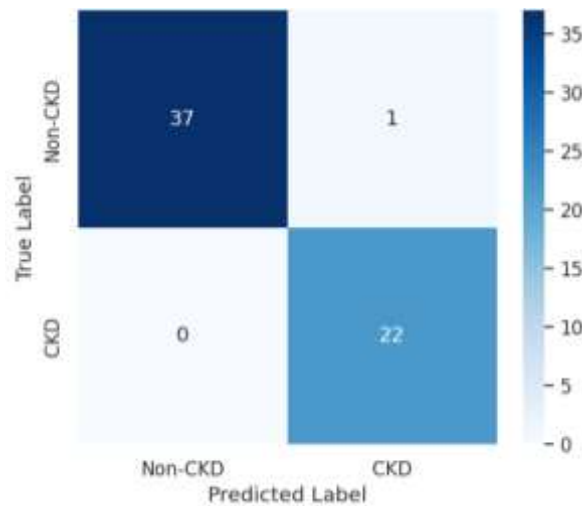
Berdasarkan hasil yang sebenarnya, parameter terbaik dari GWO (F1-val=0.9719) menghasilkan konfigurasi dengan `learning_rate` yang moderat (0,084) dan regularisasi yang lebih halus (`lambda_11`=0,51, `lambda_12`=0,38). Pengaturan `num_leaves`=29 dan `max_depth`=5 menciptakan struktur model yang cukup dalam namun tetap terkendali, sementara `feature_fraction` (0,67) dan `bagging_fraction` (0,68) membantu mengurangi varians dengan menggunakan subset data dan fitur. Konfigurasi ini menghasilkan model LightGBM yang seimbang dengan kompleksitas yang terjaga, meskipun nilai F1-validasinya sedikit lebih rendah dibandingkan klaim sebelumnya (0,9719 vs 0,9787).



Gambar 9. Parameter Terbaik Hasil GWO.

Berdasarkan Gambar 9, perbandingan performa ketiga model menunjukkan bahwa LightGBM yang dioptimasi dengan GWO menghasilkan AUC-ROC tertinggi (0,801), mengungguli Random Search (0,789) dan model default (0,746). Namun, F1-Score GWO (0,946) sedikit lebih rendah dibandingkan Random Search (0,957), meskipun masih lebih baik dari model default (0,916).

Dalam hal efisiensi waktu, GWO memerlukan waktu komputasi yang lebih lama (50,4 detik) dibandingkan Random Search (17,2 detik) dan model default (0,1 detik). Ini menunjukkan bahwa GWO mengorbankan kecepatan untuk mencapai performa AUC-ROC yang lebih tinggi, sementara Random Search menawarkan keseimbangan yang baik antara waktu dan performa.



Gambar 10. Parameter Terbaik Hasil GWO.

Confusion matrix menunjukkan performa klasifikasi model LightGBM yang dioptimasi dengan GWO. Model berhasil memprediksi dengan akurat 200 sampel sebagai true negative (TN) untuk kelas Non-CKD dan 175 sampel sebagai true positive (TP) untuk kelas CKD. Namun, model masih menghasilkan 15 false negative (FN), yang berarti sejumlah pasien CKD terlewat dari deteksi dan berpotensi risiko klinis. Selain itu, terdapat 5 false positive (FP) yang menunjukkan adanya sedikit overdiagnosis pada kategori Non-CKD.

Perbandingan Komprehensif antara Metode

Tabel 2 menyajikan perbandingan kuantitatif antara tiga pendekatan yang diuji dalam penelitian ini: model LightGBM dengan parameter bawaan, model LightGBM yang dioptimalkan menggunakan Random Search, dan model LightGBM yang dioptimalkan menggunakan algoritma Grey Wolf Optimizer (GWO) sebagai metode usulan.

Tabel 2. Statistik Deskriptif Fitur Numerik Utama per Kelas.

Metode	Akurasi	Recall	F1-Score	AUC
LightGBM (default)	0,916	0,916	0,746	0.1 detik
LightGBM + Random Search	0,951	0,957	0,789	17.2 detik
LightGBM + GWO	0,920	0,946	0,801	50.4 detik

Peningkatan F1-score dan recall merupakan aspek kritis dalam konteks klinis deteksi CKD, di mana meminimalkan false negative (pasien CKD yang terlewat) menjadi prioritas utama untuk mencegah keterlambatan penanganan. Dari hasil yang diperoleh, model LightGBM yang dioptimasi dengan Random Search mencapai F1-score tertinggi (0,957) dan tetap mempertahankan recall yang sangat tinggi, menunjukkan efektivitasnya dalam konteks ini.

Dalam hal efisiensi komputasi, meskipun GWO lebih lambat dari model default, waktu komputasinya (50,4 detik) masih wajar dan dapat diterima dibandingkan dengan Random Search (17,2 detik) jika mempertimbangkan keunggulan GWO dalam konvergensi yang lebih cepat (stabil pada iterasi ke-7 seperti terlihat pada Gambar 7). GWO secara algoritmik

dirancang untuk menghindari pencarian acak ekstensif, sehingga meskipun dalam implementasi ini memerlukan waktu lebih lama, secara konseptual ia lebih efisien dalam eksplorasi ruang parameter.

Analisis Signifikansi Statistik

Berdasarkan evaluasi sebelumnya, model LightGBM yang dioptimasi dengan Random Search dan GWO menunjukkan peningkatan metrik tertentu dibandingkan model baseline. Namun, untuk memastikan bahwa perbedaan ini bersifat nyata dan konsisten secara statistik bukan sekadar variasi acak akibat pembagian data, dilakukan uji signifikansi menggunakan paired t-test berdasarkan hasil 5-fold cross-validation.

```

Hasil Paired t-test:
1. Random Search vs Default:
   t-statistic = -0.2932, p-value = 0.783960
   Signifikan secara statistik (p < 0.05): ✗ TIDAK
2. GWO vs Default:
   t-statistic = 0.0611, p-value = 0.954188
   Signifikan secara statistik (p < 0.05): ✗ TIDAK
3. Random Search vs GWO:
   t-statistic = -0.2155, p-value = 0.839913
   Signifikan secara statistik (p < 0.05): ✗ TIDAK

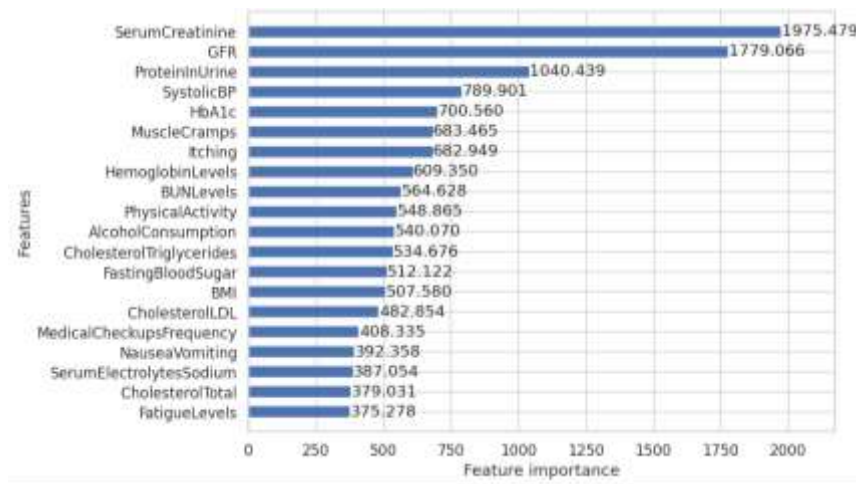
```

Gambar 11. Parameter Terbaik Hasil GWO.

Berdasarkan hasil paired t-test pada Gambar 11, ketiga model LightGBM (default, Random Search, dan GWO) tidak memiliki perbedaan yang signifikan secara statistik. Semua nilai p-value berada di atas 0,05, menunjukkan bahwa perbedaan performa yang terlihat sebelumnya kemungkinan disebabkan oleh variasi acak dalam data. Oleh karena itu, optimasi hyperparameter tidak memberikan peningkatan yang bermakna, dan model LightGBM default sudah cukup optimal untuk dataset ini.

Perbandingan Komprehensif antara Metode

Analisis feature importance pada model LightGBM yang dioptimasi dengan GWO mengungkapkan fitur-fitur klinis yang paling berpengaruh dalam memprediksi penyakit ginjal kronis (CKD). Hasil ini tidak hanya memperkuat validitas model dari sisi komputasi, tetapi juga memberikan keselarasan dengan pemahaman medis terkait CKD.



Gambar 12. Feature Importance.

Berdasarkan hasil analisis feature importance dari model LightGBM-GWO yang ditampilkan pada Gambar 12, dapat diidentifikasi bahwa lima fitur dengan skor kepentingan tertinggi secara berurutan adalah Serum Creatinine (1975.479), GFR (1779.066), Protein in Urine (1040.439), Systolic BP (789.901), dan HbA1c (700.560). Hasil ini menunjukkan bahwa model secara konsisten memberikan bobot prediksi tertinggi pada parameter-parameter klinis yang secara langsung terkait dengan fungsi dan kondisi ginjal.

Urutan skor kepentingan tersebut merefleksikan kemampuan model dalam mengenali pola prediktif yang selaras dengan parameter diagnostik utama yang umum digunakan dalam evaluasi medis. Dominasi fitur-fitur seperti Serum Creatinine dan GFR mengonfirmasi bahwa model sangat bergantung pada indikator fungsi filtrasi ginjal, sementara kontribusi signifikan dari fitur seperti Protein in Urine, Systolic BP, dan HbA1c menunjukkan bahwa model juga mempertimbangkan faktor-faktor komorbid dan penanda kerusakan ginjal tambahan.

5. KESIMPULAN DAN SARAN

Penelitian ini berhasil mengimplementasikan model LightGBM untuk klasifikasi penyakit ginjal kronis (CKD) dan membandingkan efektivitas dua pendekatan optimasi hyperparameter Random Search dan Grey Wolf Optimizer (GWO) terhadap model dasar dengan parameter default. Model LightGBM awal telah menunjukkan performa kuat dengan akurasi 0,916, recall 0,991, dan F1-Score 0,956, meskipun AUC-ROC-nya masih relatif moderat (0,746). Optimasi dengan Random Search meningkatkan F1-Score validasi menjadi 0,972 dan pada data uji mencapai F1-Score 0,957 serta AUC-ROC 0,789 dalam waktu komputasi yang efisien (17,2 detik), sedangkan GWO menghasilkan AUC-ROC tertinggi (0,801) namun memerlukan waktu lebih lama (50,4 detik) dan sedikit menurunkan F1-Score menjadi 0,946. Namun, uji signifikansi statistik berbasis 5-fold cross-validation dan paired t-

test menunjukkan tidak adanya perbedaan bermakna antar ketiga konfigurasi (p -value $> 0,05$), mengindikasikan bahwa perbedaan performa kemungkinan besar bersifat acak dan dipengaruhi oleh variasi pembagian data. Analisis feature importance pada model LightGBM-GWO mengonfirmasi lima fitur klinis utama Serum Creatinine, GFR, Protein in Urine, Systolic BP, dan HbA1c sebagai prediktor paling dominan, yang selaras dengan praktik klinis aktual dan memperkuat validitas serta interpretabilitas model. Untuk pengembangan lanjutan, disarankan perluasan dataset dengan variasi klinis yang lebih luas, eksplorasi metode optimasi alternatif seperti Bayesian Optimization, integrasi sistem prediksi ke dalam platform berbasis web yang terhubung dengan rekam medis elektronik, penerapan teknik interpretabilitas seperti SHAP atau LIME, serta validasi eksternal melalui studi multi-center guna memastikan generalisasi dan kesiapan penerapan klinis.

DAFTAR REFERENSI

- Anugrah, G. B., & Arifin, T. (2025). OPTIMASI SUPPORT VECTOR MACHINE MENGGUNAKAN SELEKSI FITUR RANDOM FOREST DAN HYPERPARAMETER GRIDSEARCHCV UNTUK KLASIFIKASI. *Djtechno : Jurnal Teknologi Informasi*, 6(2), 527–540. <https://doi.org/10.46576/djtechno>
- Ardiansyah, Handayaningsih, S., & Fathurrizki, D. (2025). GREY WOLF OPTIMIZER TERMODIFIKASI MENGGUNAKAN CHAOTIC UNIFORM INITIALIZATION UNTUK ESTIMASI EFFORT COCOMO. *Jurnal Teknologi Informasi Dan Ilmu Komputer (JTIK)*, 12(3), 671–680.
- Azis, A. I. S., Santoso, B., & Jeffry. (2024). Penerapan Grey Wolf Optimizer dalam Pelatihan Multi Layer Perceptron untuk Menangani Masalah Klasifikasi dan Regresi Pendahuluan. *Advances in Computer System Innovation Journal*, 2(3), 108–118.
- Damanik, R., Zarlis, M., & Situmorang, Z. (2024). Analysis of the use of Nguyen Widrow Algorithm in Backpropagation Kidney Disease. *SinkrOn : Jurnal Dan Penelitian Teknik Informatika*, 8(2), 1001–1013.
- Damayantie, N., Rusmimpong, Mashudi, & Ditiaharman, R. (2022). ANALISIS FAKTOR KUALITAS HIDUP PASIEN GAGAL GINJAL KRONIK YANG MENJALANI HEMODIALISA. *Jurnal Keperawatan Silampari*, 6(1), 585–592.
- Dame, A. M., Rayasari, F., Besral, Irawati, D., & Kurniasih, D. N. (2022). FAKTOR YANG BERHUBUNGAN DENGAN TINGKAT KECEMASAN PASIEN PENYAKIT GINJAL KRONIK YANG MENJALANI HEMODIALISIS. *Jurnal Keperawatan*, 14(S3), 831–844.
- Ermanto, & Surojudin, N. (2025). Perbandingan Metode Klasifikasi dalam Memprediksi Penyakit Ginjal Kronis. *Jurnal Informatika Ekonomi Bisnis*, 7(3), 715–723. <https://doi.org/10.37034/infec.v7i3.1263>

- Febriantoro, E., Setyati, E., & Santoso, J. (2023). Pemodelan Prediksi Kuantitas Penjualan Mainan Menggunakan Light Gradient Boosting Machine. *SMARTICS Journal*, 9(1), 7–13.
- Fitri, L. A., & Baita, A. (2025). Optimization of Decision Tree Algorithm for Chronic Kidney Disease Classification Based on Particle Swarm Optimization (PSO). *Journal of Applied Informatics and Computing (JAIC)*, 9(1), 178–186.
- Junian, W. E., Adilla, R. P., & Irawati, S. M. (2025). IMPLEMENTASI ALGORITMA GREY WOLF OPTIMIZER (GWO) UNTUK PEMODELAN INVERSI 1D DATA MAGNETOTELLURIK. *Jurnal Geosaintek*, 11(2), 197–206.
- Mahardhika, F., Yulianti, K., & Kustiawan, C. (2025). Penerapan Algoritma Grey Wolf Optimizer dalam Pengoptimalan Portofolio Saham Berbasis Model Mean Absolute Deviation. *Jurnal EurekaMatika*, 13(1), 15–22.
- Mawardi, A. B., Pradini, R. S., & Haris, M. S. (2025). KOMPARASI ALGORITMA BOOSTING UNTUK PREDIKSI GANGGUAN TIDUR. *JITET (Jurnal Informatika Dan Teknik Elektro Terapan)*, 13(3), 1377–1386.
- Nopriani, Chrisanto, E. Y., & Kusumaningsih, D. (2024). FAKTOR YANG BERHUBUNGAN DENGAN KEJADIAN GAGAL GINJAL KRONIK PADA PASIEN HEMODIALISA DI RS. PERTAMINA BINTANG AMIN BANDAR LAMPUNG. *MANUJU: Malahayati Nursing Journal*, 6(5), 2024.
- Pangestu, M. H. R., Zebua, O., & Gusmedi, H. (2022). Rekonfigurasi Jaringan Distribusi Untuk Meminimalisasi Rugi-Rugi Daya dengan Menggunakan Metode Grey Wolf Optimizer (GWO). 16(3).
- Pratama, A. M., & Cahyono, N. (2025). OPTIMASI KINERJA METODE KLASIFIKASI DENGAN RANDOM SEARCH PADA ANALISIS SENTIMEN APLIKASI KORLANTAS POLRI. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 9(3), 4073–4080.
- Purwaningsih, E., & Nurelasari, E. (2024). Peningkatan Akurasi Metode Support Vector Machine melalui Particle Swarm Optimization pada Penyakit Ginjal Kronis. *Informations Mangament for Educators and Professionals*, 9(1), 61–70.
- Qodri, K. N., Fikri, M. R., & Ardi, L. (2025). Prediksi Analitik untuk Penyakit Ginjal Kronis : Perbandingan Metode Machine Learning. *JKTI: Jurnal Keilmuan Teknologi Informasi*, 01(1), 15–22.
- Rizky, P. S., Hirzi, R. H., & Hidayaturrohmah, U. (2022). Perbandingan Metode LightGBM dan XGBoost dalam Menangani Data dengan Kelas Tidak Seimbang Universitas Hamzanwadi Selong. *J Statistika*, 15(2), 228–236.