

OPTIMASI *HYPERPARAMETER* LIGHTGBM MENGGUNAKAN *PARTICLE SWARM OPTIMIZATION* UNTUK KLASIFIKASI KANKER PAYUDARA

*¹Akbar Umar Alfaroq, ¹Yisti Vita Via, ¹Chrystia Aji Putra

¹Informatika, Fakultas Ilmu Komputer, Universitas Pembangunan Nasional “Veteran” Jawa Timur

DOI: <https://doi.org/10.33005/jifosi.v7i2.621>

*Corresponding Author: 19081010043@student.upnjatim.ac.id

Submit: Juli 2026 | Accepted: Agustus 2026 | Published: Agustus 2026

ABSTRAK

Performa model *Light Gradient Boosting Machine* (LightGBM) pada proses klasifikasi sangat dipengaruhi oleh konfigurasi hiperparameter yang digunakan. Penelitian ini mengusulkan penerapan *Particle Swarm Optimization* (PSO) untuk memperoleh kombinasi hiperparameter yang optimal pada model LightGBM dalam klasifikasi kanker payudara. Pengujian model dilakukan menggunakan skema *3-fold cross-validation*. Hasil eksperimen menunjukkan bahwa model LightGBM dengan hiperparameter *default* menghasilkan akurasi rata-rata sebesar 95,96%, sedangkan optimasi menggunakan PSO meningkatkan akurasi hingga 97,54% pada konfigurasi *swarm size* 60 dan *max iteration* 250. Temuan tersebut menunjukkan bahwa PSO efektif dalam meningkatkan performa klasifikasi model LightGBM.

Kata Kunci: *Particle Swarm Optimization*, *Hiperparameter Tuning*, *Breast Cancer Classification*, *LightGBM*

HYPERPARAMETER OPTIMIZATION OF LIGHTGBM USING PARTICLE SWARM OPTIMIZATION FOR BREAST CANCER CLASSIFICATION

ABSTRACT

The performance of the Light Gradient Boosting Machine (LightGBM) model in classification tasks is highly influenced by the hyperparameter configuration used during training. This study proposes the use of Particle Swarm Optimization (PSO) to obtain an optimal set of hyperparameters for the LightGBM model in breast cancer classification. The model was evaluated using a 3-fold cross-validation scheme. The experimental results show that the LightGBM model with default hyperparameters achieved an average accuracy of 95.96%, while PSO-based hyperparameter optimization improved the accuracy to 97.54% using a swarm size of 60 and max iterations of 250. These findings indicate that PSO is effective in improving the classification performance of the LightGBM model.

Keywords: Particle Swarm Optimization, Hyperparameter Tuning, Breast Cancer Classification, LightGBM.

PENDAHULUAN

Kanker menjadi penyebab utama kematian akibat penyakit tidak menular di Indonesia. Kanker payudara memiliki tingkat kematian yang tinggi, terutama pada perempuan [1]. Deteksi dini menjadi faktor penting dalam meningkatkan peluang kesembuhan pasien karena memungkinkan penanganan dilakukan sebelum kanker berkembang ke stadium yang lebih lanjut [2]. Oleh karena itu, diperlukan metode diagnosis yang cepat, akurat, dan mampu membantu proses identifikasi kanker payudara secara efektif.

Penerapan *machine learning* dalam bidang kesehatan terus berkembang, salah satunya untuk mendukung diagnosis kanker payudara melalui proses klasifikasi. Teknologi ini mampu mengolah data hasil

pemeriksaan medis secara lebih cepat dibandingkan analisis manual sehingga dapat membantu pengambilan keputusan [3]. Beberapa penelitian juga telah menerapkan algoritma seperti LightGBM untuk menyelesaikan permasalahan klasifikasi pada data medis dan menunjukkan performa yang baik. Meskipun demikian, Performa model dalam melakukan klasifikasi bergantung pada pengaturan hiperparameter yang diterapkan selama proses pembelajaran, sehingga pengaturan hiperparameter menjadi faktor penting dalam memperoleh hasil klasifikasi yang optimal.

Kemampuan LightGBM dalam menghasilkan prediksi yang akurat bergantung pada konfigurasi hiperparameter yang diterapkan saat proses pelatihan. Hiperparameter seperti *learning rate*, *num_leaves*, *max_depth*, dan *n_estimators* perlu dioptimalkan agar model mampu menghasilkan performa klasifikasi yang optimal [4]. Konfigurasi hiperparameter yang tidak sesuai dapat menyebabkan model kesulitan mempertahankan performanya ketika dihadapkan pada data yang belum pernah dipelajari sebelumnya, sehingga proses optimasi hiperparameter menjadi tahapan yang penting dalam pengembangan model LightGBM [5].

Optimasi hiperparameter pada model *machine learning* dapat dilakukan menggunakan *Particle Swarm Optimization* (PSO), model ini memanfaatkan sekumpulan partikel dalam suatu populasi untuk mencari kombinasi hiperparameter terbaik pada model machine learning. Selain digunakan untuk optimasi hiperparameter. [6]. Mengacu pada permasalahan tersebut, penelitian ini berfokus pada penerapan metode *Particle Swarm Optimization* (PSO) untuk mengoptimalkan hiperparameter model LightGBM pada klasifikasi kanker payudara. Proses evaluasi dilakukan untuk membandingkan performa model sebelum dan sesudah optimasi. Melalui pendekatan ini, diharapkan diperoleh konfigurasi hiperparameter yang lebih optimal sehingga akurasi model LightGBM dapat meningkat dibandingkan penggunaan hiperparameter *default*.

METODOLOGI

Data

Data yang digunakan dalam penelitian ini berasal dari *Breast Cancer Wisconsin (Diagnostic) Dataset* yang disediakan oleh *UCI Machine Learning Repository*. Dataset tersebut terdiri atas 569 data pasien dalam bentuk data tabular dengan 30 fitur yang merepresentasikan karakteristik inti sel (*cell nuclei*), yang meliputi 212 data malignant dan 357 data benign. Dataset ini dikembangkan berdasarkan hasil pemeriksaan klinis menggunakan metode *Fine Needle Aspiration* (FNA), kemudian sampel sel dianalisis untuk menghasilkan fitur-fitur numerik dan diberi label sesuai hasil diagnosis medis. Sebelum proses pemodelan, dilakukan analisis terhadap karakteristik dataset, meliputi struktur data, distribusi fitur, dan distribusi kelas untuk memastikan data siap digunakan dalam proses klasifikasi.

Pembagian Data

Evaluasi model pada penelitian ini menerapkan skema *3-fold cross validation*, yaitu dengan membagi seluruh data menjadi tiga kelompok yang berukuran hampir sama. Setiap kelompok secara bergantian berperan sebagai data pengujian, sementara dua kelompok lainnya dimanfaatkan untuk melatih model hingga seluruh *fold* memperoleh giliran sebagai data uji. Proses ini diulang sebanyak tiga kali hingga setiap *fold* memperoleh kesempatan menjadi data pengujian satu kali. Nilai akurasi dari setiap iterasi kemudian dirata-ratakan untuk memperoleh hasil evaluasi yang lebih andal dalam mengukur performa model klasifikasi.

Model LightGBM

LightGBM dikembangkan sebagai penyempurnaan dari *Gradient Boosting Decision Tree* (GBDT) untuk meningkatkan efisiensi pemodelan pada data terstruktur. Proses pembentukan pohonnya menggunakan pendekatan *leaf-wise*, sehingga model dapat dilatih lebih cepat dengan tetap memberikan hasil klasifikasi yang baik. [7].

Pada penelitian ini, model LightGBM menggunakan pengaturan hiperparameter *default* yang umum digunakan sebagai konfigurasi awal dalam proses pelatihan. Nilai hiperparameter tersebut meliputi *learning_rate* sebesar 0,1, *num_leaves* sebesar 31, *max_depth* bernilai -1, dan *n_estimators* sebesar 100.

Penggunaan hiperparameter *default* ini dijadikan sebagai dasar pembandingan untuk mengevaluasi pengaruh proses optimasi terhadap kinerja model dalam klasifikasi kanker payudara.

Optimasi PSO

Kennedy dan Eberhart (1995) memperkenalkan PSO yang merupakan algoritma optimasi yang memanfaatkan pergerakan sekumpulan partikel dalam suatu populasi untuk memperoleh solusi terbaik melalui proses iteratif. [8]. Karakteristik tersebut menjadikan PSO sebagai salah satu metode optimasi yang banyak dimanfaatkan untuk menentukan kombinasi hiperparameter terbaik pada berbagai model *machine learning*.

Setiap partikel merepresentasikan solusi potensial yang bergerak dalam ruang pencarian berdasarkan *personal best* (pbest) dan *global best* (gbest). Perbaruan kecepatan partikel didefinisikan sebagai:

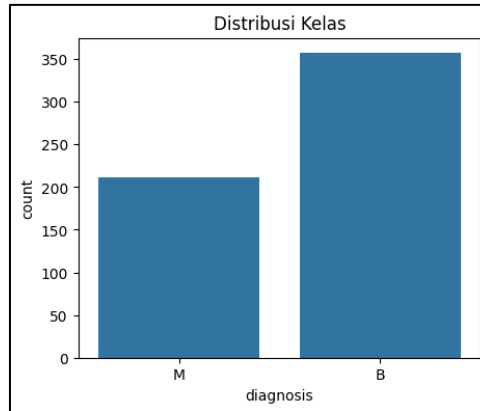
$$v(t+1) = \omega \cdot v(t) + c1 \cdot r1 \cdot (pbest - x(t)) + c2 \cdot r2 \cdot (gbest - x(t))$$

Inertia weight (ω) untuk menyeimbangkan eksplorasi global dan eksploitasi lokal.

Nilai ω yang tinggi mendorong eksplorasi, sedangkan nilai yang rendah mendorong eksploitasi. Skenario Pengujian performa model dilakukan menggunakan nilai akurasi yang diperoleh melalui skema *3-fold cross-validation*. Evaluasi mencakup model LightGBM dengan hiperparameter *default* sebagai pembandingan serta model LightGBM yang telah dioptimalkan menggunakan *Particle Swarm Optimization* (PSO). Proses optimasi dilakukan pada tiga konfigurasi PSO, yaitu *swarm size* 10 dengan *maximum iteration* 50, *swarm size* 30 dengan *maximum iteration* 150, dan *swarm size* 60 dengan *maximum iteration* 250. Seluruh hasil pengujian kemudian dianalisis untuk membandingkan performa masing-masing model dalam klasifikasi kanker payudara.

HASIL DAN PEMBAHASAN

Gambar 1 menunjukkan distribusi data target (diagnosis) pada dataset *Breast Cancer Wisconsin (Diagnostic)* yang digunakan dalam penelitian. Kelas Malignant (M) terdiri atas 212 sampel, sedangkan kelas Benign (B) terdiri atas 357 sampel, sehingga total data yang digunakan sebanyak 569 sampel. Distribusi tersebut memperlihatkan bahwa jumlah sampel pada kelas benign lebih banyak dibandingkan kelas malignan



Gambar 1. Distribusi Kelas pada data kanker Payudara

Tabel 1. Sample data

<i>Mean Radius</i>	<i>Mean Texture</i>	<i>Mean Perimter</i>	<i>Mean Area</i>	<i>Mean Smoothness</i>	...	<i>Diagnosis</i>
14.68	20.13	94.74	684.5	0.09867	...	M
16.13	20.68	108.1	798.8	0.117	...	M
19.81	22.15	130	1260	0.09831	...	M
13.54	14.36	87.46	566.3	0.09779	...	B
13.08	15.71	85.63	520	0.1075	...	B
...	...	...	...	...	...	...

Sumber : Data diolah

Tabel 1 merupakan contoh sampel data yang ditampilkan sebanyak lima baris pertama. Tabel tersebut memperlihatkan nilai dari setiap fitur beserta kelas target (diagnosis) yang digunakan sebagai data masukan dalam proses klasifikasi kanker payudara.

Hasil Pengujian Baseline

Bagian ini menyajikan hasil pengujian model LightGBM menggunakan hiperparameter default sebagai *baseline*. Pengujian dilakukan menggunakan metode *3-fold cross validation*, kemudian nilai akurasi pada setiap *fold* dirata-ratakan untuk memperoleh akurasi akhir sebagai acuan dalam mengevaluasi pengaruh optimasi hiperparameter menggunakan *Particle Swarm Optimization* (PSO).

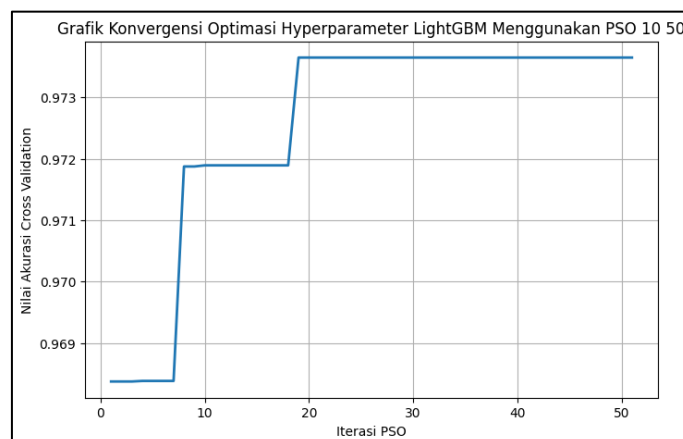
Tabel 2. Hasil Evaluasi *Baseline* LightGBM

Fold	Akurasi%
Fold-1	97,37
Fold-2	93,68
Fold-3	96,83
Rata-Rata	95,96

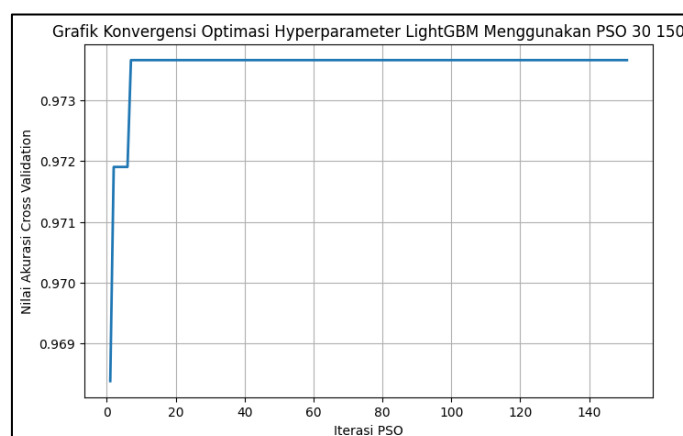
Sumber: Data diolah

Tabel 2 menyajikan hasil evaluasi awal terhadap model LightGBM sebelum dilakukan optimasi hiperparameter. Berdasarkan pengujian menggunakan *3-fold cross-validation*, diperoleh rata-rata akurasi sebesar 95,96% yang dijadikan sebagai *baseline* pada tahap evaluasi berikutnya.

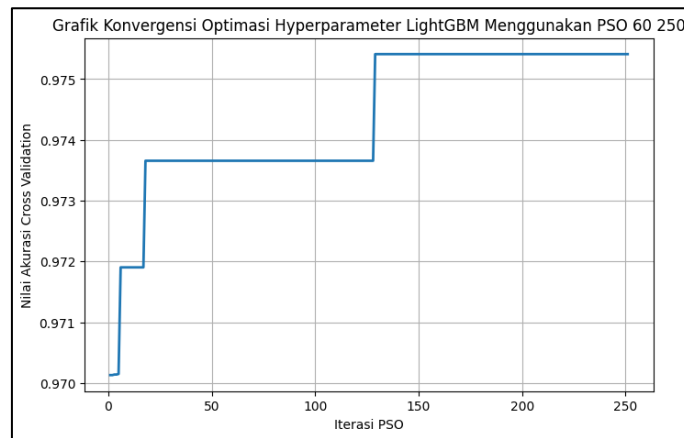
Grafik Konvergensi PSO



Gambar 2. Grafik Konvergensi PSO 10-50



Gambar 2. Grafik Konvergensi PSO 30 150



Gambar 4. Grafik Konvergensi PSO 60-250

Berdasarkan grafik konvergensi, seluruh skenario mencapai kondisi konvergen sebelum iterasi maksimum, dengan hasil terbaik pada skenario *swarm size* 60 dan *maximum iteration* 250.

Hasil Pengujian LightGBM dengan PSO

Bagian ini menyajikan hasil pengujian model LightGBM yang dioptimalkan menggunakan PSO dengan tiga kombinasi parameter, yaitu *swarm size* 10 dan *max iteration* 50, *swarm size* 30 dan *max iteration* 150, serta *swarm size* 60 dan *max iteration* 250.

Tabel 3. Hasil Evaluasi LightGBM dengan PSO 10 50

Fold	Akurasi%
Fold-1	99,47
Fold-2	94,74
Fold-3	97,88
Rata-Rata	97,36

Sumber: Data diolah

Tabel 3 menunjukkan hasil evaluasi model LightGBM dengan PSO (*swarm size* 10 dan *max iteration* 50) yang memperoleh rata-rata akurasi sebesar 97,36%..

Tabel 4. Hasil Evaluasi LightGBM dengan PSO 30 150

Fold	Akurasi%
Fold-1	98,95
Fold-2	94,74
Fold-3	98,41
Rata-Rata	97,37

Sumber: Data diolah

Tabel 4 menunjukkan hasil evaluasi model LightGBM dengan PSO (*swarm size* 30 dan *max iteration* 150) yang memperoleh rata-rata akurasi sebesar 97,37%.

Tabel 5. Hasil Evaluasi LightGBM dengan PSO 60 250

Fold	Akurasi%
Fold-1	99,47
Fold-2	94,74
Fold-3	98,41
Rata-Rata	97,54

Sumber: Data diolah

Tabel 5 menunjukkan hasil evaluasi model LightGBM dengan PSO (*swarm size* 60 dan *max iteration* 250) yang memperoleh rata-rata akurasi sebesar 97,54%.

Perbandingan dan Analisis

Tabel 6. Hasil Perbandingan Evaluasi setiap Skenario

Skenario	Akurasi%
Baseline LightGBM	95,96
LightGBM + PSO (10 50)	97,36
LightGBM + PSO (30 150)	97,37
LightGBM + PSO (60 250)	97,54

Sumber: Data diolah

Peningkatan akurasi dari 95,96% pada model *baseline* menjadi 97,54% setelah optimasi menunjukkan bahwa PSO mampu menemukan kombinasi hiperparameter yang lebih baik dibandingkan konfigurasi *default*. Dari tiga skenario yang diuji, konfigurasi *swarm size* 60 dan *maximum iteration* 250 menghasilkan performa terbaik, sedangkan dua skenario lainnya juga memberikan peningkatan akurasi yang konsisten di atas model *baseline*. Hasil ini sejalan dengan temuan Kazerani [6] yang menunjukkan bahwa penerapan PSO mampu meningkatkan performa klasifikasi kanker payudara. Meskipun digunakan pada proses seleksi fitur, temuan tersebut mendukung efektivitas PSO dalam meningkatkan kinerja model klasifikasi.

KESIMPULAN

Berdasarkan hasil penelitian, penerapan PSO mampu meningkatkan akurasi model LightGBM dalam klasifikasi kanker payudara. Model *baseline* memperoleh rata-rata akurasi sebesar 95,96%, sedangkan model yang dioptimalkan menggunakan PSO menghasilkan akurasi tertinggi sebesar 97,54%. Dari tiga skenario optimasi yang diuji, skenario dengan *swarm size* 60 dan *max iteration* 250 memberikan performa terbaik. Peningkatan akurasi yang diperoleh mengindikasikan bahwa optimasi hiperparameter menggunakan PSO memberikan kontribusi positif terhadap kinerja model LightGBM pada klasifikasi kanker payudara.

DAFTAR RUJUKAN

- [1] WHO, "Preventing and controlling cancer: WHO and partners step up support," WHO(World Health Organization). Accessed: Jul. 02, 2026. [Online]. Available: <https://www.who.int/indonesia/news/detail/24-10-2024-preventing-and-controlling-cancer--who-and-partners-step-up-support>
- [2] S. Arooj, M. Zubair, and M. F. Khan, "Breast Cancer Detection and Classification Empowered With Transfer Learning," *Front. Public Heal.*, vol. 10, no. July, pp. 1–18, 2022, doi: 10.3389/fpubh.2022.924432.
- [3] N. R. Muntari and K. H. Hanif, "Klasifikasi Penyakit Kanker Payudara Menggunakan Perbandingan Algoritma Machine Learning," *J. ILMU Komput. DAN Teknol.*, vol. 3, no. 1, pp. 1–6, 2022.
- [4] E. Ramadanti, D. A. Dinathi, C. Sri, K. Aditya, and D. R. Chandranegara, "Diabetes Disease Detection Classification Using Light Gradient Boosting (LightGBM) With Hyperparameter Tuning," *Sink. J. dan Penelit. Tek. Inform.*, vol. 8, no. 2, pp. 956–963, 2024.
- [5] M. Soni, M. W. Bhatt, and P. Ofori-amanfo, "Light-GBM based minority oversampling model using biomedical data analysis for breast cancer classification," *Discov. Appl. Sci.*, 2025.
- [6] R. Kazerani, "Improving Breast Cancer Diagnosis Accuracy by Particle Swarm Optimization Feature Selection Wisconsin Diagnostic Breast Cancer dataset," *Int. J. Comput. Intell. Syst.*, vol. 5, 2024, doi: 10.1007/s44196-024-00428-5.
- [7] S. Y. Nailendra, W. Witanti, and G. Abdillah, "Optimasi Prediksi Penjualan Retail Online Menggunakan LightGBM dan Hyperparameter Tuning," *J. Algoritma.*, vol. 22, no. 2, pp. 1931–1942, 2025, doi: 10.33364/algoritma/v.22-2.2551.
- [8] A. G. Gad, *Particle Swarm Optimization Algorithm and Its Applications : A Systematic Review*, vol. 29, no. 5. Springer Netherlands, 2022. doi: 10.1007/s11831-021-09694-4.