

Model Klasifikasi Diabetes Menggunakan XGBoost dengan Optimasi Seleksi Fitur dan Hyperparameter Berbasis PSO

Sheila Putri Aprilianti[#], Andrian Sah[#], Siti Nurhayati[#], Rasna[#], Jusmawati[#]

[#] Program Studi Sistem Informasi, Ilmu Komputer, Universitas Yapis Papua, Jayapura, 99111, Indonesia
E-mail: [sheilaputriaprilianti\[at\]gmail.com](mailto:sheilaputriaprilianti[at]gmail.com), [cyberdefance23\[at\]gmail.com](mailto:cyberdefance23[at]gmail.com), [nurhayati.siti21\[at\]gmail.com](mailto:nurhayati.siti21[at]gmail.com),
[rasna\[at\]uniyap.ac.id](mailto:rasna[at]uniyap.ac.id), [jusmawati.nr\[at\]gmail.com](mailto:jusmawati.nr[at]gmail.com)

ABSTRACTS

The rising global burden of diabetes mellitus has increased the need for accurate, technology-based early detection systems. This study develops a diabetes classification model using Extreme Gradient Boosting (XGBoost) optimized through a two-stage Particle Swarm Optimization (PSO) scheme: Binary PSO (BPSO) for feature selection and Global Best PSO (GBPSO) for hyperparameter tuning. Data were obtained from the Kaggle Diabetes Prediction Dataset (100,000 records; eight clinical attributes: gender, age, hypertension, heart disease, smoking history, BMI, HbA1c level, and blood glucose level). The extreme class imbalance (91.5% normal vs 8.5% diabetes) was addressed using the SMOTETomek hybrid technique. BPSO retained all eight features as the optimal combination (best cost 0.0329; F1-weighted 96.71%), while GBPSO produced the best hyperparameter configuration ($n_estimators = 416$, $learning_rate = 0.237$, $max_depth = 3$, $min_child_weight = 3$; best cost 0.0308, converging at the 11th iteration). The final model achieved 97.15% test-set accuracy, a ROC-AUC of 0.9779, and a diabetes-class precision of 0.93. The model was deployed as a Streamlit-based web system classifying patients into three risk categories: Not Indicated, Early Risk Indicated, and Diabetes Indicated. A preliminary functionality trial of the prototype on five real patient records from an anonymized partner hospital in Jayapura City showed classification results fully consistent with patients' clinical status (5 of 5 correct), indicating the prototype's technical feasibility rather than clinical validation, which would require larger-scale testing. These findings demonstrate that integrating XGBoost with a two-stage PSO optimization scheme produces an accurate and clinically applicable diabetes classification model.

Manuscript received Mei 8, 2026;
revised Jun 12 2026. accepted Jun
29, 2026 Date of publication Jun
30, 2026. International Journal,
JITSI : Jurnal Ilmiah Teknologi
Sistem Informasi licensed under a
Creative Commons Attribution-
Share Alike 4.0 International
License



ABSTRAK

Peningkatan kasus diabetes mellitus secara global menuntut ketersediaan sistem deteksi dini yang lebih akurat dan berbasis teknologi. Penelitian ini mengembangkan model klasifikasi diabetes menggunakan Extreme Gradient Boosting (XGBoost) yang dioptimasi melalui skema dua tahap Particle Swarm Optimization (PSO): Binary PSO (BPSO) untuk seleksi fitur dan Global Best PSO (GBPSO) untuk penyetelan hyperparameter. Data diperoleh dari Diabetes Prediction Dataset Kaggle (100.000 rekam data, delapan atribut klinis: gender, usia, hipertensi, penyakit jantung, riwayat merokok, BMI, kadar HbA1c, dan kadar glukosa darah). Ketidakseimbangan kelas yang ekstrem (91,5% normal berbanding 8,5% diabetes) ditangani menggunakan teknik hibrida SMOTETomek. BPSO mempertahankan seluruh delapan fitur sebagai kombinasi optimal (best cost 0,0329; F1-weighted 96,71%), sedangkan GBPSO menghasilkan konfigurasi hyperparameter terbaik ($n_estimators=416$, $learning_rate=0,237$, $max_depth=3$, $min_child_weight=3$; best cost 0,0308, konvergen pada iterasi ke-11). Model akhir mencapai akurasi 97,15% pada data uji, ROC-AUC 0,9779, dan precision kelas diabetes 0,93. Model

diimplementasikan ke dalam sistem berbasis web menggunakan Streamlit yang mengklasifikasikan pasien ke dalam tiga kategori risiko: Tidak Terindikasi, Terindikasi Risiko Awal, dan Terindikasi Diabetes. Uji coba fungsionalitas prototipe pada lima sampel data rekam medis pasien dari rumah sakit mitra yang dirahasiakan namanya di Kota Jayapura menunjukkan hasil klasifikasi yang seluruhnya konsisten dengan kondisi klinis pasien (5 dari 5 sesuai), mengindikasikan kelayakan teknis prototipe, bukan validasi klinis, yang memerlukan pengujian berskala lebih besar. Penelitian ini menunjukkan bahwa integrasi XGBoost dengan skema optimasi dua tahap PSO mampu menghasilkan model klasifikasi diabetes yang akurat dan aplikatif.

Keywords / Kata Kunci — *Extreme Gradient Boosting; Particle Swarm Optimization; Klasifikasi Diabetes Mellitus; Seleksi Fitur; Hyperparameter Tuning.*

CORRESPONDING AUTHOR

Sheila Putri Aprilianti
 Program Studi Sistem Informasi, Ilmu Komputer, Universitas Yapis Papua, Jayapura, 99111, Indonesia
 Email: sheilaputriaprilianti[at]gmail.com

1. PENDAHULUAN

Peningkatan jumlah kasus diabetes secara global menunjukkan kebutuhan yang semakin mendesak terhadap pendekatan yang lebih efektif dalam mendukung deteksi dini berbasis data. International Diabetes Federation (2025) melaporkan sebanyak 589 juta orang dewasa hidup dengan diabetes dan diproyeksikan meningkat menjadi 853 juta pada tahun 2050, dengan sekitar 3,4 juta kematian pada tahun 2024 [1]. World Health Organization (2024) turut mencatat peningkatan jumlah penderita dari 200 juta pada tahun 1990 menjadi 830 juta pada tahun 2022 [2]. Pada tingkat lokal, Dinas Kesehatan Kota Jayapura (2023, 2025) mencatat 6.539 kasus diabetes mellitus pada tahun 2022, namun kunjungan pasien pada 14 puskesmas justru menurun dari 4.885 (2022) menjadi 2.482 (2025) sebuah tren kontradiktif yang mengindikasikan rendahnya kesadaran masyarakat untuk pemeriksaan rutin, bukan penurunan jumlah penderita sesungguhnya [3]. Kondisi ini menegaskan perlunya pendekatan berbasis teknologi untuk membantu identifikasi status diabetes secara lebih dini, khususnya pada data klinis di lingkungan pelayanan kesehatan seperti Rumah Sakit Mitra. Berbagai penelitian telah menerapkan machine learning untuk klasifikasi diabetes, namun masih menyisakan keterbatasan. Abdurrahman et al. [4] mengoptimasi XGBoost menggunakan Grid Search dan Random Search pada dataset UCI, meningkatkan akurasi dari 75% menjadi 95%, tetapi tanpa seleksi fitur cerdas berbasis swarm intelligence. Rahman et al. [5] mengevaluasi lima metode seleksi fitur (RFE, GWO, PSO, GA, Boruta) pada LightGBM dan XGBoost, dengan performa terbaik justru diraih Boruta-LightGBM (akurasi 85,16%), namun PSO hanya digunakan untuk seleksi fitur tanpa hyperparameter tuning simultan. Maulana dan Eliyani [6] serta Maulana et al. [7] menunjukkan efektivitas PSO dalam menyetel hyperparameter Naïve Bayes, C4.5, Random Forest, dan SVM, tetapi tidak melibatkan XGBoost dalam eksperimennya. Rahmawati et al. [8] berhasil menggabungkan SVM dengan SMOTE pada data RSUD Indonesia (akurasi 95,8%; AUC 99,1), namun tanpa optimasi fitur maupun parameter berbasis metaheuristik. Pada konteks yang lebih dekat, Kurnia et al. [9] membuktikan kombinasi PSO-XGBoost meningkatkan AUC dari 0,9366 menjadi 0,9483 pada klasifikasi Parkinson, sedangkan Jawza et al. [10] menerapkan PSO sebagai seleksi fitur pada XGBoost untuk dataset diabetes Kaggle 100.000 records (AUC meningkat dari 0,7666 menjadi 0,8250), tetapi tanpa hyperparameter tuning berbasis PSO, tanpa perbandingan eksplisit sebelum-sesudah optimasi, dan tanpa implementasi pada data klinis nyata. Kaliappan et al. [11] menganalisis berbagai strategi klasifikasi dan seleksi fitur untuk prediksi diabetes pada beragam dataset, sedangkan Talebi Moghaddam et al. [12] mengidentifikasi fitur-fitur penting pada data tidak seimbang melalui studi kohort lima tahun menggunakan machine learning; kedua studi ini menegaskan pentingnya seleksi fitur dan penanganan ketidakseimbangan kelas dalam klasifikasi diabetes, namun tanpa mengintegrasikan optimasi metaheuristik seperti PSO. Berdasarkan kajian tersebut, penerapan PSO dan XGBoost pada klasifikasi diabetes umumnya masih dilakukan secara terpisah atau tidak lengkap dua fungsi PSO (seleksi fitur dan hyperparameter tuning) belum diterapkan secara terintegrasi, penanganan ketidakseimbangan kelas berskala besar melalui teknik hibrida belum banyak dieksplorasi bersamaan dengan optimasi tersebut, dan validasi model pada data klinis nyata beserta implementasi sistem berbasis web masih jarang dilaporkan. Penelitian ini mengisi celah tersebut dengan mengusulkan model klasifikasi diabetes berbasis XGBoost [13] yang dioptimasi melalui dua tahap Particle Swarm Optimization (PSO) [14], yaitu Binary PSO untuk seleksi fitur dan Global Best PSO untuk hyperparameter tuning disertai penanganan ketidakseimbangan kelas menggunakan SMOTETomek. Model yang dihasilkan selanjutnya diimplementasikan ke dalam sistem berbasis web menggunakan Streamlit dan divalidasi menggunakan data klinis nyata dari Rumah Sakit Mitra, sehingga kontribusi penelitian ini mencakup baik aspek metodologis (integrasi dua fungsi PSO pada XGBoost) maupun aspek aplikatif (sistem skrining dini yang siap digunakan pada layanan kesehatan).

2. METODOLOGI PENELITIAN

2.1. Data Penelitian

Penelitian ini menggunakan Diabetes Prediction Dataset dari platform Kaggle sebagai data utama untuk pelatihan, optimasi, dan evaluasi model, terdiri atas 100.000 rekaman data dengan delapan atribut klinis prediktor (gender, age, hypertension, heart_disease, smoking_history, bmi, HbA1c_level, blood_glucose_level) dan satu atribut target (diabetes: 0 = Normal, 1 = Diabetes). Data rekam medis pasien dari rumah sakit mitra penelitian di Kota Jayapura (nama institusi dirahasiakan atas pertimbangan etika penelitian, selanjutnya disebut Rumah Sakit Mitra) digunakan secara terpisah pada tahap validasi implementasi sistem, dengan struktur atribut yang disesuaikan agar identik dengan dataset Kaggle.

2.2. Tahapan Penelitian

Penelitian dilaksanakan mengikuti alur kerja berkelanjutan dalam satu lingkungan eksekusi (Jupyter Notebook): (1) pemuatan, pembersihan, dan encoding data; (2) standarisasi fitur; (3) evaluasi baseline XGBoost; (4) seleksi fitur menggunakan Binary PSO; (5) hyperparameter tuning menggunakan Global Best PSO; (6) pembagian data latih-uji 80:20; (7) penyeimbangan kelas pada data latih menggunakan SMOTETomek; (8) pelatihan dan evaluasi model final; serta (9) implementasi model ke dalam aplikasi web Streamlit.

2.3. Preprocessing Data

Pembersihan data dilakukan menggunakan dropna() dan penghapusan spasi tersembunyi pada kolom kategorikal. Encoding diterapkan pada atribut gender (Male=0, Female=1, Other=2) dan smoking_history (category encoding, menghasilkan kode 0–5). Seluruh fitur hasil encoding kemudian distandarisasi menggunakan StandardScaler mengikuti transformasi Z-score pada Persamaan (1), diterapkan pada keseluruhan data (100.000 baris) sebelum data dibagi menjadi data latih dan data uji, karena proses optimasi PSO memerlukan matriks fitur yang telah berskala seragam pada seluruh data. Setelah proses seleksi fitur dan hyperparameter tuning selesai, dataset dibagi menjadi data latih (80%) dan data uji (20%) menggunakan train-test split dengan stratify=y dan random_state=42, sehingga proporsi kelas Normal dan Diabetes tetap terjaga pada kedua kelompok data.

$$z = (x - \mu) / \sigma \quad (1)$$

dengan x adalah nilai asli, μ adalah rata-rata fitur, dan σ adalah standar deviasi fitur.

2.4. Penanganan Ketidakseimbangan Kelas

Ketidakseimbangan kelas yang ekstrem pada dataset (91,5% Normal berbanding 8,5% Diabetes) ditangani menggunakan teknik hibrida SMOTETomek dari library Imbalanced-learn (random_state=42). SMOTE membangkitkan sampel sintesis pada kelas minoritas, sedangkan Tomek Links menghapus pasangan data antar kelas yang saling berdekatan untuk mempertegas batas keputusan. Teknik ini diterapkan pada data latih hasil pembagian 80:20, sedangkan data uji tetap pada distribusi aslinya agar evaluasi mencerminkan kondisi nyata.

2.5. Seleksi Fitur Berbasis Binary PSO

Seleksi fitur dilakukan menggunakan kelas BinaryPSO dari library PySwarms [15], dengan setiap partikel merepresentasikan vektor biner 8-dimensi (1 = fitur disertakan, 0 = fitur diabaikan). Konfigurasi parameter BPSO disajikan pada Tabel 1. Nilai fitness setiap kombinasi fitur dihitung sebagai 1 dikurangi rata-rata F1-weighted hasil Stratified 5-Fold Cross Validation pada pipeline SMOTETomek + XGBoost (parameter default), sehingga kombinasi fitur dengan fitness terendah merupakan kombinasi terbaik.

$$Fitness = 1 - F1 - weighted \quad (2)$$

Nilai fitness dihitung sebagai 1 dikurangi rata-rata F1-weighted hasil Stratified 5-Fold Cross Validation pada pipeline SMOTETomek + XGBoost, mengikuti konvensi minimisasi yang digunakan library PySwarms; kombinasi fitur/hyperparameter dengan nilai fitness terendah merupakan solusi terbaik. Pembaruan kecepatan dan posisi setiap partikel pada BPSO mengikuti persamaan berikut:

$$v_{i,j}^{t+1} = w \cdot v_{i,j}^t + c_1 \cdot r_1 \cdot (Pbest_{i,j}^t - x_{i,j}^t) + c_2 \cdot r_2 \cdot (Gbest_{g,j}^t - x_{i,j}^t) \quad (3)$$

$$x_{i,j}^{t+1} = x_{i,j}^t + v_{i,j}^{t+1} \quad (4)$$

dengan v_{id}^{t+1} dan x_{id}^{t+1} berturut-turut adalah kecepatan dan posisi partikel i pada dimensi d untuk iterasi ke $t + 1$; w adalah bobot inersia; c_1, c_2 adalah koefisien akselerasi kognitif dan sosial; $r_1, r_2 \in [0,1]$ adalah bilangan acak; $Pbest_{id}$ adalah posisi terbaik individu partikel; dan $Gbest_d$ adalah posisi terbaik seluruh kawanan (Tabel 1 untuk BPSO, Tabel 2 untuk GBPSO). Pada tahap BPSO mekanisme ini diterapkan pada representasi posisi biner, sedangkan pada tahap GBPSO diterapkan pada representasi posisi kontinu yang berkorespondensi langsung dengan keempat hyperparameter XGBoost. Fungsi transfer sigmoid $S(\cdot)$ digunakan pada tahap BPSO untuk mengonversi nilai kecepatan kontinu menjadi keputusan biner: jika $r < S(v_{id}^{t+1})$, maka $x_{id}^{t+1} = 1$ (fitur disertakan); $r \geq S(v_{id}^{t+1})$, maka $x_{id}^{t+1} = 0$ (fitur tidak disertakan), dengan r adalah bilangan acak pada rentang $[0,1]$.

$$S(v_{id}^{t+1}) = \frac{1}{1 + e^{-v_{id}^{t+1}}} \quad (5)$$

TABEL 1. Konfigurasi Parameter Binary PSO

Parameter	Nilai	Keterangan
n_particles	20	Jumlah partikel
dimensions	8	Jumlah fitur input
iters	20	Jumlah iterasi
c1 (kognitif)	0,5	Bobot personal best
c2 (social)	0,3	Bobot global best
w (inersia)	0,9	Kontrol kecepatan gerak
k	3	Jumlah tetangga
p	2	Normal jarak (Minkowski)

TABEL 2. Konfigurasi dan Ruang Pencarian Global Best PSO

Parameter	Batas Bawah	Batas Atas
n_estimators	50	500
learning_rate	0,01	0,30
max_depth	2	10
min_child_weight	1	10

2.6. Hyperparameter Tuning Berbasis Global Best PSO

Setelah subset fitur optimal ditetapkan, hyperparameter XGBoost dioptimasi menggunakan kelas GlobalBestPSO dari PySwarms, dengan setiap partikel merepresentasikan empat nilai kontinu (n_estimators, learning_rate, max_depth, min_child_weight). Konfigurasi parameter dan ruang pencarian (search space) disajikan pada Tabel 2. Fungsi fitness dihitung dengan mekanisme yang identik dengan tahap BPSO, yaitu 1 dikurangi rata-rata F1-weighted hasil Stratified 5-Fold Cross Validation pada pipeline SMOTETomek + XGBoost, dengan posisi partikel di-clip pada setiap iterasi agar tetap berada dalam batas pencarian, dan nilai kontinu dikonversi menggunakan int() (truncation, bukan pembulatan) untuk parameter n_estimators, max_depth, dan min_child_weight.

Berbeda dengan BPSO yang bekerja pada representasi biner, pembaruan kecepatan dan posisi pada GBPSO menggunakan Persamaan (3) dan (4) yang sama, namun diterapkan pada representasi posisi kontinu yang berkorespondensi langsung dengan keempat hyperparameter XGBoost (tanpa fungsi transfer sigmoid pada Persamaan (5), karena ruang pencarian GBPSO bersifat kontinu, bukan biner). Nilai kontinu hasil optimasi selanjutnya dikonversi menggunakan fungsi int() Python (truncation) untuk parameter n_estimators, max_depth, dan min_child_weight. Fungsi fitness yang diminimumkan identik dengan Persamaan (2). Konfigurasi optimizer: n_particles=15, dimensions=4, iters=20, dengan $c_1=0,5$, $c_2=0,3$, dan $w=0,9$ identik dengan tahap BPSO.

2.7. Pelatihan dan Evaluasi Model Final

Model final XGBClassifier dikonfigurasi menggunakan hyperparameter hasil GBPSO, dilatih pada data latih yang telah diseimbangkan menggunakan SMOTETomek (X_train_res, y_train_res), lalu dievaluasi pada data uji (20%) yang tidak mengalami resampling. Metrik evaluasi meliputi accuracy, precision, recall, F1-score per kelas, ROC-AUC, dan confusion matrix, dihitung menggunakan classification_report() dan confusion_matrix() dari Scikit-learn.

Berdasarkan nilai pada confusion matrix (True Positive/TP, True Negative/TN, False Positive/FP, False Negative/FN), dihitung empat metrik evaluasi berikut: Akurasi (Persamaan (6)), Presisi (Persamaan (7)), Recall (Persamaan (8)), dan F1-Score (Persamaan (9)).

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

$$Presisi = \frac{TP}{TP+FP} \quad (7)$$

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

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

2.8. Implementasi Sistem Berbasis Web

Model final beserta scaler dan artefak fitur diserialisasi menggunakan joblib ke dalam empat berkas .pkl, kemudian dimuat ke dalam aplikasi web berbasis Streamlit (app.py). Sistem menerima delapan atribut klinis sebagai input, melakukan encoding dan scaling yang identik dengan tahap pelatihan, kemudian menghasilkan nilai probabilitas melalui predict_proba(). Probabilitas tersebut dipetakan ke tiga kategori risiko menggunakan threshold: $< 3\%$ = Tidak Terindikasi, $3\% \leq 10\%$ = Terindikasi Risiko Awal, dan $\geq 10\%$ = Terindikasi Diabetes

3. HASIL DAN PEMBAHASAN

3.1. Karakteristik Dataset dan Preprocessing

Dataset yang digunakan terisi penuh tanpa nilai kosong pada seluruh 100.000 baris dan 9 kolom. Distribusi kelas target menunjukkan ketidakseimbangan yang signifikan, sebagaimana disajikan pada Tabel 3 dan Gambar 1.

TABEL 3. Distribusi Kelas Target Sebelum SMOTETomek

Kelas	Label	Jumlah Data	Persentase
0	Normal	91.500	91,5%
1	Diabetes	8.500	8,5%

TABEL 4. Perbandingan Distribusi Data Sebelum dan Sesudah SMOTETomek

Kondisi Data	Kelas 0 (Normal)	Kelas 1 (Diabetes)	Total
Sebelum SMOTETomek	91.500	8.500	100.000
Sesudah SMOTETomek	90.905	90.905	181.810

Rasio ketidakseimbangan sekitar 10:1 ini berpotensi menyebabkan model bias terhadap kelas mayoritas, sehingga diperlukan penanganan khusus. Setelah penerapan SMOTETomek pada data latih, jumlah sampel kelas Diabetes meningkat signifikan sementara kelas Normal berkurang sedikit akibat penghapusan sampel noise oleh Tomek Links, sebagaimana dirangkum pada Tabel 4.

Tomek Links berhasil mengidentifikasi dan menghapus 595 sampel yang berada terlalu dekat dengan batas antar kelas, sehingga total data latih hasil penyeimbangan menjadi 181.810 sampel yang digunakan sebagai basis evaluasi pada proses optimasi PSO.

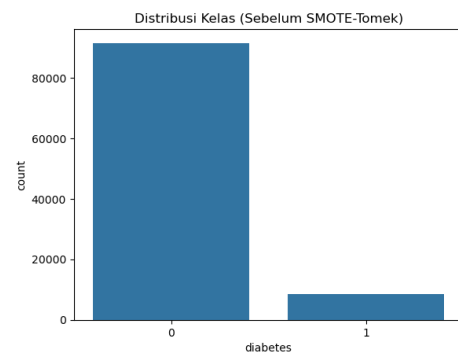
3.2. Hasil Seleksi Fitur Binary PSO

Setelah 20 iterasi dengan 20 partikel, BPSO mencapai konvergensi dengan hasil sebagaimana disajikan pada Tabel 5.

TABEL 5. Hasil Optimasi Binary PSO

Indikator	Nilai
Best Cost	0,0329
F1-Weighted Terbaik	0,9671 (96,71%)
Best Position	[1,1,1,1,1,1,1,1]
Jumlah Fitur Terpilih	8 dari 8

```
diabetes
0    91500
1     8500
Name: count, dtype: int64
```



GAMBAR 1. Distribusi Kelas Target Sebelum Resampling

TABEL 6. Hasil Optimasi Hyperparameter XGBoost

Parameter	Nilai Raw (Kontinu)	Nilai Implementasi
n_estimators	416,53	416
learning_rate	0,2370	0,237
max_depth	3,80	3
min_child_weight	3,32	3

Posisi terbaik yang ditemukan BPSO adalah [1,1,1,1,1,1,1,1], yang berarti seluruh delapan fitur klinis (gender, age, hypertension, heart_disease, smoking_history, bmi, HbA1c_level, blood_glucose_level) dipertahankan sebagai kombinasi optimal. Hasil ini mengindikasikan bahwa setiap atribut memberikan kontribusi informatif terhadap performa klasifikasi, tanpa ada eliminasi terhadap satupun atribut. Temuan ini selaras dengan karakteristik diabetes sebagai penyakit multifaktorial yang dipengaruhi kombinasi faktor demografis, antropometrik, klinis, komorbiditas, dan gaya hidup, sehingga seluruh delapan fitur digunakan sebagai input pada tahap hyperparameter tuning.

3.3. Hasil Optimasi Hyperparameter Global Best PSO

Nilai best cost pada optimasi GBPSO mengalami penurunan bertahap dari 0,0316 pada iterasi pertama hingga konvergen pada nilai 0,0308 pada iterasi ke-11 dan tetap stabil hingga iterasi ke-20. Hyperparameter terbaik yang diperoleh disajikan pada Tabel 6.

Nilai max_depth diperoleh dari konversi $\text{int}(3,80) = 3$ menggunakan fungsi `int()` Python yang memotong desimal (truncation), bukan membulatkan ke nilai terdekat, konsisten dengan parameter yang tersimpan pada berkas model final.

3.4. Perbandingan Performa Baseline vs PSO-XGBoost

Perbandingan dilakukan antara model XGBoost baseline (parameter default, tanpa seleksi fitur, SMOTETomek, maupun hyperparameter tuning eksternal) dengan model PSO-XGBoost yang mengintegrasikan BPSO, SMOTETomek, dan GBPSO, keduanya dievaluasi menggunakan Stratified 5-Fold Cross Validation pada data latih. Hasil perbandingan disajikan pada Tabel 7.

Nilai F1-weighted model baseline (0,9693) dan PSO-XGBoost (0,9692) menunjukkan hasil yang hampir identik. Hal ini dapat dijelaskan oleh dua faktor: pertama, XGBoost merupakan algoritma yang secara bawaan sudah cukup robust pada data tabular berskala besar; kedua, perbandingan ini tidak sepenuhnya setara karena model PSO-XGBoost menyelesaikan permasalahan yang lebih kompleks (penyeimbangan kelas sekaligus

penentuan fitur dan parameter optimal) namun tetap menghasilkan F1-weighted yang setara. Kontribusi PSO yang lebih signifikan terletak pada kemampuan model menangani kelas minoritas secara terstruktur dan dapat direproduksi, sebagaimana terlihat pada evaluasi data uji berikut.

3.5. Evaluasi Model Pada Data Uji

Model final PSO-XGBoost dievaluasi pada 20.000 data uji (18.300 Normal dan 1.700 Diabetes) yang tidak mengalami resampling. Confusion matrix hasil evaluasi disajikan pada Tabel 8 dan Gambar 2.

TABEL 7. Perbandingan Performa Model Baseline vs PSO-XGBoost

Aspek	Baseline (XGBoost)	PSO-XGBoost
Seleksi Fitur	Tidak Ada	BPSO (8 fitur terpilih)
Penanganan Imbalanc	Tidak Ada	SMOTETomek
Hyperparameter Tuning	Default	Global Best PSO
F1-Weighted (CV)	0,9693	0,9692
Std.Deviasi (CV)	0,0006	-

TABEL 8. Confusion Matrix Model PSO-XGBoost pada Data Uji

	Klasifikasi: Normal (0)	Klasifikasi: Diabetes (1)
Aktual : Normal (0)	TN = 18,215	FP = 85
Aktual : Diabetes (1)	FN = 485	TP = 1.215

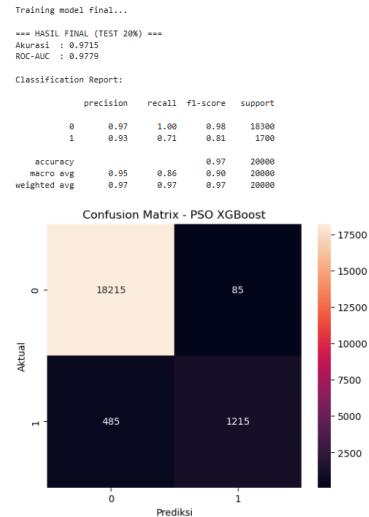
Model berhasil mengklasifikasikan 18.215 pasien Normal dan 1.215 pasien Diabetes secara benar, dengan 85 kesalahan false positive dan 485 kesalahan false negative. Rekapitulasi metrik evaluasi model final disajikan pada Tabel 9. Model mencapai akurasi 97,15% dengan ROC-AUC 0,9779, menunjukkan kemampuan diskriminasi yang sangat baik antara kelas Normal dan Diabetes. Precision kelas Diabetes sebesar 0,93 menandakan bahwa 93% pasien yang diklasifikasikan sebagai Diabetes memang benar-benar positif, sedangkan recall 0,71 berarti model berhasil mendeteksi 71% dari seluruh pasien Diabetes pada data uji hasil yang cukup baik mengingat ketidakseimbangan kelas ekstrem pada dataset asli (8,5% data Diabetes sebelum SMOTETomek). F1-score kelas Diabetes sebesar 0,81 mencerminkan keseimbangan yang baik antara precision dan recall, menjadi indikator keberhasilan penanganan data imbalanced menggunakan SMOTETomek.

TABEL 9. Rekapitulasi Metrik Evaluasi Model Final (PSO-XGBoost)

Metrik	Kelas 0 (Normal)	Kelas 1 (Diabetes)	Rata-rata (Weighted)
Precision	0,97	0,93	0,97
Recall	1,00	0,71	0,97
F1 Score	0,98	0,81	0,97
Akurasi Keseluruhan	-	-	0,9715 (97,15%)
ROC-AUC Score	-	-	0,9779

3.6. Uji Coba Fungsionalitas Prototipe pada Data Klinis Rumah Sakit Mitra

Sebagai uji coba fungsionalitas awal terhadap prototipe sistem (bukan validasi klinis formal), dilakukan pengujian menggunakan lima sampel data rekam medis pasien Rumah Sakit Mitra yang dimasukkan secara manual ke dalam antarmuka sistem. Hasil klasifikasi sistem dibandingkan dengan label aktual pasien pada rekam medis. Seluruh lima sampel (5 dari 5, 100%) menghasilkan kategori status klasifikasi yang sesuai dengan kondisi klinis aktual pasien, mengindikasikan bahwa pipeline sistem mulai dari input antarmuka hingga output klasifikasi berfungsi sebagaimana dirancang pada data klinis nyata yang tidak pernah dilihat model selama pelatihan. Mengingat jumlah sampel yang masih sangat terbatas, hasil ini merupakan indikasi awal kelayakan teknis (technical feasibility) prototipe dan belum dapat diklaim sebagai validasi klinis; pengujian pada skala yang jauh lebih besar (idealnya puluhan hingga ratusan rekam medis) diperlukan sebelum sistem dapat dipertimbangkan untuk validasi klinis formal.



GAMBAR 2. Confusion Matrix Model Final PSO-XGBoost

TABEL 10. Hasil Pengujian Sensitivitas Atribut

Kode	Atribut Diubah	Probabilitas	Selisih (point persentase)
Baseline	-	9,5%	-
S1	Gender -> Male	13,5%	+4,0
S2	Usia -> 65	17,2%	+7,7
S3	Hipertensi -> Ya	15,4%	+5,9
S4	Penyakit Jantung -> Ya	24,6%	+15,1
S5	Merokok -> Current	7,9%	-1,6
S6	BMI -> 35	12,4%	+10,9

3.7. Analisis Sensitivitas Atribut

Sebagai pelengkap evaluasi kuantitatif, dilakukan analisis sensitivitas atribut (*one-attribute-at-a-time*) menggunakan satu data pasien baseline (gender female, usia 45 tahun, tidak hipertensi, tidak memiliki riwayat penyakit jantung, tidak pernah merokok, BMI 27,5, HbA1c 6,5, glukosa darah 140 mg/dL) untuk mengamati pengaruh perubahan satu atribut terhadap probabilitas klasifikasi sistem, dengan atribut lainnya tetap konstan. Hasil pengujian pada enam skenario atribut disajikan pada Tabel 10.

Hasil pengujian menunjukkan bahwa seluruh atribut yang diuji memberikan pengaruh terhadap probabilitas klasifikasi, dengan riwayat penyakit jantung (S4) menghasilkan perubahan probabilitas terbesar di antara skenario yang diuji, diikuti oleh usia (S2), BMI (S6), dan hipertensi (S3). Perubahan status merokok menjadi *current* (S5) merupakan satu-satunya skenario yang menurunkan probabilitas dibandingkan baseline. Tidak terdapat atribut dengan kontribusi nol, yang menunjukkan bahwa model PSO-XGBoost memanfaatkan seluruh atribut input dalam proses pengambilan keputusan, konsisten dengan mekanisme algoritma berbasis pohon dalam menangkap pola dan interaksi antarfitur. Pengujian sensitivitas pada penelitian ini secara khusus dibatasi pada enam atribut kontekstual dan gaya hidup (gender, usia, hipertensi, riwayat penyakit jantung, status merokok, dan BMI); HbA1c dan kadar glukosa darah tidak disertakan dalam skema pengujian satu-atribut ini karena keduanya merupakan penanda diagnostik langsung diabetes (nilai ambang HbA1c $\geq 6,5\%$ dan glukosa darah ≥ 126 mg/dL merupakan kriteria diagnosis baku), sehingga variasi nilainya akan secara trivial mengikuti ambang diagnostik tersebut dan kurang informatif untuk menilai sensitivitas model terhadap faktor risiko kontekstual.

3.8. Keterbatasan dan Arah Penelitian Masa Depan

Penelitian ini memiliki beberapa keterbatasan yang perlu diperhatikan pada penelitian selanjutnya. Pertama, uji coba fungsionalitas prototipe pada data klinis baru dilakukan pada lima sampel rekam medis pasien Rumah Sakit Mitra, sehingga hasil kesesuaian 100% belum dapat digeneralisasi sebagai validasi klinis dan memerlukan pengujian pada skala yang jauh lebih besar dengan jumlah sampel yang secara statistik lebih memadai untuk memperoleh estimasi kinerja klinis yang lebih kredibel. Kedua, perbandingan performa antara model baseline dan PSO-XGBoost pada Tabel 7 hanya melaporkan F1-weighted agregat, tanpa metrik *precision*, *recall*, dan F1-score per kelas untuk model baseline; perbandingan per-kelas yang lebih rinci diperlukan untuk membuktikan secara kuantitatif klaim bahwa kontribusi PSO terletak pada penanganan kelas minoritas. Ketiga, pengujian sensitivitas atribut pada penelitian ini dibatasi pada enam dari delapan atribut klinis, dengan HbA1c dan glukosa darah dikecualikan karena keduanya berfungsi sebagai penanda diagnostik langsung; penelitian selanjutnya dapat mengeksplorasi desain pengujian sensitivitas yang lebih sesuai untuk kedua atribut tersebut, misalnya melalui analisis interaksi dua atribut (*two-way interaction*) alih-alih skema satu-atribut. Keempat, penelitian ini menggunakan dataset publik (Kaggle Diabetes Prediction Dataset) yang bersifat tabular dan tidak sepenuhnya merepresentasikan variasi populasi klinis riil di Indonesia, sehingga generalisasi model ke populasi yang lebih luas memerlukan validasi lebih lanjut pada data multi-institusi. Sebagai arah penelitian masa depan, penelitian lanjutan disarankan untuk: (1) melakukan validasi klinis berskala lebih besar dengan protokol etik yang formal; (2) melaporkan metrik evaluasi per kelas secara lengkap untuk seluruh model pembandingan; (3) membandingkan performa PSO dengan metaheuristik lain seperti Genetic Algorithm (GA) atau Grey Wolf Optimizer (GWO) pada kasus yang sama; dan (4) menerapkan teknik interpretasi model (misalnya SHAP atau LIME) untuk memperkuat transparansi keputusan klasifikasi pada konteks klinis.

4. KESIMPULAN

Penelitian ini berhasil mengembangkan model klasifikasi diabetes mellitus berbasis XGBoost yang dioptimasi melalui dua tahap PSO Binary PSO untuk seleksi fitur dan Global Best PSO untuk hyperparameter tuning disertai penanganan ketidakseimbangan kelas menggunakan SMOTETomek. BPSO mempertahankan seluruh delapan fitur klinis sebagai kombinasi optimal (best cost 0,0329), sedangkan GBPSO menghasilkan konfigurasi hyperparameter $n_estimators=416$, $learning_rate=0,237$, $max_depth=3$, dan $min_child_weight=3$ (best cost 0,0308). Model final mencapai akurasi 97,15%, ROC-AUC 0,9779, *precision* kelas Diabetes 0,93, dan *recall* 0,71 pada data uji. Perbandingan dengan model baseline menunjukkan bahwa kontribusi utama PSO tidak terletak pada peningkatan F1-weighted secara agregat (0,9693 berbanding 0,9692), melainkan pada kemampuannya menghasilkan konfigurasi model yang terstruktur, konsisten, dan dapat direproduksi tanpa intervensi manual. Model yang dihasilkan telah diimplementasikan ke dalam sistem berbasis web menggunakan Streamlit dengan tiga kategori indikasi risiko, dan uji coba fungsionalitas prototipe pada lima sampel data klinis Rumah Sakit Mitra menunjukkan konsistensi klasifikasi 100% terhadap kondisi aktual pasien, sebagai indikasi awal kelayakan teknis sistem.

UCAPAN TERIMAKASIH

Penulis mengucapkan terima kasih kepada Rumah Sakit Mitra atas dukungan data klinis untuk keperluan uji coba fungsionalitas prototipe, serta kepada pembimbing penelitian atas arahan yang diberikan selama proses penyusunan penelitian ini.

REFERENSI

- [1] "International Diabetes Federation, IDF Diabetes Atlas." Accessed: Apr. 12, 2026. [Online]. Available: <https://diabetesatlas.org/>
- [2] World Health Organization, "Diabetes." Accessed: May 21, 2026. [Online]. Available: <https://www.who.int/news-room/fact-sheets/detail/diabetes>
- [3] Dinas Kesehatan Kota Jayapura, Profil Kesehatan Kota Jayapura Tahun 2022. Dinas Kesehatan Kota Jayapura, 2023.
- [4] G. Abdurrahman, H. Oktavianto, and M. Sintawati, "Optimasi Algoritma XGBoost Classifier Menggunakan Hyperparameter Gridsearch dan Random Search Pada Klasifikasi Penyakit Diabetes," *INFORMAL: Informatics Journal*, vol. 7, no. 3, pp. 193–198, 2022, doi: 10.19184/isj.v7i3.35441.
- [5] F. Rahman, S. Hossain, J. J. Tiang, and A. Al Nahid, "Diabetes Prediction Using Feature Selection Algorithms and Boosting-Based Machine Learning Classifiers," *Diagnostics*, vol. 15, no. 20, pp. 1–24, 2025, doi: 10.3390/diagnostics15202622.
- [6] R. Maulana and E. Eliyani, "Diabetes Classification Algorithm Optimization Using Particle Swarm Optimization on Naïve Bayes, C4.5 and Random Forest," *Jurnal Sisfokom (Sistem Informasi dan Komputer)*, vol. 14, no. 4, pp. 499–509, 2025, doi: 10.32736/sisfokom.v14i4.2431.
- [7] M. R. Maulana, A. Sucipto, and H. Mulyo, "OPTIMISASI PARAMETER SUPPORT VECTOR MACHINE DENGAN PARTICLE SWARM OPTIMIZATION UNTUK PENINGKATAN KLASIFIKASI DIABETES," *Jurnal Informatika Teknologi dan Sains (JINTEKS)*, vol. 6, no. 4, pp. 802–812, 2024.
- [8] S. Rahmawati, A. Wibowo, and A. F. N. Masruriyah, "Improving Diabetes Prediction Accuracy in Indonesia: A Comparative Analysis of SVM, Logistic Regression, and Naive Bayes with SMOTE and ADASYN," *Jurnal RESTI*, vol. 8, no. 5, pp. 607–614, 2024, doi: 10.29207/resti.v8i5.5980.
- [9] D. Kurnia, M. Itqan Mazdadi, D. Kartini, R. Adi Nugroho, and F. Abadi, "Seleksi Fitur dengan Particle Swarm Optimization pada Klasifikasi Penyakit Parkinson Menggunakan XGBoost," *Jurnal Teknologi Informasi dan Ilmu Komputer*, vol. 10, no. 5, pp. 1083–1094, 2023, doi: 10.25126/jtiik.2023107252.
- [10] D. N. Jawza, M. I. Mazdadi, A. Farmadi, T. H. Saragih, D. Kartini, and V. Abdullayev, "Enhancing Diabetes Prediction Accuracy Using Random Forest and XGBoost with PSO and GA-Based Feature Selection," *Journal of Electronics, Electromedical Engineering, and Medical Informatics*, vol. 7, no. 2, pp. 295–306, 2025, doi: 10.35882/jeeemi.v7i2.626.
- [11] J. Kaliappan et al., "Analyzing classification and feature selection strategies for diabetes prediction across diverse diabetes datasets," *Front. Artif. Intell.*, vol. 7, 2024, doi: 10.3389/frai.2024.1421751.
- [12] M. Talebi Moghaddam et al., "Predicting diabetes in adults: identifying important features in unbalanced data over a 5-year cohort study using machine learning algorithm," *BMC Med. Res. Methodol.*, vol. 24, no. 1, 2024, doi: 10.1186/s12874-024-02341-z.
- [13] 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*, San Francisco: Association for Computing Machinery (ACM), 2016, pp. 785–794. doi: 10.1145/2939672.2939785.
- [14] T. M. Shami, A. A. El-Saleh, M. Alswaitti, Q. Al-Tashi, M. A. Summakieh, and S. Mirjalili, "Particle Swarm Optimization: A Comprehensive Survey," *IEEE Access*, vol. 10, pp. 10031–10061, 2022, doi: 10.1109/ACCESS.2022.3142859.
- [15] L.J.V. Miranda, "PySwarms: a research toolkit for Particle Swarm Optimization in Python," *The Journal of Open Source Software*, vol. 3, no. 21, p. 433, 2018, doi: 10.21105/joss.00433