Jurnal Minfo Polgan (JMP)
Volume 15, Nomor 3, September-Desember 2026
DOI : https://doi.org/10.33395/jmp.v15i3.16857
e-ISSN : 2797-3298
p-ISSN : 2089-9424

Optimasi Algoritma *K-Nearest Neighbor* Menggunakan *Particle Swarm Optimization* untuk Prediksi Penyakit Cacar Monyet

[1]Muhamad Imam, [2] Bayu Arisandi, [3]Pahla Widhiani
[4]Fitri Dian Perwitasari, [5]Shindy Zhaski Pradita
[1*,2]Program Studi Teknik Informatika dan Peternakan, Fakultas Teknik
Universitas Muhammadiyah Cirebon, Cirebon, Indonesia
[3,4,5]Manajemen Informatika, Politeknik Ganesha Medan, Medan, Indonesia
*Korespondensi: muhamad_imam@umc.ac.id

Submit : 17 Sept 2026 | **Diterima** : 30 Okt 2026 | **Terbit** : 08 Nov 2026

ABSTRACT

Cacar monyet is a zoonotic disease that can spread through direct contact with infected humans or animals. The increasing number of cases reported in various countries highlights the importance of early detection to support timely diagnosis and treatment. One promising approach is the application of Machine Learning techniques, particularly the K-Nearest Neighbor (KNN) algorithm. However, the performance of KNN is highly dependent on the selection of the optimal K parameter. Therefore, an optimization method is required to improve its classification performance. This study aims to optimize the K-Nearest Neighbor algorithm using Particle Swarm Optimization (PSO) for cacar monyet disease prediction. The dataset used consists of 500 records with 50 attributes obtained from Kaggle. The research methodology includes Exploratory Data Analysis (EDA), data preprocessing, training and testing data partitioning, baseline KNN model development, optimization of the K parameter using PSO, and model evaluation using Accuracy, Precision, Recall, F1-Score, and Confusion Matrix metrics. The results indicate that PSO optimization improves the performance of the KNN model by identifying a more optimal K value compared with the baseline model. These findings demonstrate that the integration of PSO and KNN has the potential to enhance classification performance and can be considered an effective approach for developing decision support systems for the early detection of cacar monyet disease.

***Keywords**: Machine Learning, K-Nearest Neighbor, Particle Swarm Optimization, Cacar monyet, Classification.*

ABSTRAK

Penyakit cacar monyet (Cacar monyet) merupakan penyakit zoonosis yang dapat menular dari hewan ke manusia maupun antarmanusia melalui kontak langsung dengan individu, cairan tubuh, atau benda yang terkontaminasi. Peningkatan jumlah kasus cacar monyet di berbagai negara dalam beberapa tahun terakhir menunjukkan pentingnya upaya deteksi dini untuk mendukung proses diagnosis dan penanganan yang lebih cepat serta mengurangi risiko penyebaran penyakit. Salah satu pendekatan yang dapat diterapkan dalam proses prediksi penyakit adalah pemanfaatan algoritma Machine Learning, khususnya *K-Nearest Neighbor* (KNN), yang dikenal memiliki proses klasifikasi sederhana dan efektif. Namun, performa algoritma KNN sangat dipengaruhi oleh pemilihan nilai parameter *K*, sehingga diperlukan teknik optimasi untuk memperoleh hasil klasifikasi yang lebih akurat. Penelitian ini bertujuan mengoptimalkan algoritma *K-Nearest Neighbor* menggunakan metode *Particle Swarm Optimization* (PSO) dalam prediksi penyakit cacar monyet. Dataset yang digunakan terdiri atas 500 data dengan 50 atribut yang diperoleh dari Kaggle. Tahapan penelitian meliputi *Exploratory Data Analysis* (EDA), *data preprocessing* yang mencakup penanganan data, normalisasi, dan seleksi fitur, pembagian data menjadi data latih dan data uji, pembangunan model KNN sebagai model dasar, optimasi nilai parameter *K* menggunakan PSO, serta evaluasi performa model menggunakan metrik *Accuracy*, *Precision*, *Recall*, *F1-Score*, dan *Confusion Matrix*. Hasil penelitian menunjukkan bahwa penerapan PSO mampu menentukan nilai *K* yang lebih optimal sehingga meningkatkan performa model KNN dibandingkan dengan model tanpa optimasi. Temuan ini menunjukkan bahwa integrasi PSO dan KNN berpotensi menjadi pendekatan yang

@2026 The Author Published By Politeknik Ganesha Medan
This is an open access article under the CC BY licence (http://creativecommons.org/licenses/by/4.0/). which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properlycited.

efektif dalam meningkatkan akurasi klasifikasi serta mendukung pengembangan sistem pendukung keputusan untuk deteksi dini penyakit cacar monyet secara lebih cepat dan akurat.

Kata Kunci: Machine Learning, K-Nearest Neighbor, Particle Swarm Optimization, Cacar monyet, Klasifikasi.

PENDAHULUAN

Mpox, atau yang sebelumnya dikenal sebagai cacar monyet, merupakan penyakit zoonosis yang disebabkan oleh virus cacar monyet dari genus Orthopoxvirus (Bunge et al., 2022) . Penyakit ini dapat menular melalui kontak langsung dengan penderita, cairan tubuh, lesi kulit, benda yang terkontaminasi, maupun hewan yang terinfeksi (World Health Organization [WHO], 2024). Gejala yang ditimbulkan meliputi demam, pembengkakan kelenjar getah bening, nyeri otot, sakit kepala, dan ruam kulit yang berkembang menjadi vesikel maupun pustula (Lahariya et al., 2022). Meskipun sebagian besar pasien dapat sembuh secara spontan, kelompok rentan seperti anak-anak, ibu hamil, dan individu dengan gangguan sistem imun berisiko mengalami komplikasi yang lebih berat (Lahariya et al., 2022.; Sanchez Clemente et al., 2024) . Organisasi Kesehatan Dunia (WHO, 2024) menyatakan bahwa mpox masih menjadi perhatian kesehatan masyarakat karena kasus baru masih terus dilaporkan di berbagai negara dan diperlukan upaya peningkatan deteksi dini, surveilans, serta diagnosis yang cepat untuk mengurangi penyebaran penyakit.

Proses diagnosis mpox memiliki tantangan tersendiri karena gejala klinisnya memiliki kemiripan dengan beberapa penyakit kulit lainnya, seperti *smallpox*, *chickenpox*, campak, maupun infeksi kulit akibat virus lainnya. Kesamaan manifestasi klinis tersebut dapat menyebabkan keterlambatan diagnosis apabila hanya mengandalkan observasi gejala. Pemeriksaan laboratorium seperti Polymerase Chain Reaction (PCR) masih menjadi standar konfirmasi, namun ketersediaan fasilitas laboratorium tidak selalu merata dan memerlukan waktu tertentu untuk memperoleh hasil. Oleh karena itu, diperlukan suatu pendekatan komputasi yang mampu membantu proses identifikasi awal berdasarkan karakteristik pasien dan gejala klinis sehingga dapat mendukung pengambilan keputusan secara lebih cepat.

Perkembangan teknologi Artificial Intelligence (AI), khususnya Machine Learning, telah memberikan kontribusi yang signifikan dalam bidang kesehatan (Sidey-Gibbons & Sidey-Gibbons, 2019). Berbagai penelitian menunjukkan bahwa algoritma Machine Learning mampu membantu proses diagnosis penyakit melalui analisis pola data medis secara otomatis (Sidey-Gibbons & Sidey-Gibbons, 2019; Uddin et al., 2019). Berbagai metode klasifikasi seperti Decision Tree, Naïve Bayes, Support Vector Machine (SVM), Random Forest, dan *K-Nearest Neighbor* (KNN) telah banyak diterapkan untuk memprediksi maupun mengklasifikasikan berbagai jenis penyakit (Uddin et al., 2019) . Dibandingkan dengan beberapa algoritma lainnya, KNN memiliki keunggulan berupa konsep yang sederhana, implementasi yang mudah, serta kemampuan untuk memberikan performa klasifikasi yang baik pada berbagai jenis dataset tanpa memerlukan proses pelatihan model yang kompleks (Zhang, 2016) .

Meskipun demikian, performa algoritma *K-Nearest Neighbor* sangat dipengaruhi oleh pemilihan nilai parameter *K* sebagai jumlah tetangga terdekat yang digunakan dalam proses klasifikasi. Nilai *K* yang terlalu kecil dapat menyebabkan model menjadi sensitif terhadap *noise* sehingga meningkatkan risiko *overfitting*, sedangkan nilai *K* yang terlalu besar dapat menyebabkan model kehilangan kemampuan membedakan karakteristik lokal data dan menghasilkan *underfitting*. Pada banyak penelitian, nilai *K* masih ditentukan secara manual melalui proses *trial and error*, sehingga belum tentu menghasilkan performa klasifikasi yang optimal.

Salah satu pendekatan yang dapat digunakan untuk mengatasi permasalahan tersebut adalah *Particle Swarm Optimization* (PSO). PSO merupakan algoritma optimasi berbasis populasi yang terinspirasi dari perilaku sosial kawanan burung atau ikan dalam mencari sumber makanan (Sengupta et al., 2019) . Algoritma ini memiliki kemampuan untuk mencari solusi optimal secara iteratif melalui pembaruan posisi dan kecepatan setiap partikel berdasarkan pengalaman individu (personal best) dan pengalaman terbaik seluruh populasi (global best) (Shami et al., 2022a) . Dalam penelitian ini, PSO dimanfaatkan untuk menentukan nilai parameter K yang optimal pada algoritma KNN sehingga proses klasifikasi diharapkan menghasilkan performa yang lebih baik dibandingkan dengan penggunaan nilai K yang ditentukan secara manual (Dafid et al., 2025) .

Berbagai penelitian sebelumnya telah menunjukkan bahwa algoritma KNN memiliki performa yang baik dalam menyelesaikan permasalahan klasifikasi di bidang kesehatan (Uddin et al., 2019) . Di sisi lain, sejumlah penelitian juga melaporkan bahwa penerapan algoritma optimasi seperti PSO mampu meningkatkan performa model Machine Learning melalui optimasi parameter dan seleksi fitur (Dafid et al., 2025). Namun demikian, penelitian yang secara khusus mengoptimalkan parameter K pada algoritma KNN untuk prediksi penyakit mpox menggunakan dataset klinis masih relatif terbatas (Farzipour et al., 2023). Selain itu, sebagian penelitian lebih berfokus pada klasifikasi berbasis citra kulit menggunakan deep learning (Deng et al., 2024) ; (Ozsahin et al., 2023) , sedangkan pendekatan berbasis data klinis dan gejala pasien masih memerlukan pengembangan lebih lanjut (Farzipour et al., 2023) .

Berdasarkan kondisi tersebut, penelitian ini mengusulkan optimasi algoritma *K-Nearest Neighbor* menggunakan *Particle Swarm Optimization* untuk prediksi penyakit cacar monyet. Dataset yang digunakan merupakan dataset publik dari Kaggle yang terdiri atas 500 data dengan 50 atribut. Tahapan penelitian meliputi *Exploratory Data Analysis* (EDA), *data preprocessing*, pembagian data menjadi data latih dan data uji, pembangunan model KNN sebagai model dasar (*baseline*), optimasi parameter *K* menggunakan PSO, serta evaluasi model menggunakan metrik Accuracy, Precision, Recall, F1-Score, dan Confusion Matrix.

Kontribusi utama penelitian ini adalah penerapan *Particle Swarm Optimization* sebagai metode optimasi untuk menentukan nilai parameter *K* yang optimal pada algoritma *K-Nearest Neighbor* sehingga mampu meningkatkan performa klasifikasi dibandingkan dengan model KNN tanpa optimasi. Hasil penelitian ini diharapkan dapat menjadi referensi dalam pengembangan sistem pendukung keputusan berbasis *Machine Learning* yang membantu proses deteksi dini penyakit cacar monyet secara lebih efektif.

METODE PENELITIAN

Desain Penelitian

Penelitian ini menggunakan pendekatan kuantitatif dengan metode eksperimen (*experimental research*) untuk mengevaluasi pengaruh optimasi parameter *K* menggunakan *Particle Swarm Optimization* (PSO) terhadap performa algoritma *K-Nearest Neighbor* (KNN) dalam prediksi penyakit cacar monyet. Seluruh proses penelitian diimplementasikan menggunakan bahasa pemrograman Python pada platform Google Colaboratory dengan memanfaatkan pustaka *Pandas*, *NumPy*, *Scikit-learn*, *PySwarms*, *Matplotlib*, dan *Seaborn*.

Penelitian dilakukan melalui beberapa tahapan, yaitu pengumpulan dataset, *Exploratory Data Analysis* (EDA), *data preprocessing*, pembagian data menjadi data latih dan data uji, pembangunan model KNN sebagai model dasar (*baseline*), optimasi parameter *K* menggunakan PSO, evaluasi model menggunakan beberapa metrik klasifikasi, serta analisis hasil penelitian.

Dataset Penelitian

Penelitian menggunakan dataset publik cacar monyet yang diperoleh dari platform Kaggle. Dataset berisi informasi karakteristik pasien, gejala klinis, riwayat kesehatan, serta atribut pendukung yang berkaitan dengan proses diagnosis penyakit.

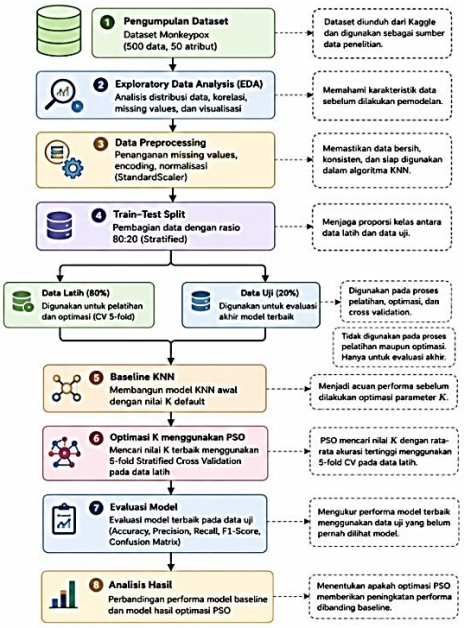
Tabel 1. Karakteristik Dataset

Karakteristik	Keterangan
Nama Dataset	Cacar monyet Disease Dataset
Sumber	Kaggle
Jumlah Data	500 data
Jumlah Atribut	50 atribut
Jenis Data	Numerik dan kategorikal
Variabel Target	Diagnosing (atau disesuaikan dengan target akhir penelitian)
Platform Implementasi	Google Colaboratory
Bahasa Pemrograman	Python

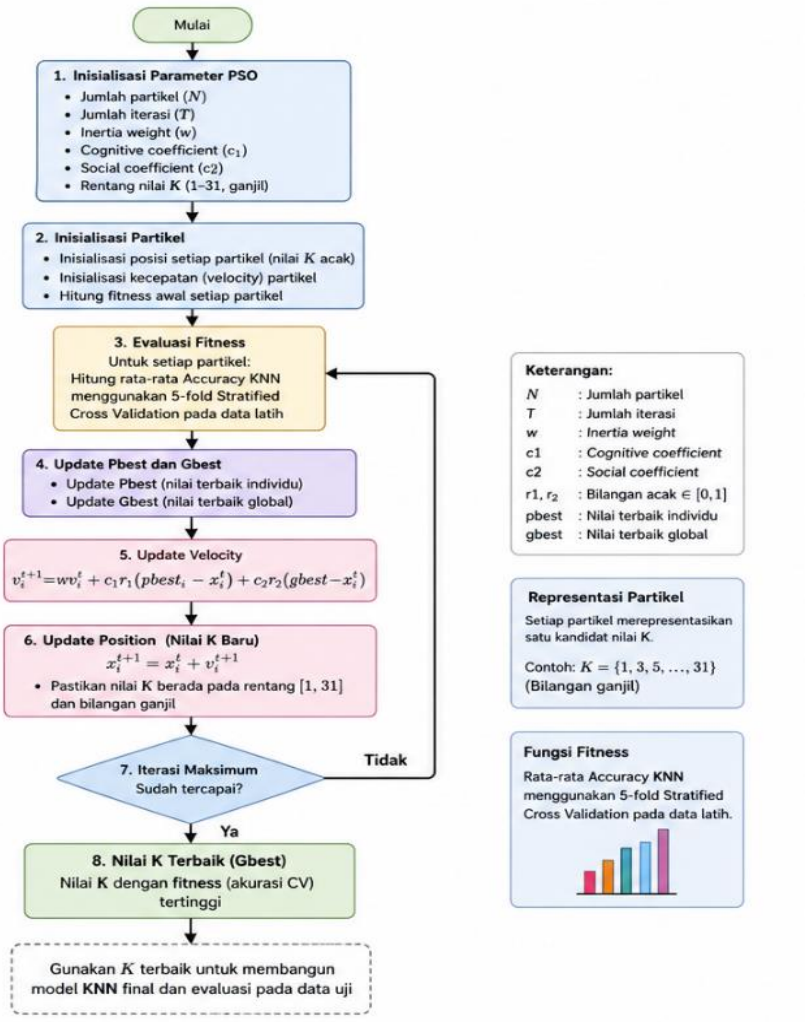
Sebelum digunakan dalam proses klasifikasi, dataset diperiksa untuk mengidentifikasi *missing values*, data duplikat, inkonsistensi format, dan distribusi kelas. Tahapan ini bertujuan memastikan kualitas data sehingga model yang dibangun memiliki tingkat keandalan yang lebih baik.

Tahapan Penelitian

Penelitian dilaksanakan melalui tahapan yang ditunjukkan pada



Gambar 1. Framework Penelitian



Gambar 2. *Flowchart* Optimasi Parameter K dengan PSO

Tahapan penelitian dijelaskan sebagai berikut.

1.Pengumpulan Dataset.
Dataset diunduh dari Kaggle dan digunakan sebagai sumber data penelitian.
2. *Exploratory Data Analysis* (EDA).
Analisis dilakukan untuk mengetahui karakteristik data, distribusi kelas, korelasi antar atribut, dan mendeteksi adanya data yang tidak konsisten.
3. Data *Preprocessing*.
Tahapan ini meliputi penanganan *missing values*, transformasi atribut kategorikal menggunakan teknik *encoding*, normalisasi data numerik menggunakan *StandardScaler*, serta pemisahan fitur dan variabel target.
4. Pembagian Dataset.
Dataset dibagi menjadi data latih dan data uji menggunakan rasio 80:20 dengan metode *Stratified Train-Test Split* agar proporsi setiap kelas tetap terjaga.
5. Pembangunan Model Baseline.
Algoritma KNN dibangun menggunakan parameter awal sebagai model pembanding sebelum dilakukan optimasi.
6. Optimasi Parameter K menggunakan PSO.
PSO digunakan untuk mencari nilai parameter *K* terbaik berdasarkan nilai rata-rata akurasi hasil 5-fold Stratified Cross Validation pada data latih.
7. Evaluasi Model.
Model terbaik dievaluasi menggunakan data uji untuk memperoleh nilai Accuracy, Precision, Recall, F1-Score, serta Confusion Matrix.
8. Analisis Hasil.
Hasil model baseline dibandingkan dengan model hasil optimasi PSO untuk mengetahui peningkatan performa klasifikasi.

Data Preprocessing

Tahapan *data preprocessing* dilakukan agar dataset siap digunakan dalam proses pembelajaran mesin. Tahapan yang dilakukan meliputi:

1. Pemeriksaan *missing values*.
2. Penghapusan data duplikat (jika ada).
3. Transformasi atribut kategorikal menjadi numerik menggunakan teknik *encoding* yang sesuai.
4. Normalisasi atribut numerik menggunakan *StandardScaler*.
5. Pembagian data menjadi data latih dan data uji.

Untuk menghindari *data leakage*, proses normalisasi dilakukan setelah pembagian data. Model *StandardScaler* hanya dilatih (*fit*) menggunakan data latih, kemudian digunakan untuk mentransformasi data latih dan data uji.

Algoritma *K-Nearest Neighbor* (KNN)

K-Nearest Neighbor merupakan algoritma klasifikasi berbasis *lazy learning* yang menentukan kelas suatu data berdasarkan mayoritas kelas dari sejumlah tetangga terdekatnya. Kedekatan antar data dihitung menggunakan jarak Euclidean [3].
Persamaan jarak Euclidean adalah:

$$d(x,y) = \sqrt{\sum_{i=1}^{n}(x_i - y_i)^2}$$

Keterangan:

- $d(x,y)$ = jarak antar data
- x_i = data training
- y_i = data testing

Nilai parameter *K* menentukan jumlah tetangga yang digunakan dalam proses klasifikasi. Nilai *K* yang terlalu kecil berpotensi menyebabkan *overfitting*, sedangkan nilai *K* yang terlalu

besar dapat menyebabkan *underfitting*. Oleh karena itu, diperlukan metode optimasi untuk memperoleh nilai *K* yang optimal.

Particle Swarm Optimization (PSO)

Particle Swarm Optimization merupakan algoritma optimasi berbasis populasi yang diperkenalkan oleh Kennedy dan Eberhart [1,2]. Dalam penelitian ini, setiap partikel merepresentasikan kandidat nilai parameter *K* pada algoritma KNN.
Kecepatan partikel diperbarui menggunakan persamaan:

$$v_i^{t+1} = wv_i^t + c_1 r_1 (pbest_i - x_i^t) + c_2 r_2 (gbest - x_i^t)$$

Sedangkan posisi partikel diperbarui menggunakan persamaan:

$$x_i^{t+1} = x_i^t + v_i^{t+1}$$

Parameter PSO yang digunakan pada penelitian ini ditunjukkan pada Tabel 2.

Tabel 2. Parameter Particle Swarm Optimization

Parameter	Nilai
Jumlah Partikel	20
Jumlah Iterasi	30
Inertia Weight (w)	0,70
Cognitive Coefficient (c1)	1,50
Social Coefficient (c2)	1,50
Rentang Nilai K	1–31 (bilangan ganjil)

Fungsi objektif (*fitness function*) yang digunakan adalah rata-rata nilai Accuracy hasil 5-fold Stratified Cross Validation pada data latih. Pendekatan ini dipilih agar proses optimasi tidak memanfaatkan data uji sehingga hasil evaluasi lebih objektif.

Evaluasi Model

Performa model dievaluasi menggunakan data uji yang tidak digunakan selama proses pelatihan maupun optimasi. Metrik evaluasi yang digunakan meliputi Accuracy, Precision, Recall, dan F1-Score.
Accuracy dihitung menggunakan persamaan:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \times 100\%$$

Keterangan:

TP = True Positive
TN = True Negative
FP = False Positive
FN = False Negative

Selain metrik tersebut, penelitian juga menggunakan **Confusion Matrix** untuk menganalisis kemampuan model dalam mengklasifikasikan setiap kelas serta membandingkan performa model KNN sebelum dan sesudah optimasi menggunakan PSO.

HASIL DAN PEMBAHASAN

Hasil

Deskripsi Dataset

Penelitian ini menggunakan dataset Cacar monyet Disease yang diperoleh dari platform Kaggle. Dataset terdiri atas 500 data dengan 50 atribut yang merepresentasikan karakteristik pasien, gejala klinis, riwayat kesehatan, serta variabel pendukung lainnya yang digunakan untuk proses klasifikasi penyakit cacar monyet. Variabel target pada penelitian ini adalah **Diagnosing**, yang menunjukkan hasil diagnosis setiap pasien.

Sebelum dilakukan proses pemodelan, dataset dianalisis melalui tahapan *Exploratory Data Analysis* (EDA) untuk mengetahui struktur data, tipe data, distribusi kelas, serta kualitas dataset. Berdasarkan hasil pemeriksaan, dataset tidak memiliki nilai hilang (*missing value*) maupun data duplikat, sehingga dapat langsung digunakan pada proses *preprocessing*.

Tabel 3. Karakteristik Dataset

Karakteristik	Nilai
Jumlah Data	500
Jumlah Atribut	50
Jenis Data	Numerik dan Kategorikal
Variabel Target	Diagnosing
Missing Value	Tidak ditemukan
Data Duplikat	Tidak ditemukan

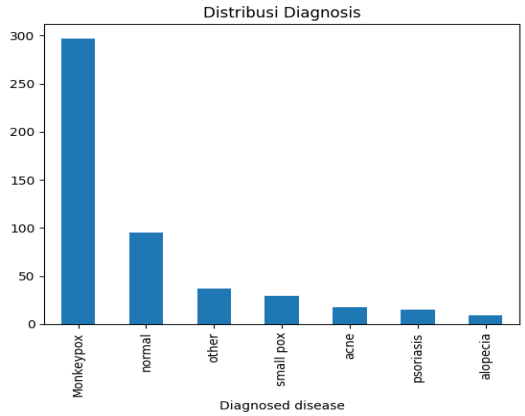
Hasil identifikasi menunjukkan bahwa dataset telah memenuhi persyaratan untuk digunakan dalam proses klasifikasi menggunakan algoritma K-Nearest Neighbor.

Exploratory Data Analysis (EDA)

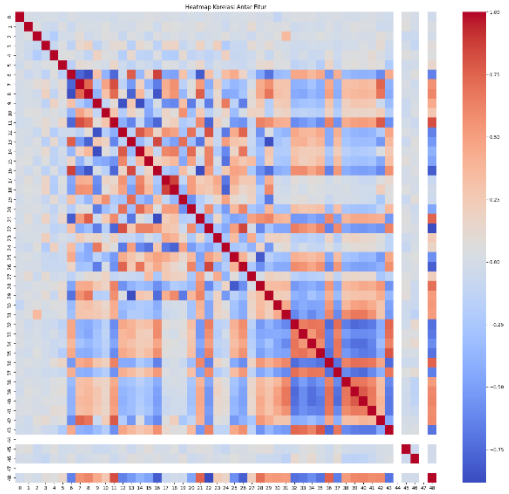
Tahap EDA bertujuan untuk memahami karakteristik data sebelum dilakukan proses pemodelan. Analisis dilakukan melalui pemeriksaan distribusi kelas, statistik deskriptif, serta hubungan antaratribut menggunakan visualisasi data.

Visualisasi distribusi kelas menunjukkan bahwa jumlah data pada setiap kelas relatif seimbang sehingga risiko *class imbalance* relatif kecil. Kondisi tersebut mendukung proses pembelajaran model karena setiap kelas memiliki representasi yang memadai.

Selanjutnya dilakukan analisis korelasi antaratribut menggunakan *heatmap*. Hasil visualisasi menunjukkan bahwa sebagian besar atribut tidak memiliki korelasi yang sangat tinggi, sehingga seluruh atribut tetap dipertahankan pada proses klasifikasi. Hal ini sejalan dengan tujuan penelitian yang berfokus pada optimasi parameter *K* pada algoritma KNN, bukan pada proses seleksi fitur.



Gambar 3. Distribusi Kelas Dataset



Gambar 4. Heatmap Korelasi Antar Fitur

Hasil Data Preprocessing

Tahapan *data preprocessing* dilakukan untuk memastikan dataset berada dalam kondisi optimal sebelum digunakan pada proses pembelajaran mesin. Tahapan yang dilakukan meliputi pemeriksaan *missing value*, pemeriksaan data duplikat, transformasi atribut kategorikal menjadi numerik menggunakan teknik *encoding*, normalisasi atribut numerik menggunakan **StandardScaler**, serta pembagian dataset menjadi data latih dan data uji dengan rasio 80:20.

Normalisasi dilakukan agar setiap atribut memiliki skala yang seragam sehingga proses perhitungan jarak Euclidean pada algoritma KNN tidak didominasi oleh atribut tertentu. Selain itu, pembagian dataset dilakukan secara *stratified* untuk menjaga proporsi setiap kelas pada data latih dan data uji.

Tabel 4. Hasil Data Preprocessing

Tahapan	Hasil
Missing Value	Tidak ditemukan
Data Duplikat	Tidak ditemukan
Encoding	Berhasil
Standardisasi	StandardScaler
Train-Test Split	80 : 20

Tahapan *preprocessing* menghasilkan dataset yang siap digunakan dalam pembangunan model klasifikasi.

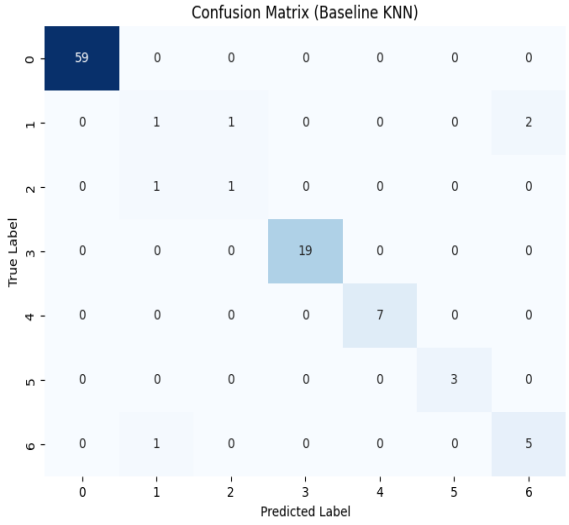
Hasil Model Baseline K-Nearest Neighbor

Sebagai model dasar (*baseline*), penelitian ini menggunakan algoritma *K-Nearest Neighbor* tanpa optimasi parameter. Model dievaluasi menggunakan data uji dengan metrik Accuracy, Precision, Recall, dan F1-Score. Hasil evaluasi menunjukkan bahwa model baseline memperoleh Accuracy sebesar 95%, Weighted Precision sebesar 94,62%, dan Weighted Recall sebesar 95%. Nilai tersebut menunjukkan bahwa algoritma KNN memiliki kemampuan yang baik dalam mengklasifikasikan data penyakit cacar monyet.

Tabel 5. Hasil Evaluasi Baseline KNN

Metrik	Nilai
Accuracy	95%
Precision (Weighted)	94,62%
Recall (Weighted)	95,00%
F1-Score	*(Macro Avg F1-score): 0.79* *(Weighted Avg F1-score): 0.95*

Analisis menggunakan *Confusion Matrix* menunjukkan bahwa sebagian besar data berhasil diklasifikasikan dengan benar. Meskipun demikian, masih terdapat beberapa kesalahan klasifikasi pada beberapa kelas, yang mengindikasikan bahwa pemilihan parameter *K* masih dapat ditingkatkan melalui proses optimasi.



Gambar 5. Confusion Matrix Baseline KNN

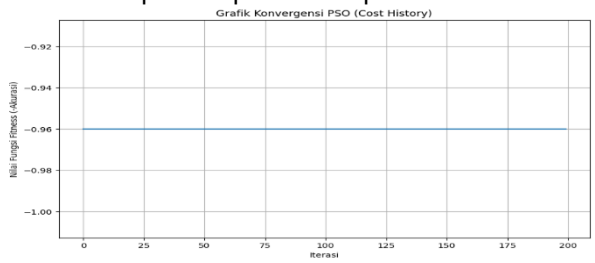
Hasil Optimasi Menggunakan Particle Swarm Optimization

Pada tahap ini dilakukan optimasi parameter *K* menggunakan algoritma *Particle Swarm Optimization* (PSO). Setiap partikel merepresentasikan kandidat nilai *K*, sedangkan fungsi objektif dirancang untuk memperoleh nilai akurasi terbaik dari model KNN. Hasil optimasi menunjukkan bahwa PSO memperoleh nilai K terbaik sebesar 1. Parameter tersebut kemudian digunakan untuk membangun model klasifikasi akhir. Model KNN hasil optimasi menghasilkan Accuracy sebesar 96%, Weighted Precision sebesar 97,08%, dan Weighted Recall sebesar 96%, yang menunjukkan adanya peningkatan dibandingkan dengan model baseline.

Tabel 6. Hasil Optimasi PSO

Parameter	Nilai
Best K	1
Accuracy	96%
Precision (Weighted)	97,08%
Recall (Weighted)	96,00%
F1-Score	*(Macro Avg F1-score): 0.83* *(Weighted Avg F1-score): 0.95*

Grafik konvergensi menunjukkan bahwa nilai *fitness* meningkat secara bertahap hingga mencapai kondisi stabil pada iterasi akhir. Hal ini mengindikasikan bahwa PSO berhasil menemukan solusi optimal dalam proses pencarian parameter *K*.



Gambar 6. Grafik Konvergensi PSO

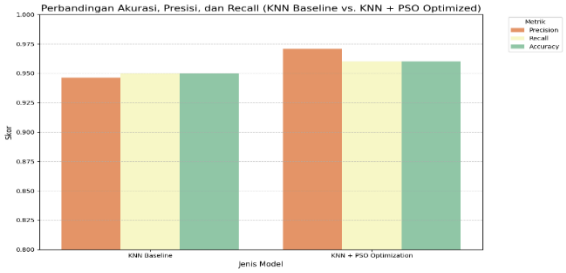
Perbandingan Performa Model

Performa model baseline dibandingkan dengan model hasil optimasi untuk mengetahui pengaruh penerapan *Particle Swarm Optimization* terhadap algoritma KNN.

Tabel 7 Perbandingan Performa Model

Model	Accuracy	Precision	Recall
KNN	95,00%	94,62%	95,00%
KNN + PSO	**96,00%**	**97,08%**	**96,00%**

Berdasarkan Tabel 7, optimasi menggunakan PSO berhasil meningkatkan seluruh metrik evaluasi. Nilai accuracy meningkat sebesar 1%, weighted precision meningkat sekitar 2,46%, dan weighted recall meningkat sebesar 1%. Peningkatan tersebut menunjukkan bahwa optimasi parameter *K* memberikan dampak positif terhadap kemampuan model dalam melakukan klasifikasi.



Gambar 7. Perbandingan *Accuracy, Precision, dan Recall*

Grafik pada Gambar 7 memvisualisasikan perbandingan akurasi, presisi, dan recall antara model KNN *Baseline* dan model KNN yang dioptimasi dengan PSO. Dari grafik ini, terlihat jelas bagaimana optimasi PSO berkontribusi pada peningkatan kinerja model secara

keseluruhan: Akurasi: Model KNN + PSO menunjukkan sedikit peningkatan akurasi dibandingkan dengan KNN Baseline, menandakan kemampuan klasifikasi yang lebih baik secara umum. Presisi: Ada peningkatan signifikan pada presisi model KNN + PSO, yang berarti model ini lebih baik dalam menghindari kesalahan positif (memprediksi kelas positif padahal sebenarnya negatif). Recall: Recall juga menunjukkan peningkatan untuk model KNN + PSO, menunjukkan bahwa model ini lebih baik dalam mengidentifikasi semua kasus positif yang ada.

Pembahasan

Hasil penelitian menunjukkan bahwa penerapan *Particle Swarm Optimization* (PSO) mampu meningkatkan performa algoritma *K-Nearest Neighbor* (KNN) dalam prediksi penyakit cacar monyet (Dafid et al., 2025). Berdasarkan hasil pengujian, model KNN tanpa optimasi memperoleh nilai akurasi sebesar 95%, sedangkan model KNN yang dioptimalkan menggunakan PSO menghasilkan akurasi sebesar 96%. Selain itu, nilai *weighted precision* meningkat dari 94,62% menjadi 97,08%, sedangkan *weighted recall* meningkat dari 95% menjadi 96%. Peningkatan pada ketiga metrik tersebut menunjukkan bahwa proses optimasi parameter memberikan kontribusi positif terhadap kemampuan model dalam mengklasifikasikan data (Dafid et al., 2025) .

Peningkatan performa tersebut berkaitan dengan karakteristik algoritma KNN yang sangat bergantung pada pemilihan nilai parameter K (Dafid et al., 2025; Zhang, 2016). Pada model *baseline*, nilai K ditentukan secara manual sehingga belum tentu sesuai dengan karakteristik dataset (Zhang, 2016). Sebaliknya, pada penelitian ini PSO melakukan pencarian nilai K secara otomatis melalui mekanisme optimasi berbasis populasi (Shami et al., 2022a) sehingga diperoleh nilai K = 1 sebagai solusi terbaik. Hasil ini menunjukkan bahwa pada dataset yang digunakan, informasi dari tetangga terdekat sudah cukup untuk membedakan kelas data secara efektif. Meskipun demikian, nilai K yang sangat kecil berpotensi sensitif terhadap *noise* dan *outlier* sehingga memunculkan risiko *overfitting* (Xiao et al., 2026); (Zhang, 2016) .

Analisis terhadap confusion matrix menunjukkan bahwa setelah proses optimasi, jumlah prediksi yang benar meningkat pada beberapa kelas(Rainio et al., 2024; Zhang, 2016). Hal ini mengindikasikan bahwa optimasi parameter tidak hanya meningkatkan nilai akurasi secara keseluruhan, tetapi juga memperbaiki kemampuan model dalam mengenali karakteristik masing-masing kelas. Meskipun masih terdapat sejumlah kecil kesalahan klasifikasi, distribusi kesalahan menjadi lebih rendah dibandingkan dengan model baseline, sehingga model hasil optimasi memiliki kemampuan generalisasi yang lebih baik terhadap data uji.

Dari sudut pandang algoritma, keberhasilan PSO disebabkan oleh kemampuannya dalam mengeksplorasi ruang pencarian parameter secara sistematis (Shami et al., 2022a). Berbeda dengan pendekatan trial and error, PSO memanfaatkan informasi terbaik dari setiap partikel (personal best) dan seluruh populasi (global best) untuk memperbarui posisi solusi secara iteratif (Sengupta et al., 2019); (Shami et al., 2022a) . Dengan mekanisme tersebut, proses pencarian nilai K menjadi lebih efisien (Dafid et al., 2025) dan berpeluang menemukan solusi yang lebih optimal daripada pemilihan parameter secara manual.

Hasil penelitian ini juga mendukung berbagai penelitian sebelumnya yang melaporkan bahwa algoritma optimasi berbasis populasi mampu meningkatkan performa model machine learning melalui optimasi hiperparameter (Dafid et al., 2025) ; (Shami et al., 2022a). Pada berbagai studi klasifikasi di bidang kesehatan, penerapan PSO dilaporkan dapat meningkatkan stabilitas model sekaligus memperbaiki nilai metrik evaluasi (Ileberi & Sun, 2024). Hasil penelitian ini memperkuat temuan tersebut dengan menunjukkan bahwa optimasi menggunakan PSO juga efektif diterapkan pada algoritma KNN untuk prediksi penyakit cacar monyet. Kontribusi utama penelitian ini terletak pada penerapan PSO sebagai metode optimasi parameter K pada algoritma KNN dalam konteks prediksi penyakit cacar monyet menggunakan dataset klinis. Berbeda dengan sebagian penelitian yang menggunakan nilai K berdasarkan percobaan manual (Zhang, 2016) atau berfokus pada klasifikasi berbasis citra (Deng et al., 2024; Uzun Ozsahin et al., 2023b), penelitian ini menunjukkan bahwa optimasi parameter secara otomatis dapat memberikan peningkatan performa pada pendekatan berbasis data tabular (Farzipour et al., 2023) . Dengan demikian, penelitian ini memberikan alternatif pendekatan yang sederhana namun efektif untuk meningkatkan kualitas model klasifikasi.

Walaupun menunjukkan hasil yang baik, penelitian ini masih memiliki beberapa keterbatasan. Pertama, dataset yang digunakan berasal dari satu sumber publik sehingga

belum sepenuhnya merepresentasikan variasi karakteristik pasien pada kondisi nyata. Kedua, optimasi hanya difokuskan pada parameter K tanpa mempertimbangkan optimasi hiperparameter lain, seperti jenis metrik jarak (distance metric) atau pembobotan tetangga (distance weighting) (Uddin et al., 2019; Zhang, 2016) . Ketiga, evaluasi model masih terbatas pada satu dataset, sehingga diperlukan pengujian pada dataset lain untuk menguji konsistensi performa model.

Sebagai arah penelitian selanjutnya, optimasi dapat dikembangkan dengan membandingkan PSO dengan algoritma optimasi lain seperti Genetic Algorithm (GA), Grey Wolf Optimizer (GWO), atau Whale Optimization Algorithm (WOA) (Mohammed et al., 2019; Shami et al., 2022b). Selain itu, penelitian juga dapat mengombinasikan optimasi hiperparameter dengan metode seleksi fitur sehingga model yang dihasilkan tidak hanya memiliki akurasi yang lebih tinggi, tetapi juga kompleksitas yang lebih rendah (Ileberi & Sun, 2024; Shami et al., 2022b). Penggunaan dataset yang lebih besar dan berasal dari berbagai institusi kesehatan juga akan meningkatkan kemampuan generalisasi model dan memperkuat validitas hasil penelitian.

KESIMPULAN

Penelitian ini bertujuan untuk menganalisis pengaruh optimasi *Particle Swarm Optimization* (PSO) terhadap performa algoritma *K-Nearest Neighbor* (KNN) dalam prediksi penyakit cacar monyet menggunakan dataset Cacar Monyet Disease. Berdasarkan hasil penelitian, penerapan PSO berhasil mengoptimalkan parameter K pada algoritma KNN sehingga diperoleh nilai K = 1 sebagai parameter terbaik untuk dataset yang digunakan. Hasil pengujian memperlihatkan bahwa KNN tanpa optimasi mencapai accuracy 95%, sedangkan KNN yang dioptimalkan dengan PSO naik menjadi 96%. Weighted precision meningkat dari 94,62% menjadi 97,08%, dan weighted recall bertambah dari 95% menjadi 96%. Kenaikan ini membuktikan bahwa PSO mampu memperbaiki kemampuan klasifikasi dibandingkan pemilihan parameter secara manual. Dengan demikian, penggabungan KNN dan PSO merupakan pendekatan yang efektif untuk meningkatkan kinerja klasifikasi pada dataset penyakit cacar monyet. Walaupun accuracy hanya naik 1%, perbaikan precision dan recall menandakan model semakin tepat dalam mengenali setiap kelas. Hal ini menegaskan bahwa optimasi hiperparameter berperan penting dalam meningkatkan performa algoritma berbasis instance-based learning. Meski begitu, penelitian ini memiliki sejumlah keterbatasan. Pertama, hanya satu dataset cacar monyet yang digunakan, sehingga hasilnya belum tentu berlaku pada dataset lain dengan karakteristik berbeda. Kedua, selisih accuracy yang kecil belum diuji signifikansi statistiknya, sehingga belum pasti bukan akibat variasi pembagian data. Ketiga, optimasi hanya memakai PSO tanpa pembanding seperti Grid Search, Random Search, atau Genetic Algorithm. Keempat, waktu komputasi, ketidakseimbangan kelas, dan seleksi fitur belum dikaji. Penelitian selanjutnya disarankan memakai dataset yang lebih beragam, uji statistik, metode optimasi pembanding, serta kombinasi dengan seleksi fitur dan penanganan ketidakseimbangan kelas agar hasil yang diperoleh lebih andal dan dapat digeneralisasi.

REFERENSI

Bunge, E. M., Hoet, B., Chen, L., Lienert, F., Weidenthaler, H., Baer, L. R., & Steffen, R. (2022). The changing epidemiology of human monkeypox—A potential threat? A systematic review. *PLOS Neglected Tropical Diseases*, *16*(2), e0010141. https://doi.org/10.1371/JOURNAL.PNTD.0010141

Dafid, A., Sudianto, A. I., Thinakaran, R., Umam, F., Adiputra, F., Debora, R. S., Madura, T., Raya, J., Po Box, T., & Kamal, I. (2025). Optimizing K-Nearest Neighbors with *Particle Swarm Optimization* for Improved Classification Accuracy Corresponding Author. *Jurnal Ilmiah Teknik Elektro Komputer Dan Informatika (JITEKI)*, *11*(2), 238–250. https://doi.org/10.26555/jiteki.v11i2.30775

Deng, J., Liu, J., Kong, C., Zang, B., Hu, Y., & Zou, M. (2024). Using novel deep learning models for rapid and efficient assistance in monkeypox screening from skin images. *Frontiers in Medicine*, *11*, 1443812. https://doi.org/10.3389/FMED.2024.1443812/TEXT

Farzipour, A., Elmi, R., & Nasiri, H. (2023). Detection of Monkeypox Cases Based on Symptoms Using XGBoost and Shapley Additive Explanations Methods. *Diagnostics 2023, Vol. 13, Page 2391*, *13*(14), 2391. https://doi.org/10.3390/DIAGNOSTICS13142391

Ileberi, E., & Sun, Y. (2024). Machine Learning-Assisted Cervical Cancer Prediction Using *Particle Swarm Optimization* for Improved Feature Selection and Prediction. *IEEE Access*, *12*, 152684–152695. https://doi.org/10.1109/ACCESS.2024.3469869

Lahariya, C., Thakur, A., & Dudeja, N. (n.d.). Monkeypox Disease Outbreak (2022): Epidemiology, Challenges, and the Way Forward. *INDIAN PEDIATRICS*, *636*, 2022.

Mohammed, H. M., Umar, S. U., & Rashid, T. A. (2019). A Systematic and Meta-Analysis Survey of Whale Optimization Algorithm. *Computational Intelligence and Neuroscience*, *2019*. https://doi.org/10.1155/2019/8718571

Rainio, O., Teuho, J., & Klén, R. (2024). Evaluation metrics and statistical tests for machine learning. *Scientific Reports 2024 14:1*, *14*(1), 6086-. https://doi.org/10.1038/s41598-024-56706-x

Sanchez Clemente, N., Coles, C., Paixao, E. S., Brickley, E. B., Whittaker, E., Alfven, T., Rulisa, S., Agudelo Higuita, N., Torpiano, P., Agravat, P., Thorley, E. V., Drysdale, S. B., Le Doare, K., & Muyembe Tamfum, J. J. (2024). Paediatric, maternal, and congenital mpox: a systematic review and meta-analysis. *The Lancet. Global Health*, *12*(4), e572. https://doi.org/10.1016/S2214-109X(23)00607-1.

Sengupta, S., Basak, S., & Peters, R. A. (2019). Particle Swarm Optimization: A survey of historical and recent developments with hybridization perspectives. *Machine Learning and Knowledge Extraction*, *1*(1), 157–191. https://doi.org/10.3390/make1010010.

Shami, T. M., El-Saleh, A. A., Alswaitti, M., Al-Tashi, Q., Summakieh, M. A., & Mirjalili, S. (2022a). Particle Swarm Optimization: A Comprehensive Survey. *IEEE Access*, *10*, 10031–10061. https://doi.org/10.1109/ACCESS.2022.3142859.

Shami, T. M., El-Saleh, A. A., Alswaitti, M., Al-Tashi, Q., Summakieh, M. A., & Mirjalili, S. (2022b). Particle Swarm Optimization: A Comprehensive Survey. *IEEE Access*, *10*, 10031–10061. https://doi.org/10.1109/ACCESS.2022.3142859.

Sidey-Gibbons, J. A. M., & Sidey-Gibbons, C. J. (2019). Machine learning in medicine: a practical introduction. *BMC Medical Research Methodology*, *19*(1). https://doi.org/10.1186/S12874-019-0681-4.

Uddin, S., Khan, A., Hossain, M. E., & Moni, M. A. (2019). Comparing different supervised machine learning algorithms for disease prediction. *BMC Medical Informatics and Decision Making*, *19*(1). https://doi.org/10.1186/S12911-019-1004-8.

Uzun Ozsahin, D., Mustapha, M. T., Uzun, B., Duwa, B., & Ozsahin, I. (2023a). Computer-Aided Detection and Classification of Monkeypox and Chickenpox Lesion in Human Subjects Using Deep Learning Framework. *Diagnostics (Basel, Switzerland)*, *13*(2). https://doi.org/10.3390/DIAGNOSTICS13020292.

Uzun Ozsahin, D., Mustapha, M. T., Uzun, B., Duwa, B., & Ozsahin, I. (2023b). Computer-Aided Detection and Classification of Monkeypox and Chickenpox Lesion in Human Subjects Using Deep Learning Framework. *Diagnostics*, *13*(2). https://doi.org/10.3390/DIAGNOSTICS13020292.

Xiao, Y., Li, Y., Zhao, Q., & Tang, Y. (2026). ϵ-Neighborhood Model: A classifier based on the k-Nearest Neighbors algorithm and semantic cell models. *Knowledge-Based Systems*, *338*. https://doi.org/10.1016/j.knosys.2026.115518.

Zhang, Z. (2016). Introduction to machine learning: k-nearest neighbors. *Annals of Translational Medicine*, *4*(11). https://doi.org/10.21037/ATM.2016.03.37.