

# Evaluasi Komparatif DFFNN, DFFNN - GA, DFFNN - GWO, dan LSTM untuk Prediksi Konsumsi Energi Listrik dengan *Sliding Window Cross-Validation*

**Giovanni Dimas Prenata**

Program Studi Teknik Elektro,  
Fakultas Teknologi Elektro dan Informatika Cerdas,  
Universitas 17 Agustus 1945 Surabaya, Surabaya  
gprenata@untag-sby.ac.id

## Ringkasan

Prediksi konsumsi energi listrik yang akurat menjadi tantangan penting dalam sistem tenaga modern karena pola beban yang dinamis dan dipengaruhi berbagai faktor, seperti pertumbuhan konsumsi, perubahan perilaku pengguna, serta kondisi eksternal. Pendekatan konvensional seperti regresi dan ANN berbasis *backpropagation* sering menghadapi keterbatasan dalam hal konvergensi dan kemampuan generalisasi terhadap data masa depan. Oleh karena itu, diperlukan pendekatan yang tidak hanya mampu meningkatkan akurasi, tetapi juga stabil dalam menghadapi variasi *temporal data*. Dalam penelitian ini dilakukan evaluasi komparatif empat metode, yaitu *deep feedforward neural network* (DFFNN), DFFNN dengan *genetic algorithm* (GA), DFFNN dengan *grey wolf optimizer* (GWO), serta *long short-term memory* (LSTM), dengan menggunakan teknik *sliding window cross-validation* (SWCV). Tiga skenario evaluasi diterapkan untuk menguji kemampuan generalisasi model terhadap data masa depan. Hasil menunjukkan bahwa DFFNN-GWO unggul dalam efisiensi dengan rata-rata iterasi kurang dari 1.000 dan MSE terbaik mencapai  $6,0 \times 10^{-5}$ , sedangkan LSTM menunjukkan kestabilan prediksi dengan akurasi uji di atas 0,90. DFFNN-GA menghasilkan akurasi tinggi pada beberapa *run* namun tidak konsisten, sementara DFFNN standar berfungsi sebagai *baseline* dengan performa stabil tetapi membutuhkan iterasi lebih besar. Secara keseluruhan, kombinasi optimisasi metaheuristik dan validasi temporal terbukti meningkatkan kinerja prediksi konsumsi energi listrik, dengan GWO unggul dalam efisiensi dan LSTM dalam kestabilan temporal.

**Kata kunci:** DFFNN, *Genetic algorithm*, *grey wolf optimizer*, LSTM, prediksi energi listrik, *sliding window cross-validation*

## Abstract

Accurate prediction of electrical energy consumption remains a critical challenge in modern power systems due to dynamic load patterns influenced by factors such as consumption growth, changes in user behavior, and external conditions. Conventional approaches, including regression models and Artificial Neural Networks (ANN) trained using backpropagation, often face limitations in convergence and generalization to unseen future data. Therefore, there is a need for approaches that not only improve prediction accuracy but also maintain stability under temporal data variations. This study presents a comparative evaluation of four methods: Deep Feedforward Neural Network (DFFNN),

Genetic Algorithm-optimized DFFNN (DFFNN-GA), Grey Wolf Optimizer-optimized DFFNN (DFFNN-GWO), and Long Short-Term Memory (LSTM), utilizing the Sliding Window Cross-Validation (SWCV) technique. Three evaluation scenarios were designed to assess the models' generalization capability on future data. The results indicate that DFFNN-GWO achieves superior efficiency, requiring fewer than 1,000 iterations on average, with the best Mean Squared Error (MSE) reaching  $6.0 \times 10^{-5}$ . Meanwhile, LSTM demonstrates strong predictive stability, achieving test accuracy above 0.90. Although DFFNN-GA occasionally produces highly accurate predictions, its performance is inconsistent across runs. In contrast, the standard DFFNN serves as a stable baseline but requires significantly higher computational effort. Overall, the combination of metaheuristic optimization and temporal validation is proven to enhance the performance of electrical energy consumption forecasting, with GWO excelling in computational efficiency and LSTM in temporal stability.

**Keywords:** Electrical energy consumption, DFFNN, genetic algorithm, grey wolf optimizer, lstm, sliding window cross-validation.

## 1. Pendahuluan

Konsumsi energi listrik merupakan elemen penting dalam pembangunan berkelanjutan dan pertumbuhan ekonomi modern. Di Indonesia, permintaan energi listrik terus meningkat seiring dengan pertumbuhan jumlah penduduk, peningkatan aktivitas industri, urbanisasi, serta meningkatnya penggunaan peralatan elektronik rumah tangga. Selain itu, faktor eksternal, seperti kondisi ekonomi, kebijakan tarif listrik, serta perkembangan teknologi juga turut mempengaruhi pola konsumsi energi listrik. Kondisi ini menuntut sistem tenaga listrik untuk mampu menyediakan pasokan yang andal sekaligus efisien dalam pengelolaan beban, khususnya dalam mengantisipasi fluktuasi beban puncak [1].

Model-model tradisional, seperti ARIMA, regresi linier, dan *artificial neural network* (ANN) berbasis backpropagation telah banyak digunakan dalam prediksi konsumsi energi listrik. Sebagai contoh, penelitian pada sistem distribusi PLN menunjukkan bahwa ANN mampu memodelkan pola konsumsi energi secara cukup baik pada jangka pendek [2]. Selain itu, model *long short-term memory* (LSTM) juga telah banyak digunakan karena kemampuannya dalam menangkap dependensi *temporal* pada data deret waktu [3]. Namun, model-model tersebut masih menghadapi keterbatasan, terutama dalam hal konvergensi pelatihan dan kemampuan generalisasi terhadap data yang belum pernah dilihat sebelumnya.

Di sisi lain, penggunaan optimisasi bobot eksternal untuk meningkatkan performa ANN masih relatif terbatas dalam penelitian Indonesia. Faktor-faktor, seperti *setting hyperparameter*, data latih yang terbatas, dan cara evaluasi model sering berpengaruh besar pada hasil prediksi. Sebagai contoh, penelitian untuk memprediksi konsumsi energi listrik di Kabupaten Kebumen tahun 2023 menunjukkan bahwa model LSTM dapat mencapai MAPE sebesar ~4,07% dengan data historis 2018-2023, namun evaluasi dilakukan dalam kerangka pembagian data latih-uji sederhana (*train-test split*), tidak menggunakan validasi temporal (*sliding window*), yang dapat menyembunyikan kelemahan generalisasi terhadap data masa depan [4].

Masalah generalisasi ini sangat penting karena pola konsumsi energi sering berubah dari tahun ke tahun dipengaruhi oleh banyak faktor eksternal, termasuk kebijakan tarif, perilaku konsumen, kondisi ekonomi, dan gangguan tak terduga seperti cuaca atau pandemi. Apabila evaluasi hanya berdasarkan data historis yang telah diketahui, model mungkin tampil baik pada data latih tetapi performanya menurun ketika menghadapi

data baru. Karena itu, teknik evaluasi yang lebih realistis dan sesuai sifat deret waktu, seperti *sliding window cross-validation* (SWCV), menjadi sangat diperlukan [5]. Teknik ini memungkinkan model diuji pada berbagai jendela temporal sehingga kinerja prediksi pada data masa depan dapat dievaluasi lebih baik.

Berdasarkan kondisi tersebut, penelitian ini bertujuan untuk melakukan evaluasi komparatif terhadap empat metode prediksi konsumsi energi listrik, yaitu *deep feedforward neural network* (DFFNN), DFFNN yang dioptimasi menggunakan *genetic algorithm* (GA), DFFNN yang dioptimasi menggunakan *grey wolf optimizer* (GWO), serta LSTM. Kontribusi utama penelitian ini adalah memberikan analisis komprehensif terkait perbandingan akurasi, efisiensi komputasi, dan stabilitas model dengan menggunakan pendekatan SWCV, yang lebih representatif untuk data deret waktu dibandingkan metode validasi konvensional. Teknik SWCV digunakan untuk mempertahankan urutan temporal data dalam proses evaluasi, sehingga kemampuan generalisasi model terhadap data masa depan dapat diuji secara lebih realistis [16]. Dengan pendekatan ini, penelitian tidak hanya berfokus pada akurasi prediksi, tetapi juga pada konsistensi performa model pada berbagai skenario data.

Tabel 1. Perbandingan penelitian terdahulu pada prediksi konsumsi energi listrik

No	Peneliti	Metode	Data/Tahun	Evaluasi	Hasil Utama	Kelemahan
1	Saputra et al. [2]	ANN ( <i>Backpropagation</i> )	Data PLN Salatiga	<i>Train-test split</i>	Akurasi cukup baik	Tidak mempertimbangkan validasi temporal
2	Dewi & Riwurohi [3]	LSTM	2012–2021	<i>Train-test split</i>	Performa tinggi	Tidak diuji generalisasi jangka waktu
3	Fikriaziz et al. [4]	LSTM	2018–2023	MAPE $\approx$ 4,07%	Akurat pada data historis	Tidak menggunakan SWCV
4	Madrid & Antonio [5]	ML (general)	<i>Time series</i> listrik	<i>Cross-validation</i>	Model ML efektif	Tidak fokus pada optimasi metaheuristik
5	Singh [6]	ANN vs GA	<i>Load forecasting</i>	Perbandingan metode	GA meningkatkan akurasi	Tidak diuji pada validasi temporal
6	Hafeez et al. [7]	<i>Deep Learning + Heuristic</i>	<i>Smart grid</i>	Eksperimen DL	Akurasi meningkat	Kompleksitas tinggi
7	Shanmugam [8]	ANN + <i>Metaheuristic</i>	<i>Power consumption</i>	<i>Multi-model</i>	<i>Error</i> menurun	Tidak dibandingkan dengan LSTM
8	Viegas et al. [9]	GA-ANN	<i>Load forecasting</i>	Eksperimen GA	Optimasi berhasil	Tidak ada GWO dan SWCV

Peramalan konsumsi energi listrik (*electric load forecasting*) berperan penting dalam perencanaan operasi sistem tenaga, penjadwalan pembangkit, dan manajemen permintaan. Pada dekade terakhir, pendekatan berbasis jaringan saraf tiruan (ANN) menjadi fondasi kuat untuk memodelkan relasi nonlinier antara beban dan variabel pendukung, serta menunjukkan kinerja yang kompetitif untuk peramalan jangka pendek [6]. Sejalan dengan kemajuan komputasi, muncul kelas model *deep learning* termasuk arsitektur berurutan yang mampu merepresentasikan pola temporal dan musiman secara lebih kaya pada data deret waktu energi [7]. Di sisi lain, berbagai studi juga menekankan bahwa kualitas pelatihan model sangat dipengaruhi oleh inisialisasi bobot dan parameter tambahan sehingga integrasi teknik optimasi heuristik/metaheuristik (misalnya GA atau

GWO) kerap diadopsi untuk meningkatkan konvergensi dan mengurangi risiko terjebak pada local minima [7] [8].

Literatur terkini menunjukkan dua arah penguatan kinerja yang saling melengkapi. Pertama, pemanfaatan arsitektur *deep learning* (misalnya jaringan berurutan) untuk menangkap dependensi jangka pendek–menengah pada deret waktu beban [7]. Kedua, penggabungan model ANN/LSTM dengan optimasi metaheuristik (misalnya GA, GWO) untuk menata bobot/parameter sehingga *loss* menurun lebih cepat dan stabil [8]. Studi pendekatan hibrida dapat menurunkan galat (MSE/MAE/MAPE) dibanding pelatihan gradien murni atau model konvensional, khususnya ketika data menunjukkan nonstasioneritas dan variasi musiman [7] [8] [9]. Namun demikian, penelitian komparatif yang mengevaluasi DFFNN murni, DFFNN+GA, DFFNN+GWO, dan LSTM secara berdampingan dengan skema validasi temporal yang ketat masih relatif terbatas.

## 2. Tinjauan Pustaka

Metode-metode yang dipilih dalam penelitian ini merepresentasikan beberapa karakteristik penting dalam prediksi konsumsi energi listrik. DFFNN dipilih sebagai model dasar karena arsitektur ini sederhana, umum digunakan, dan mampu memodelkan hubungan nonlinier pada data beban listrik. GA dipilih karena memiliki kemampuan eksplorasi global yang baik dalam mengoptimasi bobot jaringan, sehingga berpotensi mengurangi masalah local minima pada pelatihan ANN. GWO dipilih karena dikenal memiliki keseimbangan yang baik antara eksplorasi dan eksploitasi, sehingga secara teoritis dapat memberikan efisiensi konvergensi yang lebih baik dibanding metode optimasi lainnya. Sementara itu, LSTM dipilih karena dirancang khusus untuk menangani data deret waktu dan memiliki kemampuan menyimpan ketergantungan temporal yang relevan pada pola konsumsi energi listrik. Selain itu, penggunaan SWCV dalam penelitian ini memberikan nilai tambah analitis karena evaluasi tidak hanya dilakukan pada satu pembagian data latih-uji, tetapi pada beberapa jendela waktu yang bergeser.

### 2.1. Jaringan Syaraf Tiruan

Jaringan syaraf tiruan (JST) merupakan model komputasi yang terinspirasi dari struktur dan cara kerja otak manusia. Salah satu arsitektur paling sederhana adalah DFFNN, yang terdiri dari lapisan *input*, lapisan tersembunyi, dan lapisan *output*. Informasi diproses secara searah (*feedforward*) tanpa adanya umpan balik [10]. Secara matematis, keluaran dari sebuah neuron pada lapisan tersembunyi dapat dinyatakan sebagai [10]:

$$y = f(\sum_{i=1}^n w_i x_i + b) \quad (1)$$

dengan  $x_i$  adalah *input*,  $w_i$  bobot,  $b$  bias, dan  $f$  fungsi aktivasi (sigmoid, ReLU, dan lainnya). Pada tahap pelatihan, DFFNN meminimalkan *error* dengan fungsi objektif berupa *mean squared error* (MSE) [10]:

$$MSE = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2 \quad (2)$$

dimana  $t_i$  adalah target dan  $y_i$  adalah keluaran model. DFFNN sederhana tetapi efektif dalam memodelkan hubungan *non-linear*, sehingga banyak dijadikan *baseline* dalam penelitian prediksi konsumsi energi listrik.

## 2.2. Algoritma Genetika

Algoritma Genetika (GA) merupakan algoritma optimasi berbasis evolusi yang meniru proses seleksi alam [11]. GA menggunakan representasi kromosom sebagai kandidat solusi (dalam konteks ini adalah bobot-bobot jaringan). Proses optimasi dilakukan melalui tahapan: inisialisasi populasi, evaluasi *fitness*, seleksi, *crossover*, dan mutasi. *Fitness* biasanya dirumuskan sebagai kebalikan dari *error* prediksi [11]:

$$F = \frac{1}{1+MSE} \quad (3)$$

Semakin kecil nilai MSE, semakin besar *fitness* yang diperoleh. Keunggulan GA dalam optimasi bobot DFFNN adalah kemampuannya menghindari jebakan *local minima*, yang kerap terjadi pada metode *backpropagation* klasik. Namun, GA dapat membutuhkan jumlah iterasi besar untuk mencapai konvergensi, sehingga efisiensi tergantung pada desain parameter evolusi, seperti probabilitas *crossover* dan mutasi.

## 2.3. Grey Wolf Optimizer

GWO diperkenalkan oleh Mirjalili [12] dengan meniru perilaku berburu kawanan serigala abu-abu. GWO memiliki hierarki  $\alpha$  (*alpha*),  $\beta$  (*beta*),  $\delta$  (*delta*), dan  $\omega$  (*omega*), dimana tiga teratas memandu pergerakan kawanan menuju solusi optimal. Persamaan pembaruan posisi adalah [13]:

$$\vec{D}_\alpha = |\vec{C}_1 \cdot \vec{X}_\alpha - \vec{X}(t)| \quad (4)$$

$$\vec{D}_\beta = |\vec{C}_2 \cdot \vec{X}_\beta - \vec{X}(t)| \quad (5)$$

$$\vec{D}_\delta = |\vec{C}_3 \cdot \vec{X}_\delta - \vec{X}(t)| \quad (6)$$

$$\vec{X}(t+1) = \frac{\vec{X}_1 + \vec{X}_2 + \vec{X}_3}{3} \quad (7)$$

dengan [12]:

$$\vec{X}_1 = \vec{X}_\alpha - \vec{A}_1 \cdot \vec{D}_\alpha \quad (8)$$

$$\vec{X}_2 = \vec{X}_\beta - \vec{A}_2 \cdot \vec{D}_\beta \quad (9)$$

$$\vec{X}_3 = \vec{X}_\delta - \vec{A}_3 \cdot \vec{D}_\delta \quad (10)$$

parameter  $\alpha$  menyusut linier dari 2 ke 0 selama iterasi untuk menyeimbangkan eksplorasi dan eksploitasi.

## 2.4. Long Short-Term Memory

*Recurrent Neural Network* (RNN) memiliki kelemahan berupa hilangnya gradien (*vanishing gradient problem*) ketika menangani data sekuen panjang. Untuk mengatasinya, Hochreiter & Schmidhuber mengembangkan LSTM [14]. Arsitektur LSTM memiliki tiga gerbang utama:

1. *Forget gate* ( $f_t$ ): menentukan informasi yang dibuang,
2. *Input gate* ( $i_t$ ): menentukan informasi baru yang masuk ke *cell state*,
3. *Output gate* ( $O_t$ ): menentukan informasi yang dikeluarkan.

Persamaan LSTM secara matematis dirumuskan oleh Hochreiter dan Schmidhuber [14], serta dikembangkan dalam literature deep learning [15]:

$$f_t = \sigma(W_f \cdot |h_{t-1}, X_t| + b_f) \quad (11)$$

$$i_t = \sigma(W_i \cdot |h_{t-1}, X_t| + b_i) \quad (12)$$

$$\tilde{C}_t = \tanh(W_c \cdot |h_{t-1}, X_t| + b_c) \quad (13)$$

$$C_t = f_t \odot C_{t-1} + i_t \odot \tilde{C}_t \quad (14)$$

$$O_t = \sigma(W_o \cdot |h_{t-1}, X_t| + b_o) \quad (15)$$

$$h_t = o_t \odot \tanh(C_t) \quad (16)$$

dengan  $C_t$  adalah *cell state*,  $h_t$  adalah *hidden state*,  $\sigma$  fungsi sigmoid, dan  $\tanh$  fungsi aktivasi hiperbolik.

### 2.5. Sliding Window Cross-Validation

Validasi silang (*cross-validation*) merupakan teknik umum untuk mengevaluasi performa model. Namun, pada data deret waktu, pembagian data secara acak seperti *k-fold Cross-Validation* konvensional tidak relevan karena dapat merusak sifat kronologis data. Untuk itu, SWCV digunakan sebagai metode evaluasi yang lebih sesuai. Prinsip utama SWCV adalah mempertahankan urutan waktu dengan membagi data latih dan data uji secara bergeser (*rolling*). Sebagai contoh, pada *window* pertama digunakan data tahun 2015-2017 sebagai data latih dan tahun 2018 sebagai data uji, sedangkan pada *window* kedua data latih bergeser menjadi 2016-2018 dengan tahun 2019 sebagai data uji. Dengan cara ini, evaluasi model lebih realistis karena meniru kondisi prediksi aktual dimana data historis digunakan untuk memprediksi periode mendatang [16].

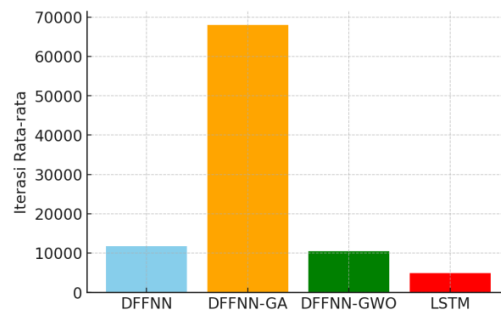
Keunggulan utama SWCV adalah kemampuannya mengurangi bias evaluasi akibat keterbatasan data latih dan memberikan gambaran lebih stabil terhadap performa model jangka panjang. SWCV menekankan bahwa *time series cross-validation* dengan *rolling window* sangat tepat diterapkan pada data energi listrik yang memiliki pola tren dan musiman, karena setiap iterasi menguji kemampuan model untuk generalisasi ke periode yang benar-benar baru [16]. Dengan demikian, SWCV tidak hanya menjaga integritas *temporal data*, tetapi juga memastikan model diuji dalam skenario yang menyerupai aplikasi nyata. Hal ini menjadikan SWCV sebagai pendekatan validasi yang lebih handal dibandingkan metode *holdout* sederhana maupun *k-fold* acak, khususnya dalam penelitian prediksi konsumsi energi listrik.

## 3. Hasil dan Pembahasan

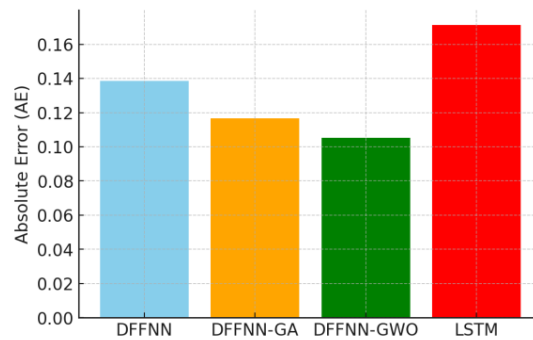
### 3.1. Tanpa Sliding Window Cross-Validation

Pada skenario pertama, seluruh data tahun 2015–2018 digunakan sebagai data latih, sedangkan tahun 2019 sebagai data uji. Kondisi ini memungkinkan model menangkap pola historis secara lebih lengkap. Berdasarkan Gambar 1, terlihat bahwa DFFNN konvensional memiliki rata-rata iterasi maksimum yang tinggi, yaitu di atas 10.000 iterasi, meskipun mampu mencapai nilai MSE mendekati ambang batas 0,0001. Hal ini menunjukkan bahwa DFFNN memiliki akurasi yang baik, namun kurang efisien secara komputasi.

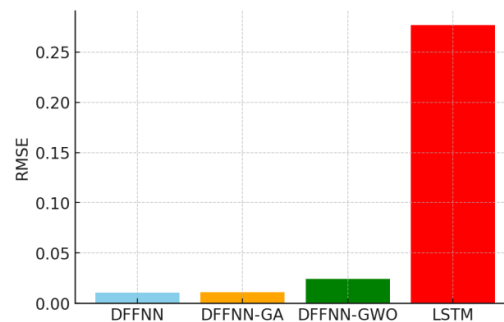
Berdasarkan Gambar 2, terlihat bahwa DFFNN-GA mampu menghasilkan akurasi data uji yang sangat tinggi pada beberapa pengujian (AE mendekati nol). Namun demikian, distribusi hasil yang ditunjukkan pada gambar tersebut juga memperlihatkan variasi yang besar, sehingga menunjukkan bahwa performa DFFNN-GA tidak stabil. Hal ini disebabkan oleh sifat evolusioner GA yang sangat bergantung pada keberagaman populasi dan proses eksplorasi solusi.



Gambar 1. Rata-rata iterasi maksimum per metode (skenario 1)



Gambar 2. Akurasi data uji (AE) per metode (skenario 1)



Gambar 3. Akurasi data latih (RMSE) per metode (skenario 1)

Berdasarkan Gambar 3, DFFNN-GWO menunjukkan performa yang paling efisien, dengan sebagian besar hasil pengujian konvergen hanya dalam ratusan iterasi. Selain itu, akurasi data uji relatif stabil dan mendekati target. Hal ini menunjukkan bahwa mekanisme eksplorasi dan eksploitasi pada DFFNN-GWO mampu menemukan solusi optimal secara lebih konsisten dibandingkan metode lainnya.

Untuk LSTM, hasilnya unik, yaitu MSE stagnan pada kisaran 0,047, yang lebih tinggi dibanding model berbasis DFFNN. Namun, grafik akurasi data uji (AE) menunjukkan kinerja cukup stabil, sebagian besar pengujian menghasilkan prediksi mendekati target (0,9-0,97). Kekuatan LSTM justru tampak pada kestabilan, meskipun tidak pernah benar-benar menurunkan *error* mendekati nol. Dengan jumlah iterasi ribuan, LSTM membutuhkan komputasi lebih panjang daripada GWO, tetapi memberikan generalisasi yang lebih dapat diandalkan pada data runtun waktu.

Berdasarkan Tabel 2, dapat dilihat bahwa DFFNN-GWO memiliki median iterasi yang jauh lebih rendah dibandingkan metode lain, sedangkan LSTM menunjukkan stabilitas *error* yang tinggi dengan variasi yang sangat kecil. Sementara itu, DFFNN-GA

memiliki variasi performa yang besar, dan DFFNN konvensional menunjukkan performa stabil namun dengan biaya komputasi tinggi.

Tabel 2. Ringkasan per metode (skenario 1)

Model	Runs	Median Iterasi	Min. Iterasi	Max. Iterasi	Median MSE	Best (Min) MSE	Worst (Max) MSE	% Lulus Threshold	Threshold
DFNN	12	10555	6848	24955	9,99553E-05	9,85556E-05	9,99877E-05	100	0,0001
DFNN-GA	12	50000	15995	262746	9,35936E-05	7,83293E-05	0,00067769	83,33333333	0,0001
DFNN-GWO	12	2289	78	70868	0,000650368	6,02271E-05	0,00090013	8,333333333	0,0001
LSTM	12	4034	2156	9982	0,0469978	0,0469947	0,0469997	100	0,047

Selain itu, perbedaan performa antar metode dapat dijelaskan dari karakteristik algoritma masing-masing. DFFNN konvensional yang dilatih menggunakan *backpropagation* cenderung mengalami konvergensi lambat karena bergantung pada gradien lokal, sehingga membutuhkan iterasi yang besar untuk mencapai *error* minimum. Pada DFFNN-GA, proses pencarian berbasis evolusi memungkinkan eksplorasi ruang solusi yang lebih luas, namun sifat stokastik dari *crossover* dan mutasi menyebabkan hasil yang tidak konsisten antar pengujian. Sebaliknya, DFFNN-GWO menunjukkan keseimbangan eksplorasi dan eksploitasi yang lebih baik melalui mekanisme pembaruan posisi berbasis hirarki  $\alpha$ ,  $\beta$ , dan  $\delta$  sehingga mampu mencapai konvergensi lebih cepat dengan hasil yang lebih stabil.

Sementara itu, LSTM tidak berfokus pada optimasi bobot secara agresif seperti metode berbasis metaheuristik, tetapi mengandalkan kemampuan memori *temporal* untuk menangkap pola runtun waktu. Hal ini menjelaskan mengapa nilai *error* LSTM relatif stagnan, namun akurasi data uji tetap tinggi dan stabil.

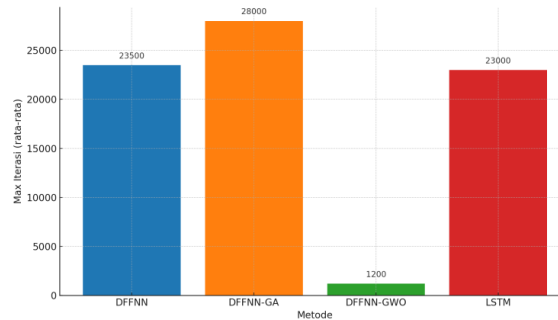
Secara keseluruhan, tabel ringkasan skenario 1 memperlihatkan bahwa DFFNN-GWO unggul dalam efisiensi iterasi, LSTM unggul dalam kestabilan *temporal*, DFFNN-GA menunjukkan potensi solusi optimal tetapi fluktuatif, sedangkan DFFNN konvensional berfungsi sebagai baseline yang stabil namun mahal secara komputasi. Hal ini menegaskan pentingnya kompromi antara akurasi, efisiensi, dan reliabilitas ketika memilih metode.

### 3.2. Sliding Window CV Mode 1

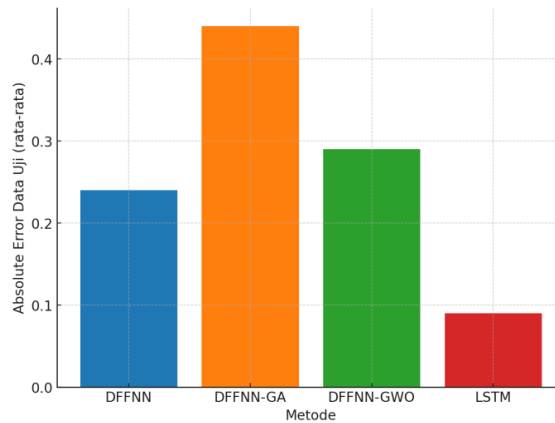
Pada skenario kedua, jumlah data latih berkurang menjadi hanya tiga tahun (2015-2017), dengan 2018 sebagai data uji. Kondisi ini menurunkan kelengkapan pola historis, sehingga risiko *overfitting* meningkat. Dari grafik rata-rata MSE terbaik per metode, terlihat bahwa semua metode mengalami sedikit kenaikan *error* rata-rata dibanding skenario pertama. Namun, grafik akurasi data latih (RMSE) memperlihatkan perbedaan mencolok dimana DFFNN konvensional tetap rendah pada data latih, tetapi AE pada data uji meningkat, menandakan kelemahan generalisasi. Berdasarkan Gambar 4, terlihat bahwa seluruh metode mengalami peningkatan *error* akibat berkurangnya data latih. DFFNN konvensional tetap menunjukkan *error* rendah pada data latih, namun performa pada data uji menurun.

DFNN-GA menunjukkan perbaikan relatif dibanding *baseline*. Beberapa pengujian menghasilkan *output* mendekati target uji, tetapi distribusi MSE yang lebih lebar pada grafik Distribusi MSE mengindikasikan hasil tidak stabil. Persentase hasil pengujian yang

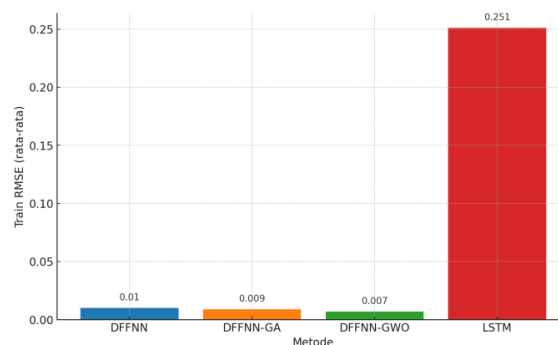
lulus *threshold* MSE pada DFFNN-GA lebih rendah dibanding DFFNN-GWO, memperlihatkan DFFNN-GA masih kesulitan menjaga konsistensi ketika data latih terbatas. Iterasi maksimum pada DFFNN-GA juga relatif tinggi, menandakan pencarian solusi optimal lebih mahal. Berdasarkan Gambar 5, DFFNN-GA menunjukkan hasil yang tidak stabil dengan distribusi *error* yang lebar. Beberapa pengujian menunjukkan performa baik, namun sebagian lainnya gagal mencapai generalisasi yang baik.



Gambar 4. Rata-rata iterasi maksimum per metode (skenario 2)



Gambar 5. Akurasi data uji (AE) per metode (skenario 2)



Gambar 6. Akurasi data latih (RMSE) per metode (skenario 2)

Sebaliknya, DFFNN-GWO tetap menampilkan efisiensi unggul. Grafik Distribusi max. iterasi memperlihatkan sebagian pengujian berhenti hanya dalam ratusan iterasi, dengan rata-rata iterasi maksimum jauh berada di bawah DFFNN dan GA. AE pada data uji lebih konsisten mendekati target, menegaskan kemampuan GWO mempertahankan keseimbangan pencarian solusi meskipun ruang latih lebih sempit. Persentase pengujian lulus *threshold* pada DFFNN-GWO lebih tinggi, menunjukkan reliabilitas kuat.

Berdasarkan Gambar 6, DFFNN-GWO tetap menunjukkan performa yang efisien dan stabil, dengan iterasi yang relatif rendah serta hasil prediksi yang konsisten.

Untuk LSTM, keunggulannya lebih nyata pada skenario ini. Dengan memori jangka panjang, LSTM memanfaatkan keterkaitan *temporal* dari data latih yang terbatas untuk menghasilkan prediksi mendekati target uji. Grafik akurasi data uji (AE) memperlihatkan performa stabil, meski error MSE tetap stagnan di kisaran 0,047. Hal ini menandakan LSTM tidak mampu menekan *error* data latih terlalu rendah, tetapi menjaga prediksi uji tetap realistis dan tidak *overfitted*. Berdasarkan Tabel 3, terlihat bahwa DFFNN-GWO memiliki rata-rata iterasi paling rendah dengan *error* yang relatif kecil, sedangkan LSTM menunjukkan kestabilan pada data uji meskipun *error* latih lebih tinggi.

Tabel 3. Ringkasan per metode (skenario 2)

Metode	<i>n runs</i>	<i>mean iter</i>	<i>std iter</i>	<i>mean mse</i>	<i>std mse</i>	<i>mean train rmse</i>	<i>std train rmse</i>	<i>mean test AE</i>	<i>std test AE</i>
DFNN	12	23446,83	44588,04	0,0001	0	0,009978	0,000029	0,237444	0,015804
DFNN-GA	12	27858,5	50155,84	0,000068	0,000026	0,008012	0,001904	0,437799	0,326306
<b>DFNN-GWO</b>	<b>12</b>	<b>1224,583</b>	<b>1271,586</b>	<b>0,000057</b>	<b>0,000028</b>	<b>0,006513</b>	<b>0,002625</b>	<b>0,286745</b>	<b>0,164155</b>
LSTM	12	23131,17	2227,121	0,047	0	0,251653	0,006089	0,090869	0,014437

Pada skenario 2 terlihat bahwa ketika data latih berkurang, DFFNN-GWO dan LSTM lebih kompetitif dibanding DFFNN-GA. DFFNN-GWO unggul dari sisi efisiensi dan konsistensi, sementara LSTM lebih baik dalam mengatasi keterbatasan *temporal*. DFFNN konvensional kehilangan kekuatan generalisasi, sedangkan DFFNN-GA hanya sesekali menemukan solusi baik tetapi tidak stabil. Jika dibandingkan secara langsung, terdapat perbedaan karakteristik yang jelas antar metode. DFFNN konvensional unggul dalam kesederhanaan dan stabilitas pelatihan, namun kurang efisien dan rentan terhadap *overfitting* ketika data latih terbatas. DFFNN-GA memiliki kemampuan eksplorasi solusi yang luas, tetapi tidak menjamin konsistensi hasil karena sangat bergantung pada parameter evolusi.

Sebaliknya, DFFNN-GWO menunjukkan keunggulan yang lebih seimbang, dimana efisiensi komputasi dan stabilitas hasil dapat dicapai secara bersamaan. Hal ini menjadikan DFFNN-GWO sebagai metode optimasi yang paling robust dalam konteks penelitian ini. Di sisi lain, LSTM menawarkan pendekatan yang berbeda, yaitu dengan memanfaatkan dependensi *temporal* data. Meskipun tidak mencapai *error* minimum seperti metode berbasis optimasi, LSTM memberikan performa prediksi yang lebih stabil pada data uji, terutama pada kondisi data latih terbatas atau berubah secara *temporal*.

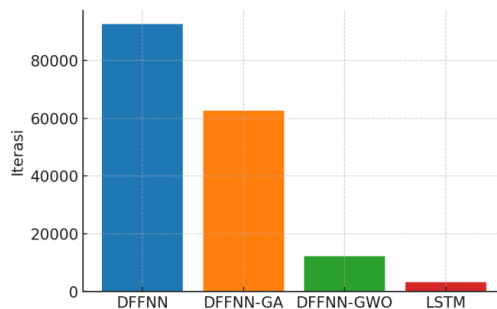
Dengan demikian, tidak terdapat satu metode yang unggul dalam semua aspek. DFFNN-GWO lebih unggul dalam efisiensi, LSTM dalam stabilitas prediksi, DFFNN-GA dalam eksplorasi solusi, dan DFFNN konvensional sebagai *baseline* yang stabil. Hal ini menunjukkan bahwa pemilihan metode harus mempertimbangkan kebutuhan aplikasi, apakah berfokus pada akurasi, efisiensi, atau stabilitas.

### 3.3. Sliding Window CV Mode 2

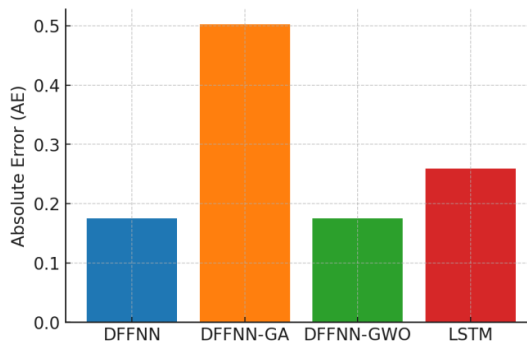
Pada skenario ketiga, data latih diambil dari tahun 2016-2018, dengan 2019 sebagai data uji. Dibanding skenario 2, kedekatan *temporal* data latih dan uji lebih tinggi, sehingga model berpotensi lebih adaptif. Dari grafik Rata-rata MSE Terbaik, terlihat bahwa nilai *error* rata-rata lebih baik dibanding skenario 2. Grafik akurasi data latih

(RMSE) dan akurasi data uji (AE) memperlihatkan tren yang lebih stabil, menunjukkan kedekatan temporal berkontribusi pada generalisasi yang lebih baik. Berdasarkan Gambar 7, terlihat bahwa semua metode mengalami peningkatan performa akibat kedekatan temporal antara data latih dan data uji.

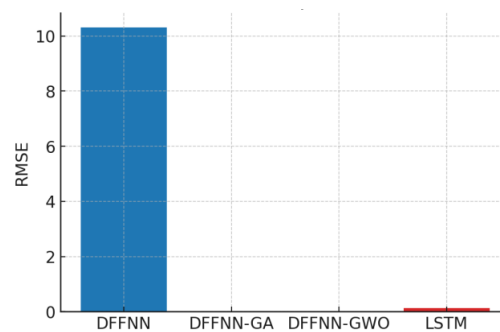
DFFNN konvensional menunjukkan perbaikan pada data uji, dengan AE lebih dekat ke target 1 dibanding skenario sebelumnya. Meski tetap membutuhkan ribuan hingga puluhan ribu iterasi, hasil ini menandakan data yang lebih relevan *temporally* membantu DFFNN konvensional menangkap pola tanpa perlu arsitektur kompleks. Distribusi iterasi tetap lebar, sehingga kelemahan efisiensi tidak hilang. Berdasarkan Gambar 8, DFFNN-GA menunjukkan peningkatan stabilitas dibanding skenario sebelumnya, meskipun variasi masih cukup besar.



Gambar 7. Rata-rata iterasi maksimum per metode (skenario 3)



Gambar 8. Akurasi data uji (AE) per metode (skenario 3)



Gambar 9. Akurasi data latih (RMSE) per metode (skenario 3)

DFFNN-GA pada skenario ini lebih stabil dibanding skenario 2. Beberapa pengujian menghasilkan *output* mendekati target uji dengan AE sangat kecil. Namun, grafik Distribusi MSE menunjukkan variasi masih besar, dan iterasi maksimum tetap tinggi. Persentase pengujian yang lulus *threshold* lebih baik, tetapi belum konsisten mendekati

DFFNN-GWO. DFFNN-GA masih unggul dalam kemampuan eksplorasi dengan harga yang lebih efisien. Berdasarkan Gambar 9, DFFNN-GWO kembali menunjukkan efisiensi terbaik dengan iterasi yang rendah dan hasil yang konsisten.

DFFNN-GWO menunjukkan efisiensi terbaik dengan rata-rata iterasi yang jauh lebih rendah dibanding DFFNN konvensional dan DFFNN-GA, tanpa mengorbankan akurasi data uji. Variasi hasil yang lebih kecil serta persentase kelulusan *threshold* MSE yang lebih tinggi menunjukkan bahwa GWO mampu menghasilkan solusi yang lebih konsisten. Berdasarkan Tabel 4, DFFNN-GWO memiliki rata-rata iterasi terendah, sedangkan LSTM memberikan kestabilan prediksi terbaik dengan nilai AE di atas 0,90.

Tabel 4. Ringkasan per metode (skenario 3)

Model	Rata-rata Iterasi	Min. Iterasi	Max. Iterasi	Mean MSE	Min. MSE	Max. MSE	Model	Rata-rata Iterasi	Min. Iterasi
DFFNN	92,732	3,039	500	0,003872	0,000099	0,022733	DFFNN	92,732	3,039
DFFNN-GA	61,679	18,257	171,85	0,000060	0,000011	0,000090	DFFNN-GA	61,679	18,257
DFFNN-GWO	16,572	774	59,286	0,000068	0,000033	0,000100	DFFNN-GWO	16,572	774
LSTM	3,194	2,25	4,97	0,009997	0,009994	0,009999	LSTM	3,194	2,25

LSTM menunjukkan performa terbaik pada skenario ini dengan nilai AE data uji di atas 0,90. Hal ini menunjukkan bahwa LSTM mampu memanfaatkan kedekatan *temporal* untuk menjaga stabilitas prediksi, meskipun nilai MSE cenderung stagnan pada kisaran 0,047. Dibandingkan metode lain, LSTM tidak seefisien DFFNN-GWO dalam jumlah iterasi, namun lebih unggul dalam konsistensi prediksi pada data runtun waktu.

Secara keseluruhan, kedekatan *temporal* antara data latih dan data uji meningkatkan performa seluruh metode. DFFNN-GWO tetap unggul dalam efisiensi komputasi, sementara LSTM unggul dalam stabilitas *temporal*. DFFNN-GA menunjukkan peningkatan konsistensi dibanding skenario sebelumnya, meskipun masih membutuhkan iterasi yang lebih besar, sedangkan DFFNN konvensional tetap berperan sebagai baseline dengan performa yang cukup baik ketika pola historis relevan dengan data uji.

#### 4. Kesimpulan

Penelitian ini mengevaluasi empat metode untuk memprediksi konsumsi energi Listrik, yaitu DFFNN (konvensional), DFFNN-GA, DFFNN-GWO, dan LSTM dengan pendekatan SWCV. Pada skenario pertama, DFFNN-GWO terbukti paling efisien, mencapai konvergensi hanya dalam ratusan iterasi dengan MSE terbaik  $6,0 \times 10^{-5}$ , sementara LSTM menunjukkan stabilitas prediksi uji dengan akurasi  $>0,90$  meski *error* stagnan di kisaran 0,047. DFFNN-GA kadang menghasilkan solusi sangat akurat (AE  $\approx 0,00$ ), tetapi performanya fluktuatif, sedangkan DFFNN standar tetap akurat namun memerlukan  $>10.000$  iterasi.

Pada skenario kedua, berkurangnya data latih menurunkan stabilitas semua model. DFFNN kehilangan kemampuan generalisasi, dan DFFNN-GA semakin tidak konsisten. DFFNN-GWO mempertahankan efisiensi dengan rata-rata iterasi rendah serta akurasi uji

yang stabil, sementara LSTM menunjukkan keunggulan dalam memanfaatkan pola *temporal*, menghasilkan prediksi mendekati target meskipun data latih terbatas. Hal ini menegaskan bahwa DFFNN-GWO dan LSTM lebih adaptif dibanding metode lain dalam kondisi data sempit.

Skenario ketiga memperlihatkan bahwa kedekatan *temporal* antara data latih dan data uji meningkatkan kinerja semua metode. DFFNN-GWO kembali unggul dalam efisiensi dengan iterasi rata-rata jauh lebih rendah, sedangkan LSTM memberikan prediksi uji terbaik dengan AE umumnya  $>0,90$ . DFFNN-GA lebih stabil dibanding skenario 2 meski tetap memerlukan iterasi besar, sementara DFFNN konvensional menjadi *baseline* yang lebih baik karena data historis lebih relevan. Secara keseluruhan, penelitian ini menegaskan bahwa DFFNN-GWO unggul dari sisi efisiensi, LSTM dari sisi stabilitas *temporal*, sementara GA dan DFFNN tetap memiliki peran sebagai pembanding, serta menekankan pentingnya penggunaan SWCV untuk evaluasi realistis pada data runtun waktu.

Hasil penelitian ini memiliki implikasi praktis dalam perencanaan dan pengelolaan sistem tenaga listrik. Metode DFFNN-GWO yang efisien dapat digunakan pada sistem dengan keterbatasan sumber daya komputasi atau kebutuhan respon cepat, seperti pada peramalan beban jangka pendek di tingkat distribusi. Sementara itu, LSTM yang menunjukkan kestabilan tinggi lebih sesuai untuk aplikasi yang memerlukan konsistensi prediksi dalam jangka waktu tertentu, seperti perencanaan operasional dan manajemen beban. Dengan demikian, pemilihan metode prediksi dapat disesuaikan dengan kebutuhan sistem, apakah lebih menekankan efisiensi komputasi atau stabilitas hasil prediksi.

Untuk pengembangan ke depan, penelitian ini dapat diperluas dengan mengeksplorasi kombinasi metode hybrid yang lebih kompleks, seperti integrasi LSTM dengan algoritma optimasi metaheuristik (misalnya LSTM-GWO), untuk menggabungkan keunggulan stabilitas temporal dan efisiensi optimasi. Selain itu, penggunaan dataset yang lebih besar dan multivariat (misalnya memasukkan variabel cuaca, ekonomi, dan perilaku konsumsi) berpotensi meningkatkan akurasi model. Pengujian dengan metode validasi yang lebih beragam, seperti *multi-step forecasting* atau *probabilistic forecasting*, juga dapat memberikan gambaran yang lebih komprehensif terhadap kinerja model dalam kondisi nyata.

## Daftar Pustaka

- [1] A. F. Insani, "Pemodelan Prediksi Konsumsi Energi Listrik pada Segmen Pelanggan Rumah Tangga Menggunakan Machine Learning : Studi Kasus PT PLN (PERSERO)", Tesis Program Magister, Institut Teknologi Bandung, Desember 2024.
- [2] R. Saputra, S. Sunardiyo, A. Nugroho, and S. Subiyanto, "Implementasi Multilayer Perceptron Artificial Neural Network untuk Prediksi Konsumsi Energi Listrik PT PLN (Persero) UP3 Salatiga," *Elektrika*, vol. 15, no. 2, p. 60, 2023, doi: 10.26623/elektrika.v15i2.6411.
- [3] N. Dewi and J. E. Riwurohi, "Forecasting the Electricity Consumptions of PLN UP3 Cengkareng using Deep Learning," *J. Sisfokom (Sistem Inf. dan Komputer)*, vol. 13, no. 1, pp. 13–20, 2024, doi: 10.32736/sisfokom.v13i1.1849.
- [4] J. Fikriaziz, M. N. Estri, S. Maryani, and I. Sihwaningrum, "Metode Long Short-Term Memory Untuk Memprediksi Konsumsi Energi Listrik Di Kabupaten Kebumen

- Tahun 2023," *Pros. Semin. Nas. Sains Data*, vol. 4, no. 1, pp. 823–831, 2024, doi: 10.33005/senada.v4i1.345.
- [5] E. A. Madrid and N. Antonio, "Short-term electricity load forecasting with machine learning," *Inf.*, vol. 12, no. 2, pp. 1–21, 2021, doi: 10.3390/info12020050.
- [6] P. K. Sarangi, N. Singh, R. K. Chauhan, dan R. Singh, "Short term load forecasting using artificial neural network: A comparison with genetic algorithm implementation," *ARPN Journal of Engineering and Applied Sciences*, vol. 4, no. 9, pp. 88–93, Nov. 2009.
- [7] G. Hafeez, K. Saleem, and I. Khan, "Electric load forecasting based on deep learning and optimized by heuristic algorithm in smart grid," *Appl. Energy*, vol. 269, no. June 2019, p. 114915, 2020, doi: 10.1016/j.apenergy.2020.114915.
- [8] D. Shanmugam, "Different Meta-Heuristic Optimized Radial Basis Function Neural Network Models for Short-Term Power Consumption Forecasting," *Advances in Engineering and Intelligence Systems*, vol. 3, no. 2, pp. 63–82, 2024. doi: 10.22034/aeis.2024.458715.1198
- [9] J. L. Viegas, S. M. Vieira, R. Melício, V. M. F. Mendes, J. M. C Sousa, "GA-ANN Short-Term Electricity Load Forecasting," in L. M. Camarinha-Matos, A. J. Falcão, N. Vafaei, S. Najdi, *Technological Innovation for Cyber-Physical Systems. DoCEIS 2016. IFIP Advances in Information and Communication Technology*, vol 470. Springer, 2016. doi: 10.1007/978-3-319-31165-4\_45
- [10] S. Haykin, *Neural networks and learning machines*, vol. 2. 1999. Upper Saddle River, NJ, USA: Pearson Education, 2009, doi: 10.1109/64.163674.
- [11] P. Waluyo, "Algoritma Genetika: Teori dan Aplikasinya untuk Bisnis dan Industri," 2018. [Online]. Available: [https://www.academia.edu/15288949/ALGORITMA\\_GENETIKA\\_Teori\\_dan\\_Aplikasinya\\_untuk\\_Bisnis\\_dan\\_Industri?bulkDownload=thisPaper-topRelated-sameAuthor-citingThis-citedByThis-secondOrderCitations&from=cover\\_page](https://www.academia.edu/15288949/ALGORITMA_GENETIKA_Teori_dan_Aplikasinya_untuk_Bisnis_dan_Industri?bulkDownload=thisPaper-topRelated-sameAuthor-citingThis-citedByThis-secondOrderCitations&from=cover_page)
- [12] S. Mirjalili, S. M. Mirjalili, and A. Lewis, "Grey Wolf Optimizer," *Adv. Eng. Softw.*, vol. 69, pp. 46–61, 2014, doi: 10.1016/j.advengsoft.2013.12.007.
- [13] Z. Mustafa, Y. Yusof, and S. S. Kamaruddin, "Application of Grey Wolf Optimizer for Time Series Forecasting," in *Proc. Asian Conf. Society, Education and Technology (ACSET)*, 2014, pp. 1–10.
- [14] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, 1997, doi: 10.1162/neco.1997.9.8.1735.
- [15] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*, Cambridge, MA, USA: MIT Press, 2016.
- [16] W. Sulandari, Y. Yudhanto, S. Subanti, E. Zukhronah, and M. Z. Subarkah, "Implementing Time Series Cross Validation to Evaluate the Forecasting Model Performance," *KnE Life Sci.*, vol. 2024, pp. 229–238, 2024, doi: 10.18502/cls.v8i1.15584.