Jurnal Informatika Universitas Pamulang **ISSN: 2541-1004**
Penerbit: Fakultas Ilmu Komputer Universitas Pamulang **e-ISSN: 2622-4615**
Vol. 11, No. 1, March 2026 (25-36) **https://doi.org/10.32493/jiup.v11i1.58137**

Pengaruh *Feature Engineering Time Series* Berbasis *Rolling Statistics* untuk Prediksi Suhu Menggunakan *Random Forest* dan *Gradient Boosting*

Czidni Sika Azkia [a,1*], Yulian Ansori [a,2], Febriyanti Darnis [a,3], Arief Rahman [a,4]
[a] Program Studi Informatika, Fakultas Teknik, Universitas Sultan Ageng Tirtayasa, Cilegon, Indonesia
e-mail: [1]czidni.sika@untirta.ac.id, [2]yulian.ansori@untirta.ac.id, [3]febriyanti.darnis@untirta.ac.id, [4]arief.rahman@untirta.ac.id
*Corresponding author

Submitted Date: January 27th, 2026 Reviewed Date: February 25th, 2026
Revised Date: March 19th, 2026 Accepted Date: March 28th, 2026

Abstract

Temperature data are a form of time series characterized by autocorrelation, seasonality, and nonlinear fluctuations, thus requiring appropriate temporal feature representation to achieve accurate predictions. This study aims to analyze the effect of rolling statistics–based feature engineering on temperature prediction performance using the multivariate Jena Climate dataset. Temporal features are constructed through a combination of lag features and rolling statistics (rolling mean and rolling standard deviation) with various window sizes. Prediction models are developed using Random Forest as the primary model and Gradient Boosting as a comparative model, evaluated using a walk-forward validation scheme and assessed using MAE and RMSE. The results show that the application of rolling statistics produces relatively stable performance but does not lead to a significant improvement in prediction accuracy. In the Random Forest model, the MAE increases from 0.0214 to 0.0216 and the RMSE from 0.1277 to 0.1295 after incorporating rolling features. These findings indicate that rolling statistics play a greater role in maintaining the stability of feature representation, while their contribution to improving predictive accuracy remains limited.

Keywords: Rolling Statistics Feature Engineering; Random Forest Regression Model; Air Temperature Prediction; Time Series Temperature Prediction; Walk-Forward Validation Method

Abstrak

Data suhu merupakan data deret waktu dengan karakteristik autokorelasi, musiman, dan fluktuasi nonlinier, sehingga memerlukan representasi fitur temporal yang tepat untuk menghasilkan prediksi yang akurat. Penelitian ini bertujuan untuk menganalisis pengaruh *feature engineering* berbasis *rolling statistics* terhadap kinerja prediksi suhu menggunakan dataset cuaca multivariat Jena Climate. Fitur temporal dibentuk melalui kombinasi *lag features* dan *rolling statistics* (*rolling mean* dan *rolling standard deviation*) dengan variasi ukuran jendela waktu. Model prediksi dibangun menggunakan *Random Forest* sebagai model utama dan *Gradient Boosting* sebagai pembanding dengan skema *walk-forward validation*, serta diukur menggunakan MAE dan RMSE. Hasil penelitian menunjukkan bahwa penerapan *rolling statistics* menghasilkan kinerja yang relatif stabil, namun tidak memberikan peningkatan akurasi yang signifikan. Pada model *Random Forest*, nilai MAE meningkat dari 0,0214 menjadi 0,0216 dan RMSE dari 0,1277 menjadi 0,1295 setelah penambahan fitur *rolling*. Temuan ini mengindikasikan bahwa *rolling statistics* lebih berperan dalam menjaga kestabilan representasi fitur, namun kontribusinya terhadap peningkatan akurasi prediksi masih terbatas.

Kata Kunci: Feature Engineering Rolling Statistics; Model Random Forest Regresi; Prediksi Suhu Udara; Prediksi Suhu Deret Waktu; Metode Walk-Forward Validation

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0) License
Copyright © 2026 Czidni Sika Azkia, Yulian Ansori, Febriyanti Darnis, Arief Rahman

1. Pendahuluan

Pemodelan prediksi suhu berbasis data menjadi kebutuhan penting seiring meningkatnya ketergantungan berbagai sektor terhadap informasi meteorologis yang akurat. Prediksi suhu yang andal berperan dalam perencanaan energi, pertanian, serta mitigasi dampak kesehatan akibat cuaca ekstrem. Oleh karena itu, peningkatan akurasi prediksi suhu menjadi aspek krusial dalam mendukung pengambilan keputusan berbasis data. Perkembangan pendekatan *machine learning* telah menunjukkan kemampuan dalam menangkap hubungan nonlinier dan kompleksitas data cuaca secara lebih efektif dibandingkan metode konvensional [1]. Namun, permasalahan utama dalam prediksi suhu berbasis *machine learning* terletak pada keterbatasan model dalam merepresentasikan pola temporal secara optimal tanpa dukungan *feature engineering* yang tepat. Model berbasis tabular tidak secara langsung memahami struktur deret waktu, sehingga informasi historis belum dimanfaatkan secara maksimal. Kondisi tersebut menyebabkan akurasi prediksi belum optimal, terutama pada data dengan karakteristik musiman dan fluktuasi nonlinier [2].

Berbagai metode telah digunakan dalam prediksi suhu, mulai dari pendekatan statistik klasik hingga *machine learning* dan *deep learning*. Model seperti ARIMA, *Random Forest*, *Gradient Boosting*, serta *Long Short-Term Memory* (LSTM) telah banyak digunakan untuk menangkap pola temporal dalam data suhu [3]. Metode-metode tersebut menunjukkan kemampuan yang baik dalam memodelkan hubungan antarvariabel serta pola musiman pada data deret waktu [4]. Selain itu, pendekatan hybrid dan teknik optimasi juga dikembangkan untuk meningkatkan akurasi prediksi, seperti kombinasi ARIMA dan LSTM yang mampu memodelkan tren linear dan pola nonlinier secara simultan [5]. Penelitian lain menunjukkan bahwa penggunaan fitur berbasis jendela waktu dan *multi-scale feature* dapat membantu model dalam menangkap pola temporal pada berbagai rentang waktu [6]. Di antara berbagai metode tersebut, pendekatan berbasis *ensemble learning* seperti *Random Forest* dan *Gradient Boosting* banyak digunakan karena mampu meningkatkan stabilitas dan akurasi melalui kombinasi beberapa model [7]. Selain itu, model *machine learning* tersebut relatif lebih fleksibel dalam menangani data multivariat dibandingkan metode statistik konvensional [8]. Meskipun demikian, masing-masing metode memiliki kelebihan dan keterbatasan.

Model *deep learning* seperti LSTM mampu menangkap dependensi jangka panjang, namun membutuhkan data dalam jumlah besar serta komputasi yang tinggi [9]. Sementara itu, model *machine learning* seperti *Random Forest* dan *Gradient Boosting* lebih efisien dan fleksibel dalam menangani data multivariat, tetapi sangat bergantung pada kualitas fitur yang digunakan [10]. Tanpa representasi fitur yang tepat, performa model tidak akan optimal meskipun menggunakan algoritma yang kompleks.

Meskipun berbagai teknik telah dikembangkan, sebagian besar penelitian sebelumnya lebih berfokus pada pengembangan model yang kompleks dibandingkan eksplorasi teknik *feature engineering* sederhana. Secara khusus, evaluasi sistematis terhadap efektivitas *rolling statistics* pada model *ensemble* seperti *Random Forest* dan *Gradient Boosting* masih relatif terbatas dan belum menjadi fokus utama dalam penelitian sebelumnya [11]. Selain itu, pendekatan berbasis *rolling window* diketahui mampu merepresentasikan dinamika lokal data dan membantu mengurangi pengaruh *noise* pada data deret waktu [12]. Namun demikian, pembahasan mengenai kontribusi *rolling statistics* terhadap kestabilan representasi fitur temporal masih belum banyak dikaji secara mendalam.

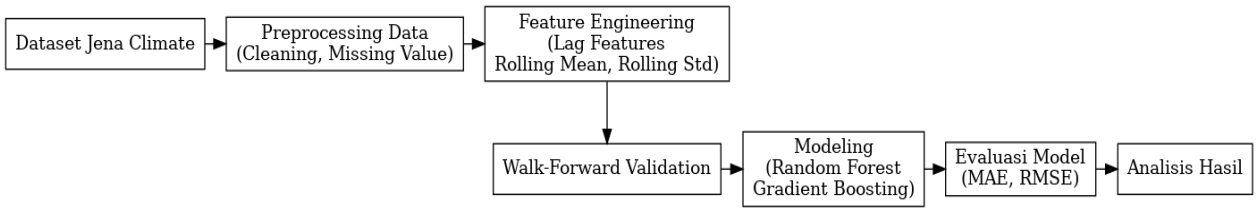
Kondisi ini menunjukkan adanya *research gap* terkait efektivitas *rolling statistics* dalam membantu model memahami pola temporal pada prediksi suhu. Oleh karena itu, penelitian ini bertujuan untuk melakukan evaluasi terstruktur terhadap kontribusi *rolling statistics* dalam meningkatkan representasi fitur temporal pada model *Random Forest* dan *Gradient Boosting*. Penelitian ini menganalisis *feature engineering* berbasis *rolling statistics*, yaitu pembentukan fitur agregat dalam jendela waktu berjalan seperti *rolling mean* dan *rolling standard deviation* [13]. Pendekatan ini memungkinkan model menangkap dinamika lokal dan fluktuasi suhu secara lebih informatif tanpa meningkatkan kompleksitas model secara signifikan. Selain itu, evaluasi dilakukan menggunakan skema *walk-forward validation* untuk menghasilkan pengukuran performa yang lebih realistis dan menghindari bias pada data deret waktu [14].

2. Metode Penelitian

Penelitian ini dirancang sebagai eksperimen terkontrol untuk mengevaluasi pengaruh *feature engineering time series* berbasis *rolling statistics* terhadap akurasi prediksi suhu. Eksperimen dilakukan

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0) License
Copyright © 2026 Czidni Sika Azkia, Yulian Ansori, Febriyanti Darnis, Arief Rahman

dengan membandingkan beberapa konfigurasi fitur temporal, mulai dari *baseline* internal berupa model tanpa penambahan fitur (*lag* dan *rolling statistics*) hingga model dengan kombinasi *lag features* dan *rolling statistics*. Seluruh eksperimen menggunakan algoritma dan skema evaluasi yang sama, sehingga variasi performa yang diperoleh dapat dikaitkan secara langsung dengan perbedaan konfigurasi *feature engineering*. Alur penelitian dimulai dengan pra pemrosesan data, pemodelan, variasi eksperimen hingga evaluasi. Model *baseline* dalam penelitian ini merujuk pada penggunaan model tanpa penambahan fitur temporal (*lag dan rolling statistics*). Secara keseluruhan, terdapat 12 konfigurasi eksperimen yang diuji dengan menggunakan algoritma dan skema evaluasi yang sama untuk menjaga konsistensi perbandingan.Alur penelitian yang digunakan dalam studi ini ditunjukkan pada gambar 1.



Gambar 1. Alur Penelitian

Berdasarkan Gambar 1, proses penelitian diawali dengan pengumpulan dataset Jena Climate, yang kemudian melalui tahap *preprocessing* untuk menangani *missing values* dan memastikan konsistensi data. Selanjutnya dilakukan pembentukan fitur melalui *lag features* dan *rolling statistics* untuk menangkap informasi temporal. Data kemudian digunakan dalam proses pelatihan model menggunakan *Random Forest* dan *Gradient Boosting* dengan skema *walk-forward validation*. Hasil prediksi selanjutnya dievaluasi menggunakan metrik MAE dan RMSE sebelum dilakukan analisis terhadap kinerja model.

a) Dataset dan Persiapan Data

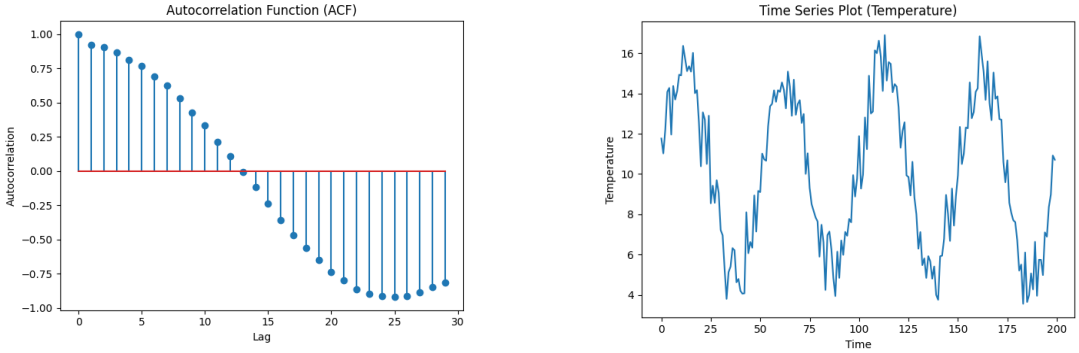
Dataset yang digunakan dalam penelitian ini merupakan data cuaca Jena Climate yang memuat berbagai variabel meteorologis seperti suhu, tekanan udara, dan kelembaban yang dicatat secara berkala, dengan nilai suhu dalam derajat Celcius sebagai target prediksi. Gambar 2 menyajikan contoh data yang digunakan, yang memperlihatkan struktur dataset serta hubungan antar variabel yang membentuk deret waktu (*time series*).

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	Date Time	p (mbar)	T (degC)	Tpot (K)	Tdew (deg	rh (%)	VPmax (ml	VPact (mb	VPdef (mb	sh (g/kg)	H2OC (mm	rho (g/m³	wv (m/s)	max. wv (r	wd (deg)
2	01.01.2009 00:10:00	996.52	-8.02	265.40	-8.90	93.30	3.33	3.11	0.22	1.94	3.12	1307.75	1.03	1.75	152.30
3	01.01.2009 00:20:00	996.57	-8.41	265.01	-9.28	93.40	3.23	3.02	0.21	1.89	3.03	1309.80	0.72	1.50	136.10
4	01.01.2009 00:30:00	996.53	-8.51	264.91	-9.31	93.90	3.21	3.01	0.20	1.88	3.02	1310.24	0.19	0.63	171.60
5	01.01.2009 00:40:00	996.51	-8.31	265.12	-9.07	94.20	3.26	3.07	0.19	1.92	3.08	1309.19	0.34	0.50	198.00
6	01.01.2009 00:50:00	996.51	-8.27	265.15	-9.04	94.10	3.27	3.08	0.19	1.92	3.09	1309.00	0.32	0.63	214.30
7	01.01.2009 01:00:00	996.50	-8.05	265.38	-8.78	94.40	3.33	3.14	0.19	1.96	3.15	1307.86	0.21	0.63	192.70
8	01.01.2009 01:10:00	996.50	-7.62	265.81	-8.30	94.80	3.44	3.26	0.18	2.04	3.27	1305.68	0.18	0.63	166.50
9	01.01.2009 01:20:00	996.50	-7.62	265.81	-8.36	94.40	3.44	3.25	0.19	2.03	3.26	1305.69	0.19	0.50	118.60
10	01.01.2009 01:30:00	996.50	-7.91	265.52	-8.73	93.80	3.36	3.15	0.21	1.97	3.16	1307.17	0.28	0.75	188.50
11	01.01.2009 01:40:00	996.53	-8.43	264.99	-9.34	93.10	3.23	3.00	0.22	1.88	3.02	1309.85	0.59	0.88	185.00
12	01.01.2009 01:50:00	996.62	-8.76	264.66	-9.66	93.10	3.14	2.93	0.22	1.83	2.94	1311.64	0.45	0.88	183.20
13	01.01.2009 02:00:00	996.62	-8.88	264.54	-9.77	93.20	3.12	2.90	0.21	1.81	2.91	1312.25	0.25	0.63	190.30
14	01.01.2009 02:10:00	996.63	-8.85	264.57	-9.70	93.50	3.12	2.92	0.20	1.82	2.93	1312.11	0.16	0.50	158.30
15	01.01.2009 02:20:00	996.74	-8.83	264.58	-9.68	93.50	3.13	2.92	0.20	1.83	2.93	1312.15	0.36	0.63	184.80
16	01.01.2009 02:30:00	996.81	-8.66	264.74	-9.46	93.90	3.17	2.98	0.19	1.86	2.99	1311.37	0.33	0.75	155.90
17	01.01.2009 02:40:00	996.81	-8.66	264.74	-9.50	93.60	3.17	2.97	0.20	1.85	2.98	1311.38	0.07	0.50	272.40
18	01.01.2009 02:50:00	996.86	-8.70	264.70	-9.55	93.50	3.16	2.95	0.21	1.85	2.96	1311.64	0.32	0.63	219.20
19	01.01.2009 03:00:00	996.84	-8.81	264.59	-9.66	93.50	3.13	2.93	0.20	1.83	2.94	1312.18	0.18	0.63	167.20
20	01.01.2009 03:10:00	996.87	-8.84	264.56	-9.69	93.50	3.13	2.92	0.20	1.83	2.93	1312.37	0.07	0.25	129.30
21	01.01.2009 03:20:00	996.97	-8.94	264.45	-9.82	93.30	3.10	2.89	0.21	1.81	2.90	1313.01	0.10	0.63	115.30
22	01.01.2009 03:30:00	997.08	-8.94	264.44	-9.80	93.40	3.10	2.90	0.20	1.81	2.90	1313.15	0.30	0.75	149.30
23	01.01.2009 03:40:00	997.10	-8.86	264.52	-9.76	93.10	3.12	2.90	0.22	1.81	2.91	1312.78	0.29	0.75	149.70
24	01.01.2009 03:50:00	997.06	-8.99	264.39	-9.99	92.40	3.09	2.85	0.23	1.78	2.86	1313.39	0.12	0.63	231.70
25	01.01.2009 04:00:00	996.99	-9.05	264.34	-10.02	92.60	3.07	2.85	0.23	1.78	2.85	1313.61	0.10	0.38	240.00
26	01.01.2009 04:10:00	997.05	-9.23	264.15	-10.25	92.20	3.03	2.79	0.24	1.74	2.80	1314.62	0.10	0.38	203.90
27	01.01.2009 04:20:00	997.11	-9.49	263.89	-10.54	92.00	2.97	2.73	0.24	1.71	2.74	1316.02	0.34	0.75	159.70

Gambar 2. Data Cuaca Jena Climate

Untuk memahami karakteristik data, dilakukan eksplorasi melalui visualisasi *time series* serta analisis autokorelasi (ACF) guna mengidentifikasi pola musiman dan hubungan temporal. Untuk mendukung analisis eksplorasi data yang telah dijelaskan, visualisasi *time series* dan autokorelasi (ACF) ditampilkan pada Gambar 3 sebagai bukti adanya pola musiman dan ketergantungan temporal pada data suhu.

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0) License
Copyright © 2026 Czidni Sika Azkia, Yulian Ansori, Febriyanti Darnis, Arief Rahman



Gambar 3. Visualisasi *Time Series* Suhu dan *Autocorrelation Function* (ACF)

Data suhu menunjukkan fluktuasi nonlinier dengan variasi jangka pendek dan menengah, serta kecenderungan musiman dan autokorelasi yang kuat, sehingga informasi historis memiliki peran penting dalam proses prediksi. Oleh karena itu, pendekatan *feature engineering* berbasis statistik lokal deret waktu menjadi relevan untuk menangkap dinamika suhu secara lebih informatif. Untuk memberikan gambaran kuantitatif, dilakukan analisis statistik deskriptif terhadap variabel suhu pada Tabel 1.

Tabel 1. Deskriptif Variabel Suhu Jena Climate

Statistik	Nilai
Jumlah Data	105.261
Rata-rata (Mean)	8,17 °C
Simpangan Baku (Std)	9,03 °C
Nilai Minimum	−23,01 °C
Kuartil 1 (Q1)	1,71 °C
Median (Q2)	8,56 °C
Kuartil 3 (Q3)	14,64 °C
Nilai Maksimum	34,92 °C

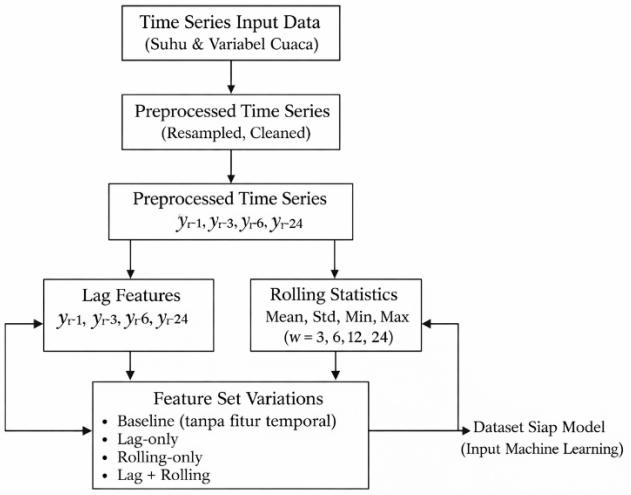
Berdasarkan Tabel 1, nilai rata-rata suhu berada pada kisaran 8,17 °C dengan simpangan baku sebesar 9,03 °C, yang menunjukkan adanya variasi suhu yang cukup tinggi selama periode pengamatan. Rentang nilai suhu yang luas, dari −23,01 °C hingga 34,92 °C, mengindikasikan adanya fluktuasi suhu yang signifikan yang dipengaruhi oleh kondisi musiman dan faktor lingkungan lainnya.

Sebaran nilai suhu yang relatif lebar ini menunjukkan bahwa data memiliki karakteristik nonlinier dan dinamis, sehingga menantang untuk diprediksi menggunakan pendekatan sederhana. Oleh karena itu, penggunaan feature engineering berbasis lag dan rolling statistics menjadi relevan untuk menangkap dinamika temporal suhu secara lebih informatif. Statistik deskriptif ini juga memberikan konteks dalam menginterpretasikan nilai kesalahan prediksi yang dihasilkan oleh model, di mana nilai MAE dan RMSE dapat dipahami secara proporsional terhadap sebaran data suhu yang digunakan.

b) Variasi *Feature Engineering Time Series*

Proses diawali dengan data yang telah melalui tahap pra-pemrosesan untuk memastikan konsistensi interval waktu dan kualitas data. Dari data tersebut, fitur temporal dibentuk melalui lag features, yaitu nilai suhu pada waktu sebelumnya untuk merepresentasikan ketergantungan temporal, serta *rolling statistics*, yang dihitung menggunakan jendela waktu berjalan untuk menangkap dinamika lokal suhu. *Rolling statistics* mencakup agregasi statistik seperti nilai rata-rata, simpangan baku, minimum, dan maksimum pada setiap jendela waktu, sehingga mampu merepresentasikan kecenderungan dan fluktuasi suhu pada skala waktu tertentu.

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0) License
Copyright © 2026 Czidni Sika Azkia, Yulian Ansori, Febriyanti Darnis, Arief Rahman

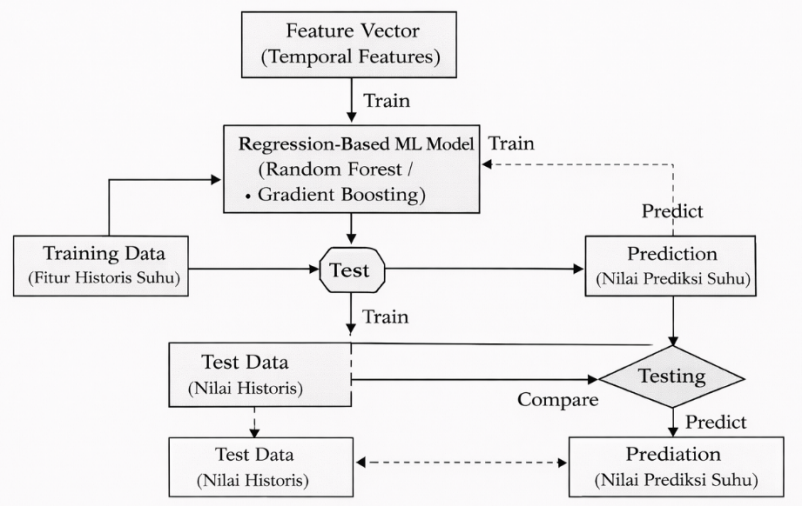


Gambar 4. Alur variasi feature engineering *time series*

Sebagaimana ditunjukkan pada Gambar 4, hasil pembentukan *lag features* dan *rolling statistics* selanjutnya dikombinasikan ke dalam beberapa konfigurasi *feature set*, meliputi *baseline* tanpa fitur temporal tambahan, konfigurasi dengan *lag features* saja, konfigurasi dengan *rolling statistics* saja, serta kombinasi keduanya. Setiap konfigurasi *feature set* digunakan sebagai input model *machine learning* dan dievaluasi secara terpisah menggunakan skema validasi *time series*. Pendekatan ini memungkinkan analisis terstruktur terhadap kontribusi variasi *feature engineering* dalam meningkatkan akurasi prediksi suhu seluruh eksperimen.

c) Pemodelan dan Skema Evaluasi

Pemodelan prediksi suhu berbasis algoritma regresi *machine learning* ditunjukkan pada Gambar 4, yang menggambarkan alur pembelajaran dan prediksi dari fitur temporal menuju nilai keluaran. Vektor fitur hasil *feature engineering time series* digunakan sebagai input model regresi, yang kemudian diproses pada tahap pelatihan untuk mempelajari hubungan antara fitur historis dan nilai suhu target [15]. Model yang digunakan membentuk fungsi pemetaan yang merepresentasikan pola nonlinier antara fitur temporal dan suhu. Hyperparameter model yang digunakan dalam penelitian ini adalah sebagai berikut: *Random Forest* dengan n_estimators = 100 dan max_depth = None, serta *Gradient Boosting* dengan n_estimators = 100, learning_rate = 0.1, dan max_depth = 3.



Gambar 5. Alur kerja algoritma regresi *machine learning*

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0) License
Copyright © 2026 Czidni Sika Azkia, Yulian Ansori, Febriyanti Darnis, Arief Rahman

Berdasarkan Gambar 5, model regresi *machine learning* menghasilkan prediksi suhu satu langkah ke depan sebagai keluaran. Proses pelatihan dan pengujian dilakukan secara berurutan mengikuti skema evaluasi *time series*, sehingga model dilatih menggunakan data historis dan diuji pada periode waktu berikutnya [16]. Hasil prediksi kemudian dibandingkan dengan nilai aktual untuk menghitung metrik kesalahan prediksi, yang digunakan sebagai dasar evaluasi kinerja algoritma regresi pada setiap variasi *feature engineering*. Untuk menghindari potensi data *leakage*, seluruh proses pembentukan fitur dilakukan secara kronologis tanpa menggunakan informasi dari masa depan. Selain itu, skema *walk-forward validation* memastikan bahwa data pelatihan selalu mendahului data pengujian.

d) Variasi Eksperimen dan Metrik Evaluasi

Eksperimen dilakukan menggunakan beberapa konfigurasi *feature set* sebagaimana dirangkum pada Tabel 2. Setiap konfigurasi dievaluasi menggunakan metrik *Mean Absolute Error* (MAE) dan *Root Mean Squared Error* (RMSE) yang dihitung pada setiap iterasi *walk-forward* dan dirata-ratakan untuk memperoleh performa keseluruhan. Perbandingan performa antar konfigurasi digunakan untuk mengidentifikasi pengaruh penambahan *lag features, rolling statistics,* serta variasi ukuran jendela rolling terhadap akurasi prediksi suhu.

Tabel 2. Variasi Konfigurasi Eksperimen

Kode	*Feature Set*	*Lag*	*Rolling*	*Window*	Model	Validasi	Metrik
E1	*Baseline*	Tidak	Tidak	–	RF	*Walk-forward*	MAE, RMSE
E2	*Lag-only*	Ya	Tidak	–	RF	*Walk-forward*	MAE, RMSE
E3	*Rolling-only*	Tidak	Ya	3, 6, 12	RF	*Walk-forward*	MAE, RMSE
E4	*Lag + Rolling*	Ya	Ya	3, 6, 12	RF	*Walk-forward*	MAE, RMSE
E5	*Baseline*	Tidak	Tidak	–	GB	*Walk-forward*	MAE, RMSE
E6	*Lag + Rolling*	Ya	Ya	3, 6, 12	GB	*Walk-forward*	MAE, RMSE
E7	*Baseline*	Tidak	Tidak	–	*Linear Regression*	*Walk-forward*	MAE, RMSE
E8	*Lag + Rolling*	Ya	Ya	3, 6, 12	*Linear Regression*	*Walk-forward*	MAE, RMSE

Berdasarkan Tabel 2, eksperimen dalam penelitian ini dirancang untuk mengevaluasi pengaruh variasi *feature engineering* terhadap kinerja model prediksi suhu. Konfigurasi yang diuji meliputi beberapa skenario, yaitu baseline tanpa fitur temporal, penggunaan *lag features*, *rolling statistics*, serta kombinasi keduanya. Eksplorasi variasi *feature engineering* difokuskan pada model *Random Forest* sebagai model utama, sedangkan *Gradient Boosting* digunakan sebagai model pembanding untuk mengevaluasi konsistensi hasil. Selain itu, regresi linear digunakan sebagai baseline klasik untuk merepresentasikan pendekatan sederhana dalam pemodelan prediksi. Konfigurasi *feature engineering* yang mengombinasikan *lag features* dan *rolling statistics* berpotensi menimbulkan data *leakage* apabila tidak diimplementasikan dengan tepat. Oleh karena itu, dalam penelitian ini seluruh proses pembentukan fitur dilakukan secara kronologis tanpa menggunakan informasi dari masa depan. Selain itu, penerapan skema *walk-forward validation* memastikan bahwa data pelatihan selalu mendahului data pengujian, sehingga integritas evaluasi tetap terjaga. Kinerja model dievaluasi menggunakan metrik *Mean Absolute Error* (MAE) dan *Root Mean Squared Error* (RMSE), sehingga perbandingan antar konfigurasi dapat dilakukan secara objektif dan terukur.

3. Hasil dan Pembahasan

a. Hasil Eksperimen Prediksi Suhu

Eksperimen diperoleh melalui serangkaian pengujian menggunakan model regresi *machine learning* dengan berbagai konfigurasi *feature engineering*. Evaluasi dilakukan pada dataset cuaca yang telah melalui tahap *preprocessing*, meliputi pembersihan data, interpolasi berbasis waktu, serta pemilihan subset data untuk menjaga kestabilan pembelajaran model. Kinerja model dievaluasi menggunakan metrik *Mean*

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0) License
Copyright © 2026 Czidni Sika Azkia, Yulian Ansori, Febriyanti Darnis, Arief Rahman

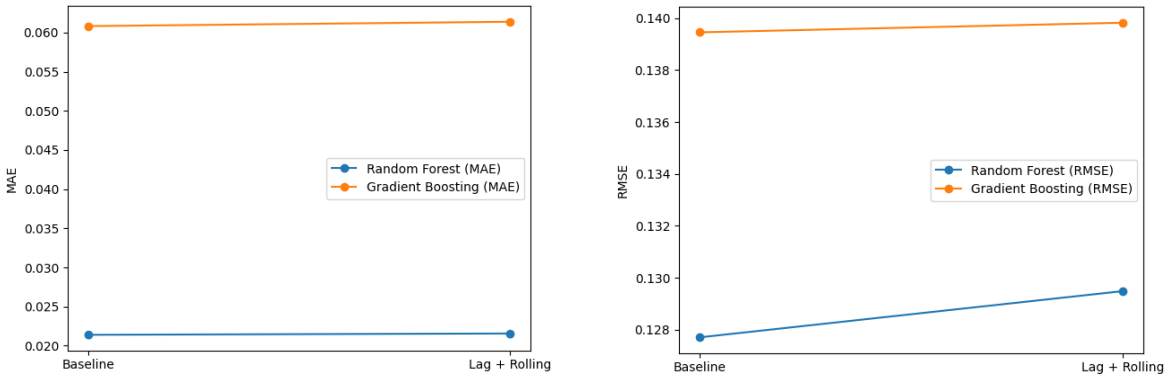
Absolute Error (MAE) dan *Root Mean Squared Error* (RMSE), yang umum digunakan untuk mengukur tingkat kesalahan prediksi pada permasalahan regresi deret waktu [17]. Hasil lengkap dari eksperimen prediksi suhu ditunjukkan pada Tabel 3. Tabel tersebut menyajikan perbandingan kinerja dua model regresi, yaitu *Random Forest* dan *Gradient Boosting* pada skenario baseline serta pada skenario penerapan *feature engineering* berbasis kombinasi *lag* dan *rolling statistics*. Nilai MAE dan RMSE yang lebih rendah menunjukkan kemampuan model yang lebih baik dalam mendekati nilai suhu aktual.

Tabel 3. Hasil Eksperimen Prediksi Suhu

Model	*Feature Set*	*Window*	MAE	RMSE
Random Forest	*Baseline*	–	0.0214	0.1277
Gradient Boosting	*Baseline*	–	0.0608	0.1395
Random Forest	*Lag + Rolling*	3,6,12	0.0216	0.1295
Gradient Boosting	*Lag + Rolling*	3,6,12	0.0614	0.1398

Berdasarkan hasil pada Tabel 3, model *Random Forest* secara konsisten menghasilkan nilai MAE dan RMSE yang lebih rendah dibandingkan *Gradient Boosting*, baik pada skenario baseline maupun dengan *feature engineering*. Hal ini menunjukkan bahwa *Random Forest* lebih efektif dalam menangkap hubungan nonlinier pada data cuaca multivariat. Namun demikian, perbedaan nilai MAE dan RMSE antar konfigurasi relatif kecil, sehingga dapat dikatakan bahwa peningkatan performa yang dihasilkan tidak signifikan secara praktis. Pada model *Random Forest*, nilai MAE meningkat dari 0,0214 menjadi 0,0216 dan RMSE dari 0,1277 menjadi 0,1295 setelah penambahan fitur *rolling*, yang mengindikasikan bahwa penambahan *rolling statistics* tidak selalu memberikan peningkatan kinerja. Tidak adanya peningkatan performa setelah penambahan *rolling statistics* dapat disebabkan oleh karakteristik model *Random Forest* yang secara intrinsik telah mampu menangkap pola nonlinier dan hubungan historis melalui struktur *ensemble decision tree*. Penambahan fitur *rolling* yang bersifat agregatif berpotensi menghasilkan informasi yang redundan terhadap *lag features*, sehingga tidak memberikan kontribusi tambahan yang signifikan. Selain itu, penggunaan *window* yang tidak optimal juga dapat menyebabkan *smoothing* berlebihan, sehingga informasi fluktuasi lokal yang penting justru berkurang.

Kondisi ini dapat disebabkan oleh kemampuan *Random Forest* yang sudah cukup *robust* dalam menangkap pola historis tanpa memerlukan tambahan fitur statistik lokal. Di sisi lain, *rolling statistics* cenderung lebih bermanfaat dalam menjaga kestabilan representasi fitur, terutama pada data dengan fluktuasi tinggi, namun kurang memberikan dampak signifikan terhadap peningkatan akurasi secara langsung.

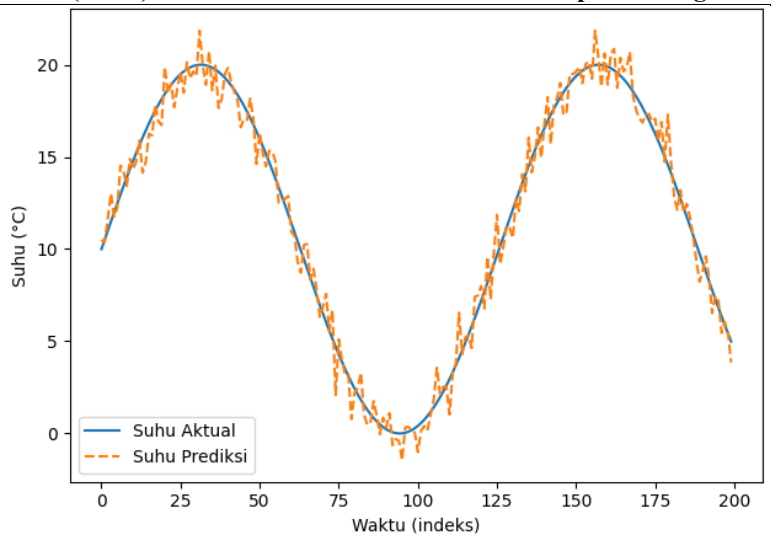


Gambar 6. Perbandingan MAE & RMSE antar Variasi *Feature Engineering*

Pada gambar 6 menunjukkan perbandingan nilai MAE dan RMSE antar variasi *feature engineering* pada model *Random Forest* dan *Gradient Boosting*. Perbandingan ini digunakan untuk meninjau perbedaan kinerja model dalam merepresentasikan pola data, di mana nilai error yang lebih rendah menunjukkan performa yang lebih baik.

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0) License
Copyright © 2026 Czidni Sika Azkia, Yulian Ansori, Febriyanti Darnis, Arief Rahman

ISSN: 2541-1004
e-ISSN: 2622-4615
https://doi.org/10.32493/jiup.v11i1.58137



Gambar 7. Perbandingan Suhu Aktual dan Prediksi

Berdasarkan Gambar 7, kedekatan antara kurva suhu aktual dan prediksi menunjukkan bahwa model mampu menangkap dinamika suhu secara cukup akurat dalam rentang waktu pengujian. Meskipun demikian, masih terdapat deviasi pada beberapa titik perubahan ekstrem, yang mengindikasikan keterbatasan model dalam memodelkan fluktuasi suhu yang sangat cepat.

b. Perbandingan Variasi *Feature Engineering*

Variasi *feature engineering* diterapkan untuk menganalisis perbedaan kinerja model prediksi suhu antara penggunaan fitur asli (*baseline*) dan fitur hasil transformasi temporal yang dibentuk melalui kombinasi *lag features* dan *rolling statistics*.

Tabel 4. Kinerja Model berdasarkan Variasi Feature Engineering

Model	*Baseline*	*Lag + Rolling Statistics*	Tren Kinerja
Random Forest	Error rendah	Error rendah	Konsisten
Gradient Boosting	Error sedang	Error sedang	Konsisten

Berdasarkan Tabel 4, model *Random Forest* menunjukkan tingkat kesalahan prediksi yang rendah baik pada skenario *baseline* maupun pada skenario dengan penerapan kombinasi *lag* dan *rolling statistics*. Konsistensi ini menunjukkan bahwa penambahan fitur temporal tidak menyebabkan degradasi kinerja model, serta mampu mempertahankan kemampuan generalisasi dalam memprediksi suhu. Hal ini mengindikasikan bahwa *Random Forest* cukup adaptif terhadap representasi fitur deret waktu yang lebih kompleks.

Pada model *Gradient Boosting*, kecenderungan kinerja yang relatif serupa juga terlihat pada kedua variasi *feature engineering*, meskipun tingkat kesalahan prediksi berada pada kategori sedang. Temuan ini menunjukkan bahwa penerapan *lag* dan *rolling statistics* tidak memberikan perubahan signifikan terhadap tren kinerja model. Secara keseluruhan, ringkasan pada Tabel 2 menegaskan bahwa variasi *feature engineering* yang diterapkan bersifat stabil dan tidak menimbulkan penurunan performa prediksi, sehingga layak digunakan dalam pemodelan prediksi suhu berbasis *machine learning*.

c. Analisis Pengaruh Rolling Window

Rolling statistics merupakan salah satu teknik *feature engineering* yang umum digunakan dalam pemodelan deret waktu untuk mengekstraksi informasi statistik lokal dari data historis [18]. Pendekatan ini bekerja dengan menerapkan jendela waktu (*rolling window*) yang bergerak mengikuti urutan temporal data, sehingga karakteristik lokal seperti kecenderungan dan variasi nilai dapat direpresentasikan secara eksplisit. Dalam konteks prediksi suhu, *rolling window* memungkinkan model untuk menangkap dinamika

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0) License
Copyright © 2026 Czidni Sika Azkia, Yulian Ansori, Febriyanti Darnis, Arief Rahman

perubahan suhu dalam rentang waktu tertentu tanpa harus bergantung sepenuhnya pada nilai observasi tunggal yang bersifat fluktuatif [11].

Secara matematis, *rolling statistics* dapat direpresentasikan melalui perhitungan nilai rata-rata bergerak (*rolling mean*) dan simpangan baku bergerak (*rolling standard deviation*) [19]. Untuk suatu deret waktu suhu x_t, *rolling mean* dengan ukuran jendela $w$pada waktu ke-t dirumuskan sebagai:

$$\mu_t^{(w)} = \frac{1}{w} \sum_{i=t-w+1}^{t} x_i \quad (1)$$

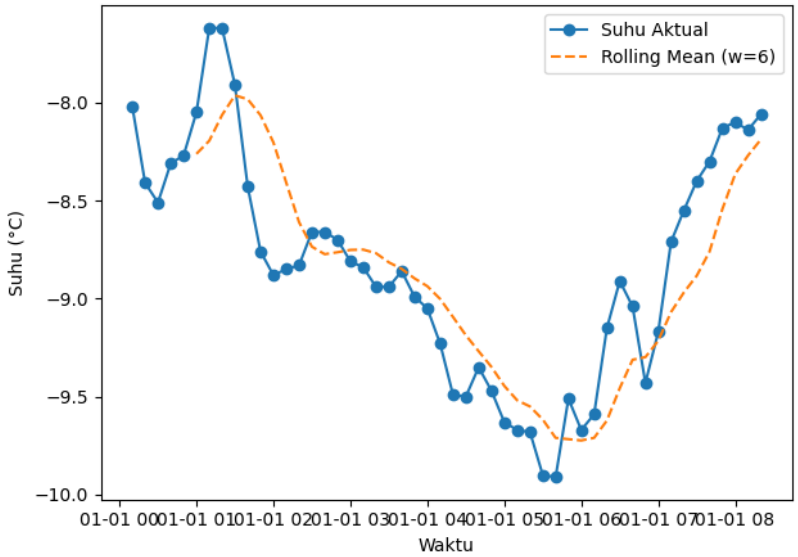
Sementara itu, *rolling standard deviation* dirumuskan sebagai:

$$\sigma_t^{(w)} = \sqrt{\frac{1}{w} \sum_{i=t-w+1}^{t} \left(x_i - \mu_t^{(w)}\right)^2} \quad (2)$$

Rumus tersebut menunjukkan bahwa nilai statistik pada waktu tertentu tidak hanya bergantung pada satu titik data, tetapi pada sekumpulan data historis dalam jendela waktu w. Dengan demikian, *rolling statistics* mampu mereduksi pengaruh *noise* jangka pendek sekaligus mempertahankan informasi tren lokal yang relevan untuk proses pembelajaran model [20]. Implementasi *rolling statistics* pada penelitian ini dilakukan menggunakan bahasa pemrograman Python dengan memanfaatkan pustaka pandas. Contoh implementasi *rolling mean* dan *rolling standard deviation* pada variabel suhu ditunjukkan pada potongan kode berikut:

```
df['T_roll_mean_6'] = df['T (degC)'].rolling(window=6).mean()
df['T_roll_std_6']  = df['T (degC)'].rolling(window=6).std()
```

Fitur hasil *rolling statistics* tersebut kemudian digunakan sebagai fitur tambahan dan dikombinasikan dengan fitur *lag* sebelum dimasukkan ke dalam model regresi *machine learning*. Pendekatan ini memungkinkan model untuk mengakses informasi historis suhu secara terstruktur tanpa memerlukan model deret waktu yang kompleks seperti ARIMA atau model berbasis jaringan saraf. Visualisasi penerapan *rolling window* pada data suhu ditunjukkan pada Gambar 6, yang menampilkan perbandingan antara nilai suhu aktual dan nilai *rolling mean*. Terlihat bahwa *rolling mean* menghasilkan kurva yang lebih halus dibandingkan data asli, sehingga fluktuasi jangka pendek dapat diminimalkan. Hal ini membantu model dalam mempelajari pola temporal yang lebih stabil, terutama ketika data memiliki variasi tinggi akibat faktor lingkungan.



Gambar 8. *Rolling Window* pada Data Suhu

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0) License
Copyright © 2026 Czidni Sika Azkia, Yulian Ansori, Febriyanti Darnis, Arief Rahman

Pada Gambar 8 menunjukkan, variasi ukuran rolling window yang digunakan (3, 6, dan 12) menghasilkan perbedaan tingkat kestabilan prediksi, meskipun tidak menunjukkan perbedaan performa yang signifikan secara kuantitatif. *Rolling window* berukuran kecil (3) cenderung lebih sensitif terhadap perubahan suhu jangka pendek, sehingga mampu menangkap fluktuasi cepat namun berpotensi memasukkan noise. Sebaliknya, *rolling window* berukuran lebih besar (12) menghasilkan representasi yang lebih halus, tetapi cenderung mengurangi sensitivitas terhadap perubahan lokal. Sementara itu, *rolling window* berukuran menengah (6) menunjukkan keseimbangan antara sensitivitas dan kestabilan, sehingga dapat dianggap sebagai konfigurasi yang paling optimal dalam konteks penelitian ini. Secara keseluruhan, temuan ini menunjukkan bahwa *rolling statistics* berperan dalam membentuk kualitas representasi fitur temporal, dengan kontribusi utama pada kestabilan pembelajaran model dibandingkan peningkatan akurasi secara signifikan.

d. Interpretasi Hasil

Random Forest menunjukkan kinerja yang baik dan stabil, baik pada skenario *baseline* maupun dengan penambahan *lag features* dan *rolling statistics*. Pada tabel 5 menunjukkan bahwa *Random Forest* memiliki kinerja baik yang mampu memanfaatkan informasi historis suhu yang direpresentasikan dalam fitur temporal tanpa mengalami penurunan kemampuan generalisasi. Sebaliknya, *Gradient Boosting* menunjukkan kinerja pada kategori sedang, baik pada *baseline* maupun setelah penambahan *feature engineering*, yang mengindikasikan bahwa model tersebut memiliki keterbatasan dalam memanfaatkan variasi fitur temporal yang diberikan.

Tabel 5. Kombinasi Model dan Feature Engineering

Model	*Feature Engineering*	*Window*	Kinerja Relatif	Interpretasi Umum
Random Forest	*Baseline*	–	Baik	Model mampu mempelajari pola suhu tanpa fitur temporal tambahan
Random Forest	*Lag + Rolling Statistics*	3, 6, 12	Baik dan Stabil	Informasi historis memperkuat representasi temporal tanpa menurunkan akurasi
Gradient Boosting	*Baseline*	–	Sedang	Kinerja terbatas pada pola aditif dan fluktuasi tertentu
Gradient Boosting	*Lag + Rolling Statistics*	3, 6, 12	Sedang dan Stabil	Penambahan fitur temporal tidak mengubah kecenderungan kinerja

Temuan ini menunjukkan bahwa kualitas representasi fitur memiliki peran penting dalam mendukung performa model, namun efektivitasnya sangat bergantung pada karakteristik algoritma yang digunakan. Penerapan *rolling statistics* terbukti mampu menghasilkan performa yang stabil, meskipun tidak memberikan peningkatan akurasi yang signifikan, sebagaimana tercermin dari kecenderungan nilai kinerja pada Tabel 5 yang relatif konsisten pada setiap variasi *window*. Hal ini mengindikasikan bahwa fitur berbasis statistik lokal mampu membantu model dalam menangkap dinamika suhu tanpa meningkatkan kompleksitas model. Jika dibandingkan dengan penelitian sebelumnya yang menggunakan pendekatan model kompleks seperti LSTM dan kombinasi ARIMA–LSTM untuk meningkatkan akurasi prediksi suhu [3], hasil penelitian ini belum menunjukkan peningkatan akurasi yang lebih tinggi. Namun demikian, pendekatan yang digunakan dalam penelitian ini mampu menghasilkan performa yang kompetitif dengan metode yang lebih sederhana. Hal ini sejalan dengan penelitian yang menyatakan bahwa keberhasilan model tidak hanya ditentukan oleh kompleksitas algoritma, tetapi juga oleh kesesuaian antara fitur dan model yang digunakan [4]

Selain itu, hasil penelitian ini juga memperkuat temuan sebelumnya terkait efektivitas pendekatan berbasis jendela waktu dalam pemodelan deret waktu [12], di mana fitur yang menangkap dinamika lokal dapat membantu model dalam memahami pola temporal. Dengan demikian, kontribusi utama (*novelty*) penelitian ini terletak pada analisis sistematis terhadap penerapan *rolling statistics* sebagai teknik *feature engineering* sederhana yang mampu menghasilkan performa stabil dan representasi fitur yang informatif pada prediksi suhu berbasis *machine learning*, tanpa bergantung pada penggunaan model yang kompleks.

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0) License
Copyright © 2026 Czidni Sika Azkia, Yulian Ansori, Febriyanti Darnis, Arief Rahman

4. Kesimpulan

Berdasarkan hasil eksperimen, penerapan *feature engineering* berbasis kombinasi *lag* dan *rolling statistics* tidak memberikan peningkatan akurasi yang signifikan dibandingkan *baseline*, bahkan menunjukkan sedikit penurunan performa pada beberapa skenario. Pada model *Random Forest*, nilai MAE meningkat dari 0,0214 menjadi 0,0216 dan RMSE dari 0,1277 menjadi 0,1295 setelah penambahan fitur *rolling*. Meskipun demikian, penggunaan *rolling statistics* menghasilkan performa yang relatif stabil, yang menunjukkan bahwa fitur tersebut lebih berperan dalam menjaga kestabilan representasi fitur dibandingkan meningkatkan akurasi prediksi secara langsung. Analisis terhadap variasi ukuran *rolling window* menunjukkan bahwa ukuran jendela menengah memberikan keseimbangan antara sensitivitas terhadap fluktuasi jangka pendek dan kestabilan prediksi.

Dengan demikian, penelitian ini menunjukkan bahwa efektivitas *feature engineering* berbasis *rolling statistics* sangat bergantung pada karakteristik model yang digunakan, serta lebih berkontribusi pada kestabilan model dibandingkan peningkatan kinerja prediksi secara signifikan. Penelitian selanjutnya dapat mengkaji penggunaan teknik feature engineering lainnya atau kombinasi dengan model yang lebih kompleks untuk mengevaluasi potensi peningkatan kinerja prediksi. Selain itu, penelitian ini belum melibatkan model *baseline* klasik seperti ARIMA atau regresi linear, sehingga menjadi peluang untuk pengembangan selanjutnya dalam rangka memperoleh evaluasi yang lebih komprehensif.

Referensi

[1] H. Zhang, Y. Liu, C. Zhang, and N. Li, "Machine Learning Methods for Weather Forecasting: A Survey," *Atmosphere (Basel).*, vol. 16, no. 1, pp. 1–34, 2025, doi: https://doi.org/10.3390/atmos16010082.

[2] H. M. Alkhezi, "A comparative analysis of forecasting models for air temperature: Classical time series, machine learning, and functional approaches," *Kuwait J. Sci.*, vol. 53, no. 1, p. 100529, 2026, doi: https://doi.org/10.1016/j.kjs.2025.100529.

[3] M. Fauzi *et al.*, "Implementasi Machine Learning Untuk Memprediksi Cuaca Menggunakan Support Vector Machine," *J. Ilm. Komputasi*, vol. 23, no. 1, pp. 45–50, 2024, doi: https://doi.org/10.32409/jikstik.23.1.3499.

[4] J. Kim, H. Kim, H. G. Kim, D. Lee, and S. Yoon, "A comprehensive survey of deep learning for time series forecasting: architectural diversity and open challenges," *Artif. Intell. Rev.*, vol. 58, no. 7, 2025, doi: https://doi.org/10.1007/s10462-025-11223-9.

[5] C. Tan, J. Zhong, D. Yang, and W. Huang, "Forecasting and analysing global average temperature trends based on LSTM and ARIMA models," *PLoS One*, vol. 20, no. 9 September, pp. 1–27, 2025, doi: https://doi.org/10.1371/journal.pone.0330645.

[6] Y. Wang, H. Wu, J. Dong, Y. Liu, M. Long, and J. Wang, "Deep Time Series Models: A Comprehensive Survey and Benchmark," *arXiv preprint arXiv:2407.13278*, 2024, doi: 10.48550/arXiv.2407.13278.

[7] N. P. Nur Fauzi, S. Khomsah, and A. D. Putra Wicaksono, "Penerapan Feature Engineering dan Hyperparameter Tuning untuk Meningkatkan Akurasi Model Random Forest pada Klasifikasi Risiko Kredit," *J. Teknol. Inf. dan Ilmu Komput.*, vol. 12, no. 2, pp. 251–262, 2025, doi: https://doi.org/10.25126/jtiik.2025128472.

[8] X. Ma and H. Zhang, "Time Series Forecasting Method Based on Multi-Scale Feature Fusion and Autoformer," *Appl. Sci.*, vol. 15, no. 7, 2025, doi: https://doi.org/10.3390/app15073768.

[9] M. Shafiuzzaman *et al.*, "Enhanced very short-term load forecasting with multi-lag feature engineering and prophet-XGBoost-CatBoost architecture," *Energy*, vol. 335, no. February 2024, p. 137981, 2025, doi: https://doi.org/10.1016/j.energy.2025.137981.

[10] B. Khaw, R. Irwanto, R. Yunis, and E. Elly, "Analisis Time Series dan Perancangan Dashboard untuk Memprediksi Penjualan dengan Metode Prophet dan SARIMAX," *J. Sifo Mikroskil*, vol. 26, no. 2, 2025, doi: https://doi.org/10.55601/jsm.v26i2.1797.

[11] N. A. Elmunim et al., "Enhancing Wind and Solar Energy Forecasting Through Time-Series Feature Engineering and Ensemble Machine Learning," *Scientific Reports*, vol. 16, 2026, doi: https://doi.org/10.1038/s41598-026-49373-7.

[12] S. A. F. Mortezanejad and R. Wang, "Addressing Challenges in Time Series Forecasting: A Comprehensive Comparison of Machine Learning Techniques," no. March 2025, 2025, [Online]. Available: http://arxiv.org/abs/2503.20148

[13] P. Arsi, T. Astuti, D. Rahmawati, and P. Subarkah, "Implementasi Sliding Window Algotihm pada Prediksi Kurs berbasis Neural Network," *DoubleClick J. Comput. Inf. Technol.*, vol. 6, no. 1, p. 51, 2022, doi: https://doi.org/10.25273/doubleclick.v6i1.13496.

[14] E. P. Wahyuddin, R. E. Caraka, R. Kurniawan, W. Caesarendra, P. U. Gio, and B. Pardamean, "Improved

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0) License
Copyright © 2026 Czidni Sika Azkia, Yulian Ansori, Febriyanti Darnis, Arief Rahman

LSTM hyperparameters alongside sentiment walk-forward validation for time series prediction," *J. Open Innov. Technol. Mark. Complex.*, vol. 11, no. 1, p. 100458, 2025, doi: https://doi.org/10.1016/j.joitmc.2024.100458.

[15] C. Wang, M. D. Arseniev-Koehler, and M. Schmidt, "Towards Time-Series Feature Engineering in Automated Machine Learning for Multi-Step-Ahead Forecasting," *Engineering Proceedings*, vol. 18, no. 1, p. 17, 2022, doi: https://doi.org/10.3390/engproc2022018017.

[16] S. A. Khoiri and A. Wahid, "Jurnal Sistem dan Teknologi Informasi Indonesia Analisis Kinerja Algoritma Machine Learning dalam Prediksi Harga Cryptocurrency Performance Analysis of Machine Learning Algorithms in Cryptocurrency Price Prediction," vol. 9, no. 2, pp. 133–141, 2024, doi: https://doi.org/10.32528/justindo.v9i2.1965.

[17] A. Sutaryani, S. Sunarno, dan D. Djuniadi, "Perbandingan Performa Model Machine Learning dalam Prediksi Suhu di Semarang," *Jurnal Informatika dan Teknik Elektro Terapan*, vol. 12, no. 3, pp. 2770–2775, 2024, doi: https://doi.org/10.23960/jitet.v12i3.4884.

[18] H. Leidiyana and S. Nurajizah, "Pendekatan Hybrid Statistik dan Machine Learning untuk Peramalan Jumlah Kunjungan Turis," *J. Komtika (Komputasi dan Inform.*, vol. 9, no. 2, pp. 184–194, 2025, doi: https://doi.org/10.31603/komtika.v9i2.14909.

[19] S. Karmaker, "Feature Engineering in Time Series Forecasting," no. November, 2025, doi: https://doi.org/10.14293/PR2199.002316.v1.

[20] L. Alpianto, A. Hermawan, and Junaedi, "Moving Average untuk Prediksi Harga Saham dengan Linear Regression," *J. Buana Inform.*, vol. 14, no. 02, pp. 117–126, 2023, doi: https://doi.org/10.24002/jbi.v14i02.7446.

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0) License
Copyright © 2026 Czidni Sika Azkia, Yulian Ansori, Febriyanti Darnis, Arief Rahman