



**IMPLEMENTASI *RECURRENT NEURAL NETWORK* (RNN) UNTUK
MEMPREDIKSI JUMLAH KUNJUNGAN WISATAWAN DI PROVINSI DAERAH
ISTIMEWA YOGYAKARTA (DIY)**

***IMPLEMENTATION OF RECURRENT NEURAL NETWORK (RNN) FOR PREDICTING
THE NUMBER OF TOURIST VISITS IN THE SPECIAL REGION OF YOGYAKARTA
(DIY)***

Nabila 'Aqil Mufida*, Program Studi Matematika Universitas Negeri Yogyakarta
Syarifah Inayati, Program Studi Matematika Universitas Negeri Yogyakarta

*e-mail: nabilaaqil.2022@student.uny.ac.id

Abstrak

Sektor pariwisata di Provinsi Daerah Istimewa Yogyakarta (DIY) memiliki peran strategis terhadap perekonomian daerah, tetapi jumlah kunjungan wisatawan seringkali mengalami fluktuasi signifikan akibat berbagai faktor eksternal. Penelitian ini bertujuan untuk membangun model prediksi jumlah kunjungan wisatawan di Provinsi DIY menggunakan metode *Recurrent Neural Network* (RNN) dengan jaringan Elman dan mengetahui hasil prediksi. Data yang digunakan adalah data bulanan jumlah kunjungan wisatawan periode Januari 2019 hingga Desember 2024 (72 data) yang bersumber dari Badan Pusat Statistik Provinsi DIY, mencakup tiga periode yaitu prapandemi, pandemi, dan pasca pandemi. Penentuan variabel *input* dilakukan berdasarkan plot *Autocorrelation Function* (ACF) yang menghasilkan 3 *lag* signifikan, sehingga model dibangun dengan 3 variabel *input* dan 1 variabel *output*. Data dibagi dengan proporsi 80% data *training* dan 20% data *testing*. Pemilihan model terbaik dilakukan melalui *trial* dan *error* terhadap jumlah neuron lapisan tersembunyi dari 1 hingga 20 neuron. Model terbaik diperoleh dengan arsitektur 3-9-1 yang menghasilkan nilai MAPE data *training* sebesar 8,51% (kategori sangat baik) dan MAPE data *testing* sebesar 13,88% (kategori baik), sehingga menunjukkan kemampuan model yang baik dalam melakukan prediksi pada data yang belum pernah dilihat sebelumnya. Uji diagnostik *white noise* secara visual menunjukkan bahwa residual tidak memiliki *lag* yang signifikan, mengindikasikan model telah menangkap seluruh pola sistematis dari data. Hasil prediksi jumlah kunjungan wisatawan tahun 2025 menunjukkan pola musiman yang konsisten dengan data historis, dengan prediksi tertinggi pada bulan Desember 2025 sebesar 5.628.235 kunjungan dan terendah pada bulan Maret 2025 sebesar 1.042.415 kunjungan.

Kata kunci: Jaringan Elman, MAPE, Prediksi Wisatawan, *Recurrent Neural Network*

Abstract

The tourism sector in Daerah Istimewa Yogyakarta (DIY) Province plays a strategic role in the regional economy, but the number of tourist visits often fluctuates significantly due to various external factors. This study aims to build a prediction model for the number of tourist visits in DIY Province using the *Recurrent Neural Network* (RNN) method with an Elman network and to determine the prediction results for 2025. The data used are monthly tourist visit figures for January 2019 to December 2024 (72 data points) from the Central Statistics Agency of DIY Province, covering three periods: pre-pandemic, pandemic, and post-pandemic. Input variables were determined based on the *Autocorrelation Function* (ACF) plot, which identified 3 significant lags, so the model was built with 3 input variables and 1 output variable. Data was split 80% training and 20% testing. The best model was selected through trial and error on the number of hidden layer neurons 1 to 20. The optimal model had a 3-9-1 architecture yielding a training MAPE of 8.51% (very good) and testing MAPE of 13.88% (good). These result indicate that the model demonstrates satisfactory predictive performance and a good generalization ability when applied to unseen data. Visual white noise diagnostics on training residuals showed no significant lags, confirming the model captured all systematic patterns. Predictions for 2025 display seasonal patterns consistent with historical data, with the highest forecast of 5,628,235 visits in December 2025 and the lowest of 1,042,415 visits in March 2025.

Keywords: Elman Network, MAPE, Tourist Visit Prediction, *Recurrent Neural Network*

PENDAHULUAN

Indonesia merupakan negara kepulauan dengan lebih dari 17.000 pulau yang memiliki kekayaan alam, budaya, dan sejarah yang melimpah. Sektor pariwisata menjadi salah satu sektor unggulan yang mendorong pertumbuhan ekonomi nasional melalui peningkatan pendapatan negara, pembangunan daerah, penciptaan lapangan pekerjaan, serta penguatan citra Indonesia di ranah Internasional. Menurut BPS (2025), jumlah kunjungan wisatawan mancanegara ke Indonesia pada tahun 2024 mencapai 13 juta kunjungan, mencerminkan pemulihan sektor pariwisata yang signifikan pascapandemi Covid-19.

Provinsi Daerah Istimewa Yogyakarta (DIY) merupakan salah satu destinasi wisata unggulan Indonesia yang dikenal dengan keberagaman budaya, sejarah, dan keindahan alam. Daya tarik wisata Provinsi DIY yang terkenal, antara lain Candi Prambanan, Keraton Yogyakarta, Malioboro, Pantai Parangtritis, dan ekowisata Gunung Merapi. Data BPS DIY (2025) menunjukkan bahwa jumlah kunjungan wisatawan mancanegara ke DIY melalui Bandara YIA mengalami kenaikan 10,88 persen, yaitu mencapai 12.554 kunjungan. Sedangkan, jumlah kunjungan wisatawan nusantara ke DIY turun 14,73 persen, yaitu 3.001.094 perjalanan. Meskipun tren secara kumulatif meningkat, fluktuasi jumlah kunjungan wisatawan masih sering terjadi karena dipengaruhi kondisi ekonomi global, kebijakan visa, promosi pariwisata, serta kondisi sosial dan kesehatan masyarakat.

Sebagai daerah yang memiliki ketergantungan terhadap sektor pariwisata, perubahan jumlah kunjungan wisatawan memberikan dampak yang signifikan terhadap perekonomian daerah. Pada triwulan IV 2024, pertumbuhan ekonomi DIY menempati posisi tertinggi dibandingkan provinsi lain se-Pulau Jawa. Hal ini tercermin dari struktur perekonomian DIY, di mana Lapangan Usaha Penyediaan Akomodasi dan Makan Minum sebagai sektor yang erat kaitannya dengan kegiatan pariwisata mengalami pertumbuhan tertinggi sebesar 13,40 persen pada triwulan IV 2024. Perekonomian DIY turut digerakkan oleh beberapa Proyek Strategi Nasional (PSN) yang berjalan, serta kunjungan wisatawan (Pemda DIY, 2025).

Prediksi jumlah kunjungan wisatawan menjadi penting sebagai dasar pengambilan keputusan dan strategi pembangunan pariwisata. Kemampuan prediksi yang akurat dapat mendukung pengurangan risiko akibat ketidakpastian eksternal, termasuk fluktuasi ekonomi dan perubahan kebijakan perjalanan, sehingga memastikan model mampu menyesuaikan diri terhadap perubahan mendadak yang terjadi dalam sektor pariwisata (Zhang et al., 2025). Data kunjungan wisatawan yang dicatat secara periodik merupakan bentuk data deret waktu (*time series*) yang perlu dianalisis menggunakan metode yang mampu menangkap dinamika temporal data.

Penelitian terdahulu mengenai prediksi kunjungan wisatawan di Provinsi DIY umumnya menggunakan metode statistik klasik. Penelitian Noho & Fajriyah (2023) menggunakan Holt-Winter's Exponential Smoothing dengan data sebelum pandemi dan berhasil memperoleh model yang cukup baik. Difitria et al. (2020) menggunakan *Support vector Regression* dan *Particle Swarm Optimization* dan memperoleh nilai MAPE terkecil sebesar 1,088%. Adapun Berlinditya & Noeryanti (2019) menggunakan metode ARIMAX untuk memprediksi pengunjung wisata di Kabupaten Gunungkidul dengan akurasi 79%, serta Wibowo & Purwaningsih (2023) menggunakan SARIMA pada data wisata Kabupaten Bantul dengan MAPE 13,37%.

Meskipun metode statistik klasik seperti ARIMA dan SARIMA populer untuk analisis deret waktu, pendekatan tersebut memiliki keterbatasan dalam menangkap pola nonlinear dan dependensi temporal jangka panjang yang menjadi karakteristik data kunjungan wisatawan. Penelitian Sun et al. (2019) menunjukkan bahwa model *machine learning* berbasis *deep learning* secara konsisten lebih unggul dibandingkan model statistik model klasik dalam

prediksi kunjungan wisatawan. Hal ini dikonfirmasi oleh Lu et al. (2025) yang menyatakan bahwa *deep learning* telah merevolusi prediksi deret waktu dengan meningkatkan kemampuan pemodelan nonlinear secara signifikan, serta Safitri (2024) yang menunjukkan metode *deep learning* cenderung memberikan hasil lebih baik dalam hal akurasi.

Recurrent Neural Network (RNN) merupakan salah satu metode *deep learning* yang dirancang khusus untuk data berurutan (*sequential*). Keunggulan utama RNN terletak pada arsitekturnya yang memiliki mekanisme umpan balik (*feedback loop*) sehingga informasi dari langkah sebelumnya dapat digunakan pada langkah waktu berikutnya (Rizal & Soraya, 2018). Dalam keluarga RNN jaringan Elman merupakan arsitektur standar yang paling umum digunakan dan telah terbukti memiliki kemampuan lebih baik dibandingkan Fully RNN, Jordan RNN, LSTM, maupun *feedforward* untuk data kunjungan wisatawan (Sugiartawan & Hartati, 2019). Jaringan Elman memiliki unit tambahan sebagai memori internal yang menyimpan nilai hidden layer dari langkah sebelumnya, memungkinkan pemrosesan informasi secara sekuensial dan pembelajaran ketergantungan temporal data dari historis secara lebih efektif (Hsieh, 2021).

Beberapa penelitian menggunakan model RNN untuk permasalahan prediksi data deret waktu. Penelitian Yuhono et al. (2024) menggunakan RNN untuk memprediksi harga Steel Hot-Rolled Coil dengan akurasi 90,90% pada *training* dan 90,02% pada *testing*. Penelitian Siregar et al. (2021) menggunakan RNN untuk memprediksi kesehatan masyarakat Indonesia dan memperoleh prediksi yang sangat baik dengan kedekatan antara data aktual dan prediksi. Penelitian Ghazi et al. (2022) prediksi kasus Covid-19 di Provinsi Lampung menggunakan RNN. Diperoleh hasil prediksi dari penelitian tersebut adalah model kasus aktif kumulatif RMSE sebesar 0,001, kasus kesembuhan kumulatif sebesar 0,0027, dan kasus kematian kumulatif sebesar 0,001. Namun, penelitian menggunakan RNN khususnya jaringan Elman untuk prediksi kunjungan wisatawan di Provinsi DIY masih sangat terbatas.

Berdasarkan uraian di atas, penelitian ini bertujuan untuk membangun model prediksi jumlah kunjungan wisatawan di Provinsi DIY menggunakan RNN jaringan Elman dan mengetahui hasil prediksi jumlah kunjungan wisatawan di Provinsi DIY. Diharapkan penelitian ini dapat menjadi referensi bagi pemerintah daerah dan dinas pariwisata dalam menyusun kebijakan, mengalokasikan sumber daya, serta mengoptimalkan fasilitas sarana dan prasarana pariwisata secara lebih terencana.

METODE

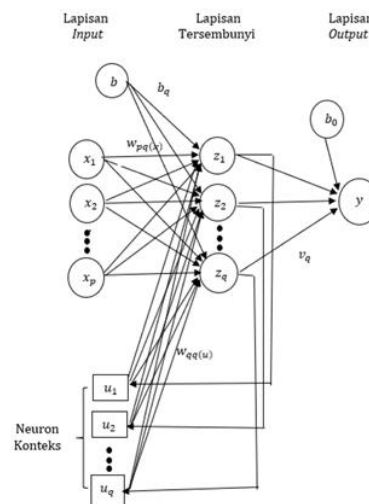
Data

Penelitian ini menggunakan data sekunder berupa data bulanan jumlah kunjungan wisatawan di Provinsi DIY yang diperoleh dari situs resmi Badan Pusat Statistik (BPS) Provinsi DIY. Data mencakup total jumlah kunjungan wisatawan yang terdiri dari wisatawan mancanegara dan wisatawan nusantara, selama periode Januari 2019 sampai dengan Desember 2024, sehingga terdapat 72 data. Rentang data ini mencakup tiga periode yang berbeda karakteristiknya, yaitu prapandemi (Januari 2019–Maret 2020), pandemi Covid-19 (April 2020–Maret 2022), dan pascapandemi/pemulihan (April 2022–Desember 2024).

Recurrent Neural Network

Recurrent Neural Network (RNN) adalah jaringan yang dirancang untuk memproses data berurutan dengan memanfaatkan mekanisme umpan balik. Pada model RNN terdapat jaringan yang dikenal dengan jaringan Elman. Jaringan Elman merupakan arsitektur RNN standar yang dikembangkan oleh Jeffrey Elman pada tahun 1990. Arsitektur ini memiliki unit konteks (*context unit*) sebagai ciri utama, yaitu neuron tambahan yang menyimpan nilai output

lapisan tersembunyi dari langkah waktu sebelumnya untuk digunakan kembali pada langkah waktu berikutnya. Arsitektur jaringan Elman mencakup lapisan input, lapisan tersembunyi, lapisan berulang yang menyimpan informasi, dan lapisan output. Setiap lapisan terdiri dari satu atau lebih neuron yang memproses data dengan menghitung akumulasi bobot input dan menerapkan fungsi aktivasi nonlinear untuk mentransmisikan informasi (Wang et al., 2016). Input pada model RNN jaringan Elman untuk prediksi data deret waktu merupakan *lag* dari variabel deret waktu. Pada deret waktu y merupakan neuron lapisan *output*, maka $x_1, x_2, \dots, x_p$ merupakan p lapisan *input* dari model RNN jaringan Elman. Misalkan $z_1, z_2, \dots, z_q$ merupakan neuron lapisan tersembunyi dan $u_1, u_2, \dots, u_q$ sebagai neuron tambahan yang merupakan *feedback loop* dari lapisan tersembunyi. Arsitektur RNN jaringan Elman untuk data deret waktu dapat dilihat pada Gambar 1.



Gambar 1. Arsitektur RNN jaringan Elman

Fungsi aktivasi yang digunakan pada lapisan tersembunyi adalah sigmoid bipolar dan pada lapisan output menggunakan fungsi aktivasi linear. Model RNN jaringan Elman dapat dirumuskan sebagai berikut

$$y = \sum_{k=1}^q v_k z_k + b_0$$

dengan,

$$z_k = \frac{1 - \exp\left(-\left(\sum_{j=1}^p x_j w_{jk} + \sum_{k'=1}^q u_{k'} w_{k'k} + b_k\right)\right)}{1 + \exp\left(-\left(\sum_{j=1}^p x_j w_{jk} + \sum_{k'=1}^q u_{k'} w_{k'k} + b_k\right)\right)}$$

v_k merupakan bobot neuron lapisan tersembunyi ke lapisan *output*, w_{jk} bobot dari lapisan *input* ke- j ke lapisan lapisan tersembunyi ke- k , $u_{k'}$ variabel *input* lapisan tambahan, $w_{k'k}$ bobot dari lapisan tambahan ke- k' ke lapisan tersembunyi ke- k , b_k bobot bias neuron lapisan tersembunyi ke- k , dan b_0 bobot bias pada neuron lapisan *output*.

Tahapan Analisis Data

Proses pemodelan dilakukan melalui beberapa tahap yang berurutan. Pertama, penentuan variabel *input* berdasarkan plot ACF dari data jumlah kunjungan wisatawan. Apabila terdapat k *lag* yang signifikan, maka variabel *input* yang digunakan adalah $y_{t-k}, \dots, y_{t-2}, y_{t-1}$

dengan variabel *output* y_t . Variabel *input* dan *output* tersebut dinotasikan dengan x_1, x_2, x_3 , dan y . Kedua, pembagian data dengan proporsi 80% data *training* (55 data) dan 20% data *testing* (14 data).

Ketiga, normalisasi data menggunakan metode standarisasi (z-score) dengan bantuan fungsi `prestd` pada *software*. Parameter normalisasi (*mean* dan standar deviasi) dihitung dari data *training* kemudian digunakan untuk menormalisasi data *testing*. Normalisasi dilakukan agar data berada pada rentang yang sesuai dengan fungsi aktivasi tansig, yaitu $[-1, 1]$. Rumus normalisasi yang digunakan adalah:

$$x'_t = \frac{(x_t - \bar{x})}{s}$$

dengan $\bar{x}$ adalah *mean* dan s adalah standar deviasi. Setelah proses prediksi, *output* jaringan yang masih dalam skala ternormalisasi dikembalikan ke skala asli menggunakan fungsi `poststd`.

Keempat, pembentukan model RNN jaringan Elman menggunakan perintah fungsi `newelm`. Penentuan jumlah neuron optimal pada lapisan tersembunyi dilakukan melalui proses *trial* dan *error* dari 1 hingga 20 neuron. Kelima, evaluasi model menggunakan *Mean Absolute Percentage Error* (MAPE). Model terbaik dipilih berdasarkan nilai MAPE terkecil dengan mempertimbangkan keseimbangan antara performa *training* dan kemampuan generalisasi. Kriteria MAPE yang digunakan ditunjukkan pada Tabel 1 (Hajjah & Marlim, 2021).

Tabel 1. Kriteria nilai MAPE

Nilai MAPE	Kriteria
$x < 10\%$	Sangat baik
$10\% \leq x < 20\%$	Baik
$20\% \leq x < 50\%$	Cukup baik
$\geq 50\%$	Buruk

Keenam, analisis diagnostik *white noise* secara visual menggunakan plot ACF dan PACF dari residual data *training*. Terakhir, prediksi jumlah kunjungan wisatawan menggunakan model terpilih.

HASIL DAN PEMBAHASAN

Statistik Deskriptif

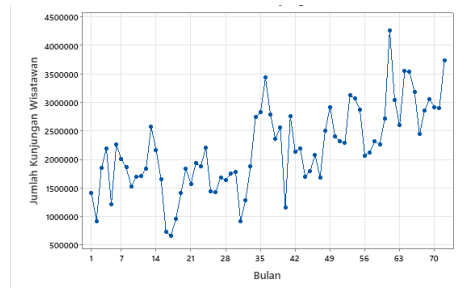
Sebelum dilakukan pemodelan, terlebih dahulu dilakukan analisis statistik deskriptif terhadap data jumlah kunjungan wisatawan di Provinsi DIY periode Januari 2019 hingga Desember 2024. Hasil analisis statistik deskriptif ditunjukkan pada Tabel 2.

Tabel 2. Statistik deskriptif data

Statistik	Nilai
Jumlah data	72
<i>Mean</i>	2.186.078
<i>Median</i>	2.157.422

Nilai minimum	665.050
Nilai maksimum	4.274.114
Standar deviasi	744.859

Berdasarkan Tabel 2, *mean* jumlah kunjungan wisatawan per bulan adalah 2.186.078 kunjungan dengan standar deviasi sebesar 744.859. Nilai minimum sebesar 665.050 kunjungan terjadi pada bulan Mei 2020, sedangkan nilai maksimum sebesar 4.274.114 kunjungan terjadi pada bulan Januari 2024.

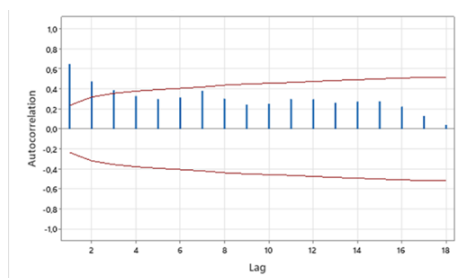


Gambar 2. Plot kunjungan wisatawan Provinsi DIY

Berdasarkan Gambar 2. pola data jumlah kunjungan wisatawan di Provinsi DIY menunjukkan karakteristik yang bersifat dinamis dan tidak stasioner dengan tren meningkat secara umum dari awal hingga akhir periode. Penurunan tajam pada pertengahan tahun 2020 mencerminkan dampak pandemi Covid-19, sedangkan fluktuasi periodik yang berulang mengindikasikan adanya komponen musiman yang berkaitan dengan musim liburan, hari raya nasional, dan periode puncak wisata tahunan. Karakteristik data yang tidak stasioner, nonlinear, dan mengandung pola musiman ini menjadi justifikasi pemilihan model RNN jaringan Elman yang mampu menangkap ketergantungan temporal kompleks.

Penentuan *Input* dan Pembagian Data

Langkah pertama dalam pemodelan adalah menentukan variabel *input* berdasarkan analisis plot ACF data jumlah kunjungan wisatawan. Hasil plot ACF ditunjukkan pada Gambar 3.



Gambar 3. Plot ACF jumlah kunjungan wisatawan

Berdasarkan Gambar 3, terdapat 3 *lag* yang keluar dari batas signifikan, yaitu *lag 1*, *lag 2*, dan *lag 3*. Hal ini mengindikasikan bahwa jumlah kunjungan wisatawan pada suatu bulan memiliki korelasi yang signifikan dengan tiga bulan sebelumnya, yang mencerminkan adanya ketergantungan temporal jangka pendek pada data. Berdasarkan hasil tersebut, model RNN dibangun dengan 3 variabel *input*, yaitu y_{t-3} , y_{t-2} , dan y_{t-1} dengan *output* y_t . Untuk data *input* y_{t-3} , y_{t-2} , y_{t-1} dan *output* y_t diberi simbol secara berurutan x_1 , x_2 , x_3 , dan y .

Dengan menggunakan *lag* 3 sebagai *input*, data ke-1 hingga ke-3 tidak dapat digunakan sehingga total data yang dianalisis menjadi 69 data (data ke-4 hingga ke-72). Pembagian data dilakukan dengan proporsi 80% data *training* dan 20% data *testing*, sehingga diperoleh 55 data *training* dan 14 data *testing*. Selanjutnya, normalisasi data dilakukan pada data *input* dan target *output* data *training* dan *testing*.

Parameter dan Pembentukan Model RNN

Pembentukan model RNN jaringan Elman dilakukan menggunakan fungsi `newelm` dengan perintah `net = newelm(minmax(Pn), [i 1], {'tansig' 'purelin'}, 'traingdx')`, di mana *i* adalah jumlah neuron pada lapisan tersembunyi yang divariasikan. Parameter pembelajaran yang digunakan disajikan pada Tabel 3.

Tabel 3. Parameter model RNN

Parameter	Nilai
Maksimum <i>epoch</i>	5.000
Target <i>error</i>	0,00001
<i>Learning rate</i>	0,1
Momentum	0,8
<i>Epoch show</i>	500

Proses *trial* dan *error* dilakukan dengan memvariasikan jumlah neuron lapisan tersembunyi dari 1 hingga 20 neuron. Setiap konfigurasi dilatih dan dievaluasi berdasarkan nilai MAPE pada data *training* maupun *testing*. Diperoleh model terbaik dengan 9 neuron pada lapisan tersembunyi. Model RNN tersebut dipilih sebagai model terbaik karena menghasilkan nilai MAPE *testing* terkecil, yaitu 13,88%. Kriteria pemilihan model tidak hanya berdasarkan MAPE *training* yang rendah, melainkan terutama berdasarkan keseimbangan antara MAPE *training* dan MAPE *testing*. Beberapa model dengan neuron lebih banyak memang menghasilkan MAPE *training* yang lebih kecil, tetapi MAPE *testing* yang sangat tinggi mengindikasikan terjadinya *overfitting*. Contoh ekstrem tampak pada model dengan 11 neuron yang menghasilkan MAPE *training* 6,43% tetapi MAPE *testing* melonjak hingga 93,47%.

Arsitektur model terpilih adalah 3-9-1, yaitu 3 neuron *input*, 9 neuron pada lapisan tersembunyi, dan 1 neuron *output*. Setelah proses *training* selesai, diperoleh bobot-bobot akhir yang akan digunakan dalam proses prediksi.

Uji White Noise

Setelah model terbaik dipilih, dilakukan analisis diagnostik *error* melalui uji *white noise* secara visual menggunakan plot ACF dan PACF dari residual data *training*. Pada model RNN jaringan Elman, pengujian *white noise* bukan merupakan syarat mutlak kelayakan model sebagaimana halnya pada model statistik klasik seperti ARIMA, melainkan berfungsi sebagai analisis diagnostik tambahan untuk menilai kualitas model (Zhang, 2003). Berbeda dengan model ARIMA yang mensyaratkan residual bersifat *white noise* sebagai kondisi formal, model berbasis jaringan syaraf tiruan tidak memiliki asumsi distribusi residual yang baku. Oleh karena itu, hasil uji ini digunakan sebagai bukti tambahan yang memperkuat layaknya model digunakan.

Hasil plot ACF dan PACF dari residual data *training* model terpilih, tidak ditemukan *lag* yang melebihi batas signifikan pada kedua plot tersebut. Kondisi ini mengindikasikan bahwa residual bersifat acak atau *white noise*, yang berarti model telah berhasil mengekstrak seluruh pola sistematis yang terkandung dalam data. Dengan diperolehnya MAPE data *training* 8,51% dan MAPE data *testing* 13,88% serta konfirmasi residual *white noise*, model RNN jaringan Elman dengan arsitektur 3-9-1 dinyatakan layak untuk digunakan dalam prediksi jumlah kunjungan wisatawan di Provinsi DIY.

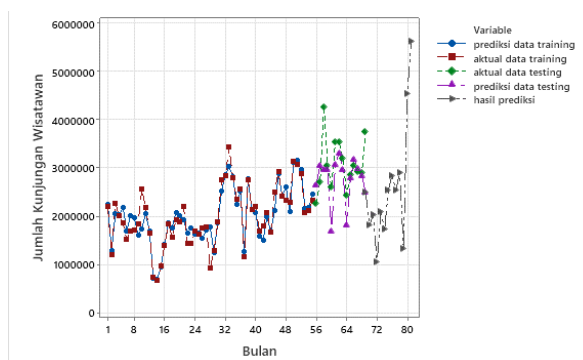
Prediksi Jumlah Kunjungan Wisatawan Tahun 2025

Model RNN jaringan Elman terpilih digunakan untuk memprediksi jumlah kunjungan wisatawan di Provinsi DIY untuk bulan Januari hingga Desember 2025 dengan menggunakan bobot-bobot akhir yang diperoleh pada proses *training*. Prediksi dilakukan secara bertahap dengan memanfaatkan tiga data bulan sebelumnya sebagai *input* pada setiap langkah prediksi. Sebagai ilustrasi, prediksi untuk bulan Januari 2025 menggunakan data Oktober, November, dan Desember 2024 sebagai *input*. Hasil prediksi jumlah kunjungan wisatawan ditunjukkan oleh Tabel 4.

Tabel 4. Hasil prediksi jumlah kunjungan wisatawan

Periode	Hasil Prediksi
Januari 2025	1.815.171
Februari 2025	2.041.021
Maret 2025	1.042.415
April 2025	2.087.637
Mei 2025	1.724.727
Juni 2025	2.546.702
Juli 2025	2.843.139
Agustus 2025	2.533.542
September 2025	2.911.042
Oktober 2025	1.328.551
November 2025	4.554.946
Desember 2025	5.628.235

Berdasarkan Tabel 4, hasil prediksi menunjukkan bahwa jumlah kunjungan wisatawan tertinggi diperkirakan terjadi pada bulan Desember 2025 sebesar 5.628.235 kunjungan, diikuti November 2025 sebesar 4.554.946 kunjungan dan September 2025 sebesar 2.911.042 kunjungan. Prediksi terendah terjadi pada bulan Maret 2025 sebesar 1.042.415 kunjungan, diikuti Oktober 2025 sebesar 1.328.551 kunjungan. Secara keseluruhan, pola prediksi menunjukkan variasi musiman yang cukup besar antarbulan dengan kecenderungan meningkat pada akhir tahun.



Gambar 4. Perbandingan data aktual dengan hasil prediksi

Gambar 4 menunjukkan perbandingan antara data aktual dengan hasil prediksi data *training-testing* serta hasil prediksi tahun 2025. Secara umum, garis prediksi mampu mengikuti pola dan arah data aktual dengan cukup baik, meskipun pada beberapa titik terdapat selisih yang cukup besar. Hal ini sesuai dengan nilai MAPE yang diperoleh oleh model RNN terbaik.

Pembahasan

Model RNN jaringan Elman dengan arsitektur 3-9-1 berhasil dibangun untuk memprediksi jumlah kunjungan wisatawan di Provinsi DIY. Penentuan 3 variabel *input* berdasarkan plot ACF menunjukkan bahwa jumlah kunjungan wisatawan pada suatu bulan memiliki ketergantungan yang signifikan terhadap tiga bulan sebelumnya. Secara substantif, kondisi ini menunjukkan bahwa pola kunjungan wisatawan pada suatu periode dipengaruhi oleh pola kunjungan pada periode-periode sebelumnya, yang berkaitan dengan efek musiman, tren jangka panjang, maupun faktor eksternal yang berlangsung secara berkelanjutan dalam beberapa periode.

Pemilihan 9 neuron pada lapisan tersembunyi didasarkan pada keseimbangan antara performa *training* dan kemampuan generalisasi pada data *testing*. Hasil *trial* dan *error* mengkonfirmasi fenomena *overfitting* pada model dengan jumlah neuron yang berlebihan, yaitu kondisi di mana jaringan terlalu menyesuaikan diri dengan pola spesifik pada data *training* sehingga kemampuan generalisasinya terhadap data baru menurun drastis. Hal ini sejalan dengan teori yang dikemukakan oleh Haykin (1999) bahwa kemampuan generalisasi model dipengaruhi oleh keseimbangan ukuran data *training* dan kompleksitas arsitektur jaringan. Tidak ada rumus baku untuk menentukan jumlah neuron optimal pada lapisan tersembunyi, sehingga prosedur *trial* dan *error* merupakan pendekatan yang umum digunakan dalam penelitian (Aiziyah et al., 2025).

Nilai MAPE data *training* sebesar 8,51% masuk dalam kategori 'sangat baik' ($\text{MAPE} < 10\%$), sedangkan MAPE data *testing* sebesar 13,88% masuk dalam kategori 'baik' ($10\% \leq \text{MAPE} < 20\%$) berdasarkan kriteria Hajjah & Marlim (2021). Selisih MAPE antara data *training* dan *testing* sebesar sekitar 5 poin persentase masih tergolong wajar, menunjukkan bahwa model tidak mengalami *overfitting* secara signifikan dan memiliki kemampuan generalisasi yang memadai. Nilai tersebut cukup baik mengingat kompleksitas data yang mencakup tiga periode berbeda (prapandemi, pandemi, dan pasca pandemi) dengan variasi yang sangat ekstrem. Nilai minimum data (665.050) merupakan kurang dari sepertiga nilai maksimum (4.274.114), menunjukkan betapa besarnya rentang variasi yang harus ditangkap oleh model.

Perbandingan dengan penelitian sebelumnya menunjukkan performa yang kompetitif. Sugiartawan & Hartati (2019) yang menggunakan Elman RNN untuk prediksi kunjungan

wisatawan di Tanah Lot Bali memperoleh MSE sebesar 0,05105, menunjukkan keunggulan jaringan Elman untuk data kunjungan wisatawan. Siregar et al. (2021) yang menggunakan RNN untuk prediksi kesehatan masyarakat Indonesia menunjukkan hasil prediksi yang sangat baik dengan nilai aktual dan prediksi yang sangat berdekatan pada interval tertentu. Yuhono et al. (2024) yang menggunakan RNN untuk prediksi harga Steel Hot-Rolled Coil memperoleh akurasi 90,90% pada *training* dan 91,02% pada *testing*, yang merupakan performa tinggi untuk data keuangan yang relatif lebih stabil dibandingkan data kunjungan wisatawan.

Hasil analisis diagnostik residual mengkonfirmasi bahwa model telah berhasil menangkap seluruh pola sistematis dalam data. Tidak adanya *lag* yang signifikan pada plot ACF dan PACF residual data *training* menunjukkan bahwa residual bersifat *white noise*, artinya model tidak menyisakan pola yang masih dapat dimodelkan. Meskipun pada model RNN jaringan Elman pengujian ini bukan merupakan syarat wajib sebagaimana pada model ARIMA, hasil yang diperoleh menjadi bukti tambahan yang memperkuat kualitas dan kelayakan model untuk digunakan dalam prediksi.

Hasil prediksi tahun 2025 menunjukkan pola musiman yang konsisten dengan karakteristik data historis. Prediksi terendah pada bulan Maret 2025 sebesar 1.042.415 kunjungan dapat dikaitkan dengan berakhirnya periode puncak wisata awal tahun (libur Natal dan Tahun Baru) dan belum dimulainya musim liburan berikutnya, sehingga secara historis bulan Maret cenderung memiliki jumlah kunjungan yang lebih rendah. Peningkatan pada periode Juni–Juli 2025 berkaitan dengan musim liburan pertengahan tahun (libur sekolah), sementara lonjakan tajam pada November–Desember 2025 mencerminkan pola musiman yang kuat, sejalan dengan peningkatan aktivitas wisata pada periode libur akhir tahun. Fenomena ini sesuai dengan teori komponen musiman dalam analisis deret waktu, di mana pola berulang secara periodik umumnya dipengaruhi oleh faktor kalender (Hanke & Wichern, 2014).

Adanya lonjakan prediksi yang sangat tajam pada bulan November dan Desember 2025 perlu disikapi dengan bijaksana. Lonjakan ini kemungkinan dipengaruhi oleh pola data akhir tahun 2023–2024 dalam data *training* yang menunjukkan peningkatan signifikan, sehingga model cenderung mengekstrapolasi tren tersebut. Informasi prediksi ini tetap dapat menjadi dasar perencanaan bagi pelaku industri pariwisata, pemerintah daerah, dan dinas pariwisata dalam mempersiapkan akomodasi, infrastruktur, dan layanan wisata secara lebih terencana dan antisipatif.

SIMPULAN

Berdasarkan hasil penelitian dapat disimpulkan bahwa model RNN jaringan Elman terbaik untuk memprediksi jumlah kunjungan wisatawan di Provinsi DIY adalah model dengan arsitektur 3-9-1, yaitu 3 variabel *input*, 9 neuron pada lapisan tersembunyi, dan 1 variabel *output*. Model ini menghasilkan MAPE data *training* sebesar 8,51% (kategori sangat baik) dan MAPE data *testing* sebesar 13,88% (kategori baik). Analisis diagnostik residual menunjukkan bahwa *error* model bersifat *white noise*, mengkonfirmasi bahwa model telah berhasil menangkap seluruh pola sistematis dari data.

Hasil prediksi jumlah kunjungan wisatawan di Provinsi DIY tahun 2025 secara berturut-turut dari Januari hingga Desember adalah: 1.815.171; 2.041.021; 1.042.415; 2.087.637; 1.724.727; 2.546.702; 2.843.139; 2.533.542; 2.911.042; 1.328.551; 4.554.946; dan 5.628.235 kunjungan. Pola prediksi menunjukkan variasi musiman yang konsisten dengan data historis, dengan kecenderungan peningkatan yang kuat pada periode libur pertengahan dan akhir tahun. Informasi prediksi ini diharapkan dapat menjadi acuan bagi pemerintah daerah dan dinas pariwisata Provinsi DIY dalam menyusun kebijakan, mengalokasikan anggaran, serta mengoptimalkan fasilitas sarana dan prasarana pariwisata secara lebih terencana.

Untuk penelitian selanjutnya disarankan untuk: (1) menambahkan variabel eksternal seperti hari libur nasional, event pariwisata, dan data sosial media sebagai variabel input tambahan untuk meningkatkan akurasi prediksi; (2) membandingkan model RNN jaringan Elman dengan arsitektur deep learning lainnya seperti LSTM, GRU, atau Transformer untuk mengetahui model yang paling optimal untuk data kunjungan wisatawan di Provinsi DIY.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada Badan Pusat Statistik Provinsi Daerah Istimewa Yogyakarta yang telah menyediakan data kunjungan wisatawan secara terbuka melalui portal resminya. Penulis juga menyampaikan terima kasih kepada Program Studi Matematika Fakultas MIPA Universitas Negeri Yogyakarta atas segala dukungan akademik yang diberikan dalam penyelesaian penelitian ini.

DAFTAR PUSTAKA

- Aiziyah, E., Lukmana, A. H., & Abadi, A. M. (2025). Prediksi rate of penetration pada pengeboran minyak bumi dengan Elman recurrent neural network. *Journal MIND Journal*, 10(2), 145–161. <https://doi.org/10.26760/mindjournal.v10i2.145-161>
- Berlinditya, B., & Noeryanti. (2019). Pemodelan time series dalam peramalan jumlah pengunjung objek wisata di Kabupaten Gunungkidul menggunakan metode ARIMAX efek variasi kalender. *Jurnal Statistika Industri Dan Komputasi*, 4(1), 81–88.
- BPS. (2025). *Statistik kunjungan wisatawan mancanegara 2024*. <https://www.bps.go.id/id/publication/2025/03/20/a85d584df19ea65a5e2b3d0b/statistik-kunjungan-wisatawan-mancanegara-2024>
- BPS DIY. (2025). *Perkembangan pariwisata DIY Agustus 2025*. <https://yogyakarta.bps.go.id/id/pressrelease/2025/10/01/1663/jumlah-kunjungan-wisatawan-mancanegara-ke-indonesia-melalui-bandara-yia-selama-bulan-agustus-2025-naik-10-88-persen-dibandingkan-juli-2025.html>
- Difitria, R., Cholissodin, I., & Indriati. (2020). Penerapan support vector regression dan particle swarm optimization untuk prediksi jumlah kunjungan wisatawan mancanegara ke Daerah Istimewa Yogyakarta. *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, 4(5), 1364–1371.
- Ghozi, A. A., Aprianti, A., Dimas, A. D. P., & Fauzi, R. (2022). Analisis prediksi data kasus covid-19 di Provinsi Lampung menggunakan recurrent neural network (RNN). *Indonesian Journal of Applied Mathematics*, 2(1), 25–32. <https://journal.itera.ac.id/index.php/indojam/>
- Hajjah, A., & Marlim, Y. N. (2021). Analisis error terhadap peramalan data penjualan. *Techno.Com*, 20(1), 1–9. <https://doi.org/10.33633/tc.v20i1.4054>
- Hanke, J. E. ., & Wichern, D. W. . (2014). *Business forecasting* (9th ed.). Pearson Educated Limited.
- Haykin, S. (1999). *Neural network: A comprehensive foundation* (2nd ed.). Prentice Hall.

- Hsieh, S.-C. (2021). Tourism demand forecasting based on an LSTM network and its variants. *Algorithms*, 14(8), 243. <https://doi.org/10.3390/a14080243>
- Lu, G., Ou, Y., Wang, Z., Qu, Y., Xia, Y., Tang, D., Kotenko, I., & Li, W. (2025). A survey of deep learning for time series forecasting: Theories, datasets, and state-of-the-art techniques. *Computers, Materials and Continua*, 85(2), 2403–2441. <https://doi.org/10.32604/cmc.2025.068024>
- Noho, C. T., & Fajriyah, R. (2023). Analisis peramalan jumlah wisatawan ke DIY menggunakan Holt-Winter's exponential smoothing. *Emerging Statistics and Data Science Journal*, 1(1).
- Pemda DIY. (2025, Februari 2). Pertumbuhan ekonomi DIY 2024 masih tertinggi se-Pulau Jawa. Portal Pemerintah Daerah Daerah Istimewa Yogyakarta. <https://jogjaprov.go.id/berita/detail-berita/pertumbuhan-ekonomi-diy-2024> masih-tertinggi-se-pulau-jawa
- Rizal, A. A., & Soraya, S. (2018). Multi time steps prediction dengan recurrent neural network long short term memory. *MATRIK : Jurnal Manajemen, Teknik Informatika Dan Rekayasa Komputer*, 18(1), 115–124. <https://doi.org/10.30812/matrik.v18i1.344>
- Safitri, B. (2024). Perbandingan kinerja antara metode deep learning dan metode klasik dalam pengenalan pola pada citra medis. *Logicloom.Id*, 1(2), 1–21.
- Siregar, A. M., Jaman, J. H., & Mufti, A. (2021). Analisa prediksi kesehatan masyarakat Indonesia menggunakan recurrent neural network. *INTERNAL (Information System Journal)*, 4(1), 28–34. <https://doi.org/10.32627>
- Sugiartawan, P., & Hartati, S. (2019). Time series data prediction using Elman recurrent neural network on tourist visits in Tanah Lot tourism object. *International Journal of Engineering and Advanced Technology*, 9(1), 314–320. <https://doi.org/10.35940/ijeat.A1833.109119>
- Sun, S., Wei, Y., Tsui, K.-L., & Wang, S. (2019). Forecasting tourist arrivals with machine learning and internet search index. *Tourism Management*, 70, 1–10. <https://doi.org/10.1016/j.tourman.2018.07.010>
- Wang, J., Wang, J., Fang, W., & Niu, H. (2016). Financial time series prediction using Elman recurrent random neural networks. *Computational Intelligence and Neuroscience*, 2016, 1–14. <https://doi.org/10.1155/2016/4742515>
- Wibowo, D. B., & Purwaningsih, T. (2023). Analisis peramalan jumlah pengunjung daya tarik wisata Kabupaten Bantul dengan metode seasonal autoregressive integrated moving average (SARIMA). *Emerging Statistics and Data Science Journal*, 1(1), 117–127.
- Yuhono, T., Hassolthine, C. ramadhona, & Sahara, R. (2024). Prediksi harga steel hot-rolled (HRC) dengan model recurrent neural network (RNN). *JEKIN (Jurnal Teknik Informatika)*, 4(1), 1–9.
- Zhang, G. P. (2003). Time series forecasting using a hybrid ARIMA and neural network model. *Neurocomputing*, 50, 159–175. [https://doi.org/10.1016/S0925-2312\(01\)00702-0](https://doi.org/10.1016/S0925-2312(01)00702-0)

Zhang, Y., Tan, W. H., & Zeng, Z. (2025). Tourism demand forecasting based on a hybrid temporal neural network model for sustainable tourism. *Sustainability*, 17(5), 2210. <https://doi.org/10.3390/su17052210>