E-ISSN : 3025-1311
https://ejournal.cahayailmubangsa.institute/index.php/kohesi

Kohesi: Jurnal Multidisiplin Saintek
Volume 11 No 6 Tahun 2026

PREDIKSI TREN PAYLOAD DAN JUMLAH PENGGUNA SUATU NETWORK PADA DAERAH SURABAYA DENGAN MENGGUNAKAN METODE *LONG SHORT TERM MEMORY* (LSTM)

Gusti Jogishwara Adji, Farell Nandana Tahir,
Shindi Shella May Wara, Andri Fauzan Adziima
Sains Data, UPN "Veteran" Jawa Timur
Jl. Raya Rungkut Madya, Gunung Anyar, Kota Surabaya, Jawa Timur, Indonesia

Email: 24083010107@student.upnjatim.ac.id

ABSTRAK

Perubahan volume trafik data pada jaringan seluler memerlukan peramalan untuk mendukung perencanaan kapasitas layanan. Penelitian ini bertujuan mengevaluasi model peramalan *payload* harian jaringan di wilayah operasional Surabaya untuk horizon 92 hari. Data yang digunakan terdiri atas 731 pengamatan harian pada periode 1 Januari 2024 hingga 31 Desember 2025. Tahapan penelitian mencakup validasi data, deteksi anomali, pembentukan 23 fitur historis dan musiman, serta pembagian data secara kronologis menjadi 639 hari pelatihan dan 92 hari pengujian. Lima kandidat model dibandingkan melalui tiga pembagian validasi pada *backtesting*, yaitu rata-rata 28 hari, musiman tujuh hari, Ridge, Random Forest, dan CatBoost. Berdasarkan hasil yang dilaporkan, Ridge memperoleh rata-rata *Total Absolute Percentage Error* terendah pada *backtesting* sebesar 4,37%. Pada pengujian Oktober–Desember 2025, model menghasilkan *Total Absolute Percentage Error* sebesar 0,24%, *Weighted Absolute Percentage Error* sebesar 3,99%, *Mean Absolute Percentage Error* sebesar 3,97%, dan *Root Mean Squared Error* sebesar 9.548,02 GB/hari. Hasil tersebut menunjukkan bahwa prediksi akumulasi *payload* mendekati nilai aktual, meskipun kesalahan harian tetap terjadi. Evaluasi pada periode tambahan dan pengembangan interval prediksi diperlukan untuk menilai konsistensi model dalam mendukung perencanaan kapasitas jaringan..

Kata kunci *: payload, jaringan seluler, peramalan deret waktu, regresi Ridge, backtesting, Surabaya*

Article History
Received: November 2026
Reviewed: November 2026
Published: November 2026

Plagiarism Checker
No 237.u873247.fh8374

Copyright : Author
Publish by : Kohesi

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License

1. PENDAHULUAN

Tingginya mobilitas dan aktivitas digital masyarakat di wilayah metropolitan seperti Surabaya memicu lonjakan jumlah pengguna aktif (*user*) serta volume trafik data (*payload*) yang sangat dinamis setiap harinya [1]. Dalam operasional jaringan telekomunikasi, fluktuasi *payload* dan jumlah pengguna merupakan indikator utama yang menentukan kualitas layanan (*Quality of Service*). Apabila lonjakan trafik pada jam sibuk maupun periode tertentu tidak diantisipasi dengan tepat, jaringan berisiko mengalami kelebihan beban (*overload*) dan kemacetan trafik (*network congestion*) yang menurunkan kecepatan akses pengguna [2]. Sebaliknya, penyediaan kapasitas berlebih tanpa perhitungan yang matang akan mengakibatkan inefisiensi biaya infrastruktur, sehingga diperlukan sistem prediksi tren jaringan yang akurat untuk mendukung perencanaan kapasitas (*capacity planning*).

Metode peramalan statistik konvensional seperti ARIMA umumnya kurang optimal dalam menangkap pola data trafik jaringan yang bersifat non-linier dan memiliki fluktuasi kompleks [3]. Sebagai solusi, metode *Long Short-Term Memory*(LSTM) yang merupakan pengembangan dari arsitektur *Recurrent Neural Network* (RNN) mampu mengatasi masalah *vanishing gradient*

melalui mekanisme sel memori dan gerbang informasi (*input, forget,* dan *output gate*), sehingga sangat efektif dalam mempelajari pola data deret waktu (*time-series*) berjangka panjang [4], [5]. Berdasarkan hal tersebut,penelitian ini bertujuan untuk mengimplementasikan dan mengevaluasi kinerja metode LSTM dalam memprediksi tren *payload* serta jumlah pengguna jaringan di wilayah Surabaya guna menghasilkan proyeksi trafik yang akurat sebagai dasar optimasi kapasitas jaringan.

2. TINJAUAN PUSTAKA

2.1. Trafik Jaringan Seluler dan Payload

Trafik jaringan seluler merupakan aktivitas pertukaran informasi melalui jaringan komunikasi seluler dalam periode tertentu. Pengamatan trafik diperlukan untuk memahami perubahan beban layanan jaringan. Dalam penelitian ini, *payload* didefinisikan secara operasional sebagai volume trafik data yang tercatat selama interval pengamatan. Pengukurannya perlu menetapkan satuan, arah trafik, serta tingkat agregasi, seperti *cell*, sektor, atau *site*. Riwayat *payload* dapat digunakan untuk mempelajari perubahan volume data dan memperkirakan nilainya pada periode berikutnya. Syam dan Aziz menggunakan data pemantauan pada tingkat *cell* untuk memprediksi kebutuhan *bandwidth* jaringan 4G, sehingga mendukung pemanfaatan data historis trafik dalam perencanaan kapasitas, meskipun *bandwidth* dan *payload* merupakan besaran berbeda [6].

2.2. Jumlah Pengguna Jaringan Seluler

Jumlah pengguna jaringan seluler dalam penelitian ini merupakan indikator pengguna berdasarkan definisi KPI pada interval pengamatan tertentu. Indikator tersebut dapat berupa rata-rata atau maksimum pengguna aktif sesuai variabel dalam dataset. Keduanya menggambarkan pengukuran yang berbeda sehingga target prediksi perlu ditetapkan secara konsisten. Nilainya tidak langsung menunjukkan jumlah pelanggan unik di wilayah penelitian. Hubungan indikator pengguna dengan *payload* perlu dianalisis karena kebutuhan trafik juga berkaitan dengan aktivitas pengguna. Syam dan Aziz menyebutkan jumlah pengguna, jenis aplikasi, dan waktu akses sebagai variabel tambahan untuk pengembangan prediksi kebutuhan *bandwidth*, sehingga mendukung eksplorasi statistik pengguna bersama perubahan trafik [6].

2.3. Prediksi Deret Waktu

Prediksi deret waktu merupakan proses memperkirakan nilai mendatang berdasarkan pengamatan historis yang tersusun menurut urutan waktu. Dalam penelitian ini, *payload* dan jumlah pengguna diperlakukan sebagai deret waktu dengan interval konsisten. Sejumlah pengamatan sebelumnya digunakan sebagai masukan untuk memprediksi nilai pada horizon yang ditetapkan. Panjang riwayat masukan dan horizon prediksi perlu dijelaskan agar tujuan pemodelan dapat dievaluasi secara terukur. Arissinta dkk. membandingkan ARIMA, LSTM, dan GRU untuk peramalan *web traffic*, yang memperlihatkan penerapan pendekatan statistik dan jaringan saraf pada data trafik berurutan. Penelitian tersebut menjadi rujukan metodologis bagi perbandingan model, tetapi hasil pada *web traffic* tidak dapat langsung digeneralisasikan terhadap trafik jaringan seluler [7].

2.4. Long Short-Term Memory

Long Short-Term Memory (LSTM) merupakan jaringan saraf berulang yang memanfaatkan memori untuk mempelajari data sekuensial. *Forget gate* mengatur informasi lama yang dipertahankan. *Input gate* mengatur penambahan informasi dari *candidate state*. Keduanya memperbarui *cell state*, sedangkan *output gate* mengatur pembentukan *hidden state*. Masukan penelitian ini dapat berupa riwayat *payload* dan jumlah pengguna. Persamaan (1)–(6) diadaptasi secara ekuivalen dari persamaan (1)–(10) pada Akbar dkk., dengan menggabungkan transformasi linear dan aktivasi [8].

$$f_t = \sigma(W_{xf}x_t + W_{hf}h_{t-1} + b_f) \quad (1)$$

$$i_t = \sigma(W_{xi}x_t + W_{hi}h_{t-1} + b_i) \quad (2)$$

$$g_t = \boldsymbol{tanh}\,(W_{xg}x_t + W_{hg}h_{t-1} + b_g) \quad (3)$$

$$C_t = f_t \odot C_{t-1} + i_t \odot g_t \quad (4)$$

$$o_t = \sigma(W_{xo}x_t + W_{ho}h_{t-1} + b_o) \quad (5)$$

$$h_t = o_t \odot \boldsymbol{tanh}\,(C_t) \quad (6)$$

Keterangan: x_t adalah masukan (*input*), h_t adalah *Hidden stat*, C_t adalah *cell state*, f_t adalah keluaran gerbang kelupaan (*forget gate*), i_t adalah keluaran gerbang masukan (*input gate*), o_t adalah keluaran gerbang keluaran (*output gate*), g_t adalah *candidate state*, W adalah matriks bobot, b adalah bias, σ adalah fungsi aktivasi Sigmoid, $tanh$ adalah fungsi aktivasi Tanh, $\odot$ adalah perkalian antarelemen, $t-1$ adalah indeks langkah sebelumnya

2.5. Fungsi Aktivasi

Fungsi aktivasi mengolah hasil transformasi linear dalam LSTM. Sigmoid menghasilkan nilai antara nol dan satu. Nilainya mengatur informasi melalui gerbang. Tanh menghasilkan nilai antara negatif satu dan satu. Fungsi ini digunakan pada *candidate state* dan pembentukan *hidden state*. Persamaan (7)-(8) mengacu pada persamaan (11)-(12) dalam Akbar dkk. [8].

$$\sigma(z) = \frac{1}{1+e^{-z}} \quad 7)$$

$$\boldsymbol{tanh}\,(z) = \frac{e^{z}-e^{-z}}{e^{z}+e^{-z}} \quad 8)$$

Keterangan: $\sigma(\boldsymbol{z})$ adalah fungsi aktivasi Sigmoid dengan masukan z, $\boldsymbol{tanh}\,(\boldsymbol{z})$ adalah fungsi aktivasi Tangent Hiperbolik (*hyperbolic tangent*) dengan masukan z, $\boldsymbol{z}$ adalah masukan (*input*) fungsi, $\boldsymbol{e}$ adalah bilangan euler / konstanta eksponensial (bernilai $\approx 2{,}718$), $-\boldsymbol{z}$ adalah pangkat eksponen negatif dari masukan z.

2.6. Normalisasi Min-Max

Normalisasi Min-Max menyesuaikan skala berdasarkan nilai minimum dan maksimum. Nilai dalam rentang acuan dipetakan menjadi nol hingga satu. Transformasi dilakukan per variabel. Selisih maksimum dan minimum harus tidak nol. Dalam rancangan penelitian ini, batas acuan ditentukan hanya dari data latih agar informasi pengujian tidak digunakan dalam penyiapan model. Persamaan (9) mengacu pada persamaan (11) dalam Khairunisa dan Hendikawati [9].

$$x' = \frac{x - x_{min}}{x_{max} - x_{min}} \quad 9)$$

Keterangan: $\boldsymbol{x}'$ adalah nilai hasil normalisasi, $\boldsymbol{x}$ adalah nilai awal (sebelum dinormalisasi), $\boldsymbol{x_{min}}$ adalah batas nilai minimum acuan, $\boldsymbol{x_{max}}$ adalah batas nilai maksimum acuan.

2.7. Root Mean Squared Error

Root Mean Squared Error (RMSE) mengukur akar rata-rata kuadrat kesalahan prediksi. Penguadratan memperbesar kontribusi kesalahan besar. Satuannya sama dengan variabel target. Nilai lebih kecil menunjukkan kesalahan lebih rendah pada pengujian yang sama. Kedua target dievaluasi secara terpisah pada skala aslinya. Persamaan (10) mengacu pada persamaan (12) dalam Khairunisa dan Hendikawati [9].

$$RMSE = \sqrt{\frac{1}{n}\sum_{t=1}^{n} (y_t - \hat{y}_t)^2} \qquad 10)$$

Keterangan: $RMSE$ adalah *Root Mean Squared Error* (akar kuadrat rata-rata galat/kesalahan prediksi), n adalah jumlah total data / sampel / pengamatan, $\sum_{t=1}^{n}$ adalah sumasi (penjumlahan akumulatif) dari data ke-$t = 1$ hingga data ke-n, y_t adalah nilai aktual (nilai sebenarnya) pada langkah waktu ke-t, $\hat{y}_t$ adalah nilai hasil prediksi (peramalan) pada langkah waktu ke-t, $(y_t - \hat{y}_t)^2$ adalah selisih kuadrat antara nilai aktual dan nilai prediksi.

2.8. Mean Absolute Percentage Error

Mean Absolute Percentage Error (MAPE) mengukur rata-rata kesalahan absolut relatif. Hasilnya berupa persentase. Nilai lebih kecil menunjukkan kesalahan relatif lebih rendah. MAPE tidak terdefinisi ketika nilai aktual nol. Nilai mendekati nol juga dapat memperbesar hasilnya. Persamaan (11) mengacu pada persamaan (13) dalam Khairunisa dan Hendikawati [9].

$$MAPE = \frac{100}{n}\sum_{t=1}^{n} \left|\frac{y_t - \hat{y}_t}{y_t}\right| \qquad 11)$$

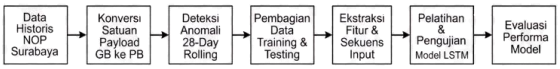
Keterangan: $\boldsymbol{MAPE}$ adalah *Mean Absolute Percentage Error* (Rata-rata Persentase Kesalahan Mutlak), $\boldsymbol{100}$ adalah Faktor pengali persentase, $\boldsymbol{n}$ adalah Jumlah pengamatan / total data, $\sum_{t=1}^{n}$ adalah Sumasi (penjumlahan) dari data ke-$\boldsymbol{t = 1}$ hingga data ke-$\boldsymbol{n}$, $|\cdots|$ adalah Nilai absolut / mutlak (memastikan nilai selisih selalu bernilai positif), $\boldsymbol{y_t}$ adalah Nilai aktual (sebenarnya) pada langkah waktu ke-$\boldsymbol{t}$, $\hat{\boldsymbol{y}}_{\boldsymbol{t}}$ adalah Nilai hasil prediksi pada langkah waktu ke-$\boldsymbol{t}$.

2.9. Penelitian Terdahulu

Rasyadan dkk. membandingkan LSTM, ARIMA, dan SARIMA untuk memprediksi jam sibuk trafik internet PT Comtelindo Balikpapan. Penelitian tersebut menggunakan horizon peramalan 28, 56, dan 84 hari. Pada horizon 28 hari, LSTM menghasilkan MAPE sebesar 3,34%, sedangkan kesalahannya meningkat pada horizon lebih panjang. Temuan tersebut menunjukkan pentingnya mempertimbangkan horizon dalam penilaian hasil prediksi. Berbeda dari fokus prediksi jam sibuk tersebut, penelitian ini diarahkan pada prediksi *payload* dan indikator jumlah pengguna jaringan seluler di Surabaya. Kontribusi yang diusulkan terletak pada evaluasi LSTM terhadap kedua target berdasarkan karakteristik data wilayah penelitian, bukan pada anggapan bahwa LSTM selalu lebih unggul daripada metode lain [10].

3. METODE PENELITIAN

Penelitian ini menggunakan pendekatan kuantitatif melalui eksperimen peramalan deret waktu pada data historis *payload* jaringan wilayah operasional Surabaya. Tahapan penelitian meliputi pengumpulan dan pemeriksaan data, deteksi anomali, pembentukan fitur historis dan musiman, pembagian data secara kronologis, pemilihan model melalui *backtesting*, serta evaluasi pada data uji. Data Januari 2024 hingga September 2025 digunakan untuk pengembangan model, sedangkan data Oktober hingga Desember 2025 digunakan sebagai periode pengujian akhir. Kandidat model yang dibandingkan meliputi rata-rata 28 hari, musiman tujuh hari, Ridge, Random Forest, dan CatBoost. Pemilihan model didasarkan pada rata-rata *Total Absolute Percentage Error* pada tiga pembagian validasi dalam *backtesting* dengan *Daily Weighted Absolute Percentage Error* sebagai pemutus seri. Model terpilih kemudian digunakan untuk menghasilkan prediksi *payload* selama 92 hari dan dievaluasi terhadap nilai aktual.



3.1 Pengumpulan dan Pemrosesan Data

Penelitian ini menggunakan dataset historis trafik jaringan harian di wilayah operasional (*NOP*) Surabaya selama dua tahun (1 Januari 2024 – 31 Desember 2025) yang mencakup 731 baris data tanggal transaksi, volume *payload* serta jumlah pengguna. Tahap pra-pemrosesan meliputi validasi kelengkapan rentang tanggal harian dan deteksi anomali (*outlier*) menggunakan metode *28-day rolling median*serta *Median Absolute Deviation* (MAD).

3.2.Feature Engineering

Ekstraksi fitur dilakukan untuk menangkap pola historis dan musiman pada data payload serta jumlah pengguna. Fitur yang digunakan meliputi lag historis pada interval 1, 2, 7, 14, 30, 60, dan 90 hari, statistik bergerak (rolling statistics) berupa rata-rata, standar deviasi, dan median dengan jendela 7–90 hari, serta pengkodean siklis (cyclical encoding) menggunakan fungsi sinus dan kosinus untuk merepresentasikan pola mingguan dan tahunan. Seluruh fitur numerik dinormalisasi menggunakan Min-Max Scaler ke rentang [0,1], kemudian disusun menjadi sekuens temporal sebagai masukan model LSTM.

3.3.Split Dataset untuk LSTM

Dataset dibagi secara kronologis menjadi data latih sebanyak 639 hari (1 Januari 2024 – 30 September 2025) dan data uji sebanyak 92 hari (1 Oktober – 31 Desember 2025). Pemodelan dilakukan menggunakan metode Long Short-Term Memory (LSTM) yang memanfaatkan mekanisme forget gate, input gate, dan output gate untuk mempelajari pola temporal pada data historis. Proses pelatihan menggunakan optimizer Adam dengan Mean Squared Error (MSE) sebagai fungsi kerugian untuk meminimalkan kesalahan prediksi payload dan jumlah pengguna.

3.4.Validasi Model Menggunakan Metode *Backtesting*

Validasi model dilakukan menggunakan pendekatan time series backtesting dengan tiga titik cutoff, yaitu 30 September 2024, 31 Maret 2025, dan 30 Juni 2025. Pada setiap tahap, model dilatih menggunakan data historis sebelum titik cutoffdan divalidasi terhadap 92 hari berikutnya. Pendekatan ini bertujuan untuk mengukur konsistensi kinerja model pada periode yang berbeda serta mencegah kebocoran data (data leakage) dengan mempertahankan urutan kronologis selama proses pelatihan dan validasi.

3.5 Evaluasi Model

Evaluasi kinerja model LSTM dilakukan dengan membandingkan hasil prediksi terhadap data aktual pada periode pengujian. Pengukuran menggunakan enam metrik, yaitu Total Absolute Percentage Error (Total APE) untuk mengukur kesalahan prediksi secara agregat, Daily Weighted Absolute Percentage Error (Daily WAPE) dan Daily Mean Absolute Percentage Error (Daily MAPE) untuk menilai kesalahan relatif harian, serta Bias (%) untuk mengidentifikasi kecenderungan prediksi berlebih atau kurang. Selain itu, Mean Absolute Error (MAE) dan Root Mean Squared Error(RMSE) digunakan untuk mengukur besarnya kesalahan absolut. Hasil evaluasi menjadi dasar penilaian akurasi dan konsistensi model dalam memprediksi tren payload serta jumlah pengguna jaringan.

4. HASIL DAN PEMBAHASAN

4.1 Pengumpulan dan Pemrosesan Data

Dataset terdiri atas 731 baris data harian volume *payload* NOP Surabaya (1 Januari 2024 – 31 Desember 2025). Hasil validasi menunjukkan rentang tanggal lengkap, tidak ada duplikasi, dan seluruh nilai bersifat numerik serta tidak negatif, sehingga tidak diperlukan imputasi maupun penghapusan baris. Pada Oktober 2024 – September 2025, total *payload* bulanan berkisar antara 5,42 PB (Februari 2025) dan 6,57 PB (Desember 2024).

Deteksi anomali dengan median bergulir 28 hari dan MAD menandai empat titik (Tabel 1), yaitu tiga hari berturut-turut pada 10–12 April 2024 dengan deviasi sekitar +27–29%, dan satu hari pada 22 September 2025 dengan penurunan 51,06%..

Tabel 1. Statistik Deskriptif

Tanggal	Payload (GB)	Deviasi (%)	Robust z
10 Apr 2024	235.994,58	+27,65	8,55
11 Apr 2024	236.607,36	+27,21	7,26
12 Apr 2024	241.724,04	+29,48	6,88
22 Sep 2025	98.575,70	−51,06	−10,79

4.2 Feature Engineering

Ekstraksi fitur menghasilkan 23 fitur numerik untuk setiap titik asal prakiraan (Tabel 2). Seluruh fitur dihitung hanya dari data hingga titik asal tersebut, sehingga tidak ada informasi masa depan yang masuk ke dalam masukan model.

Tabel 2. Fitur masukan model

Kelompok fitur	Rincian	Jumlah
Lag historis	1, 2, 7, 14, 30, 60, 90 hari	7
Statistik bergerak	Rata-rata, standar deviasi, median pada jendela 7, 30, 60, 90 hari	12
Pengkodean siklis	Sinus dan kosinus hari dalam minggu dan hari dalam tahun	4
Total		23

Setiap sampel latih berpasangan dengan target berupa vektor *payload* 92 hari berikutnya. Dengan syarat riwayat minimal 90 hari, data latih 639 hari menghasilkan 458 sampel (matriks fitur 458 × 23 dan matriks target 458 × 92). Pada tiga *fold* validasi, jumlah sampel latih

masing-masing 93, 275, dan 366. Seluruh label target pada setiap sampel berakhir paling lambat pada titik potong, sehingga kebocoran data dapat dihindari.

4.3. Hasil Split Dataset dan Backtesting

Dataset dibagi secara kronologis menjadi 639 hari data latih (1 Januari 2024 – 30 September 2025) dan 92 hari data uji (1 Oktober – 31 Desember 2025). Data uji tidak digunakan untuk pelatihan maupun pemilihan model. Pemilihan model dilakukan melalui *backtesting* dengan tiga titik potong pada Tabel 3.

Tabel 3. Konfigurasi fold backtesting

Fold	Data latih	Sampel latih	Periode validasi (92 hari)
1	1 Jan – 30 Sep 2024	93	1 Okt – 31 Des 2024
2	1 Jan 2024 – 31 Mar 2025	275	1 Apr – 1 Jul 2025
3	1 Jan 2024 – 30 Jun 2025	366	1 Jul – 30 Sep 2025

Lima kandidat model dibandingkan, yaitu rata-rata 28 hari terakhir, musiman 7 hari, *Ridge*, *Random Forest*, dan *CatBoost*. Peringkat dilihat dari rata-rata Total APE, dengan Daily WAPE sebagai pemutus seri.

Tabel 4. Peringkat model pada backtesting (rata-rata 3 fold)

Model	Total APE (%)	Daily WAPE (%)	Daily MAPE (%)
Ridge langsung 92 hari	4,37	5,99	6,01
Rata-rata 28 hari	4,81	7,46	7,52
CatBoost	5,93	7,55	7,43
Random Forest	6,33	7,45	7,26
Musiman 7 hari	7,07	10,00	10,18

Ridge terpilih karena memiliki rata-rata Total APE dan Daily WAPE terendah. Kinerjanya bervariasi antar *fold*: Total APE sebesar 8,85%, 0,25%, dan 3,99% pada *fold* 1, 2, dan 3. Kesalahan terbesar terjadi pada *fold* 1 dengan bias −8,85%, yaitu prediksi cenderung lebih rendah dari aktual. *Fold* ini hanya memiliki 93 sampel latih, sehingga keterbatasan data kemungkinan berpengaruh. Model rata-rata 28 hari yang sederhana hanya terpaut 0,44 poin persentase dan memiliki kesalahan terburuk yang lebih kecil. Dengan hanya tiga *fold* yang tumpang tindih, peringkat ini bersifat indikatif.

4.4. Hasil Prediksi dan Evaluasi pada Data Uji

Model *Ridge* dilatih menggunakan seluruh data hingga 30 September 2025, lalu memprediksi *payload* harian selama 92 hari (1 Oktober – 31 Desember 2025). Hasil prediksi dibandingkan dengan data aktual yang tidak digunakan untuk pelatihan maupun pemilihan model (Tabel 5).

Tabel 5. Metrik evaluasi pada data uji Q4 2025

Metrik	Nilai
Total APE	0,24%
Daily WAPE	3,99%
Daily MAPE	3,97%
Bias	−0,24%
MAE	7.893,38 GB/hari
RMSE	9.548,02 GB/hari

Total APE yang sangat kecil menunjukkan prediksi akumulasi kuartal mendekati aktual. Nilai Daily MAPE sebesar 3,97% menunjukkan kesalahan harian rata-rata sekitar 4% dari nilai aktual. Selisih antara keduanya terjadi karena kesalahan positif dan negatif saling meniadakan saat dijumlahkan. Bias −0,24% menunjukkan prediksi sedikit lebih rendah dari aktual secara keseluruhan.

Secara bulanan (Tabel 6), prediksi Oktober dan November lebih rendah dari aktual, sedangkan Desember lebih tinggi. Kesalahan terbesar terjadi pada Desember (4,85%) dan Oktober (4,47%).

Tabel 6. Prediksi dan aktual bulanan (PB)

Bulan	Prediksi	Aktual	APE (%)
Oktober 2025	5,66	5,93	4,47
November 2025	5,52	5,58	1,08
Desember 2025	6,13	5,84	4,85
Total Q4	17,30	17,34	0,24

Dengan data aktual Januari–September 2025 sebesar 53,10 PB, proyeksi *payload* tahunan 2025 adalah 70,40 PB, dengan APE tahunan 0,06% terhadap total aktual 70,44 PB. Kinerja pada data uji lebih baik daripada rata-rata *backtesting* (Total APE 4,37%; Daily WAPE 5,99%). Namun, evaluasi ini hanya berasal dari satu jendela 92 hari dan model belum

dilengkapi interval prediksi, sehingga hasilnya belum dapat dianggap konsisten pada periode lain.

5. KESIMPULAN DAN SARAN

Berdasarkan hasil evaluasi yang dilaporkan, model Ridge menghasilkan rata-rata *Total Absolute Percentage Error* terendah sebesar 4,37% pada tiga pembagian validasi dalam *backtesting* dibandingkan empat kandidat lainnya. Pada periode pengujian Oktober–Desember 2025, model memperoleh *Total Absolute Percentage Error* sebesar 0,24%, *Daily Weighted Absolute Percentage Error* sebesar 3,99%, *Daily Mean Absolute Percentage Error* sebesar 3,97%, serta *Root Mean Squared Error* sebesar 9.548,02 GB/hari. Nilai bias sebesar −0,24% menunjukkan bahwa prediksi akumulasi *payload* sedikit lebih rendah daripada nilai aktual. Kesalahan agregat yang lebih kecil daripada kesalahan harian menunjukkan bahwa penyimpangan positif dan negatif dapat saling meniadakan ketika dijumlahkan, sehingga evaluasi akumulasi perlu disertai penilaian pada tingkat harian. Hasil penelitian dapat menjadi bahan pertimbangan dalam perencanaan kapasitas jaringan, tetapi belum membuktikan konsistensi model di luar periode pengujian. Penelitian selanjutnya disarankan memperluas periode evaluasi, menambahkan interval prediksi, serta menguji *Long Short-Term Memory* (LSTM) dan indikator jumlah pengguna agar kemampuan model terhadap kedua target dapat dibandingkan secara terukur.

DAFTAR PUSTAKA

Asosiasi Penyelenggara Jasa Internet Indonesia (APJII), *Laporan Survei Penetrasi dan Perilaku Pengguna Internet Indonesia 2024*, Jakarta, Indonesia: APJII, 2024.

N. M. Aslam, I. U. Khan, S. A. AlQahtani, and N. Aslam, "Cellular traffic prediction for 5G and beyond networks using deep learning," *IEEE Access*, vol. 11, pp. 45120–45132, 2023.

M. K. M. Shapi, N. A. Ramli, and L. J. Awalin, "Time series forecasting of network traffic using ARIMA and deep learning approaches: A comparative study," *Journal of Communications*, vol. 17, no. 6, pp. 482–490, Jun. 2022.

A. R. Pratama, B. Setiawan, and D. R. Wijaya, "Prediksi trafik jaringan telekomunikasi seluler menggunakan metode Long Short-Term Memory (LSTM)," *Jurnal Nasional Teknik Elektro dan Teknologi Informasi (JNTETI)*, vol. 12, no. 2, pp. 115–122, Mei 2023.

Y. Liu, Z. Wang, and H. Chen, "Long-term mobile traffic forecasting based on LSTM neural networks for network capacity planning," in *Proc. IEEE Int. Conf. Communications (ICC)*, Denver, CO, USA, 2024, pp. 2110–2115.

R. F. Syam and F. Aziz, "Peramalan trafik untuk alokasi bandwidth jaringan seluler 4G menggunakan model hybrid ARIMA-LSTM," *Advances in Computer System Innovation Journal*, vol. 2, no. 2, pp. 78–84, 2024, doi: 10.51577/acsijournal.v2i2.585.

I. O. Arissinta, I. D. Sulistiyawati, D. Kurnianto, and I. Kharisudin, "Pemodelan time series untuk peramalan web traffic menggunakan algoritma Arima, LSTM, dan GRU," *PRISMA, Prosiding Seminar Nasional Matematika*, vol. 5, pp. 693–700, 2022. Accessed: Oct. 16, 2026. [Online]. Available: https://journal.unnes.ac.id/sju/prisma/article/view/54712

R. Akbar, R. Santoso, and B. Warsito, "Prediksi tingkat temperatur Kota Semarang menggunakan metode Long Short-Term Memory (LSTM)," *Jurnal Gaussian*, vol. 11, no. 4, pp. 572–579, Feb. 2023, doi: 10.14710/j.gauss.11.4.572-579.

N. Khairunisa and P. Hendikawati, "Long Short-Term Memory and Gated Recurrent Unit Modeling for Stock Price Forecasting," *Jurnal Matematika, Statistika dan Komputasi*, vol. 21, no. 1, pp. 321–333, Sep. 2024, doi: 10.20956/j.v21i1.35930.

M. T. Rasyadan, A. R. Utami, and B. Hasanah, "Penerapan metode LSTM, ARIMA dan SARIMA untuk memprediksi jam sibuk pada trafik internet," *InComTech: Jurnal Telekomunikasi dan Komputer*, vol. 16, no. 1, pp. 13–31, Apr. 2026, doi: 10.22441/incomtech.v16i1.27716.