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

Kohesi: Jurnal Multidisiplin Saintek
Volume 11 No 6 Tahun 2026

PREDIKSI PAYLOAD SELULER PER SITE SAAT NATAL DAN TAHUN BARU MENGGUNAKAN METODE UPLIFT DAN XGBOOST DI JAWA TIMUR

Baihaqi Nur Muhammad, Praja Lohphinesti Subiarto, Trimono, Muhammad Nasrudin
Sains Data, UPN "Veteran" Jawa Timur
Jl. Raya Rungkut Madya, Gunung Anyar, Kota Surabaya, Jawa Timur, Indonesia

Email: *24083010111@student.upnjatim.ac.id*

ABSTRAK

Periode Natal dan Tahun Baru (Nataru) mengubah trafik data seluler secara tidak seragam antarwilayah, sehingga operator memerlukan prediksi pada tingkat *site* untuk menentukan prioritas penambahan kapasitas. Prediksi tersebut harus disusun sekitar tiga bulan sebelum Nataru, sedangkan riwayat Nataru setiap *site* hanya mencakup empat musim. Penelitian ini bertujuan memprediksi rata-rata *payload* harian per *site* selama Nataru 2026/27 dan mengevaluasi kegunaannya dalam memilih *site* prioritas di Kota Surabaya, Jember, dan Jombang. Data yang digunakan adalah indikator kinerja harian 1.243 *site* pada periode Januari 2022 hingga September 2026. *Payload* Nataru dimodelkan sebagai level 28 hari terakhir dikalikan *uplift*, yaitu rasio *payload* Nataru terhadap *baseline* pra-Nataru. Logaritma *uplift* diprediksi dengan tujuh model, mulai dari *naive* musiman hingga XGBoost dan *ensemble*, kemudian dievaluasi melalui *backtest* berurutan waktu pada tiga musim (2023/24 hingga 2025/26). *Ensemble* menghasilkan *Weighted Absolute Percentage Error* (WAPE) rata-rata 0,101, lebih rendah daripada prediksi tanpa *uplift* (0,138), tetapi hanya sedikit di bawah *naive* musiman (0,104) dan berbeda signifikan berdasarkan uji Diebold-Mariano terkoreksi Holm pada satu dari tiga musim. Keunggulan model lebih jelas pada pemilihan prioritas: ketepatan sepuluh *site* teratas mencapai 0,867 (setara *naive* musiman) dibandingkan 0,756 tanpa *uplift*, dan Macro- F1 klasifikasi kuadran beban meningkat dari 0,66–0,68 menjadi 0,75–0,76. Sebaliknya, bias total per kabupaten berubah tanda antarmusim (−8,8% hingga +13,6%). Dengan demikian, model lebih andal untuk menentukan urutan *site* prioritas daripada untuk menetapkan total trafik, yang sebaiknya dibaca sebagai rentang.

Kata kunci : *Nataru; trafik data seluler; peramalan deret waktu; uplift; ensemble learning; backtesting*

ABSTRACT

The Christmas and New Year (Nataru) period changes mobile data traffic unevenly across regions, so operators need site-level forecasts to prioritise capacity expansion. Such forecasts must be prepared about three months before Nataru, whereas the Nataru history of each site covers only four seasons. This study aims to forecast the average daily payload per site during Nataru 2026/27 and to evaluate its usefulness for selecting priority sites in Surabaya City, Jember, and Jombang. The data are daily performance indicators of 1,243 sites from January 2022 to September 2026. Nataru payload is modelled as the last 28-day level multiplied by an uplift, defined as the ratio of Nataru payload to a pre-Nataru baseline. The log-uplift is predicted with seven models, ranging from seasonal naive to XGBoost and an ensemble, and evaluated through a time-ordered backtest over three seasons (2023/24 to

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

2025/26). The ensemble achieves a mean Weighted Absolute Percentage Error (WAPE) of 0.101, lower than forecasts without uplift (0.138) but only slightly below seasonal naive (0.104), with a significant difference under the Holm-corrected Diebold–Mariano test in one of three seasons. The model's advantage is clearer in priority selection: top-ten site precision reaches 0.867 (on par with seasonal naive) compared with 0.756 without uplift, and the Macro-F1 of load-quadrant classification rises from 0.66–0.68 to 0.75–0.76. In contrast, the district-level total bias changes sign across seasons (−8.8% to +13.6%). Therefore, the model is more reliable for ranking priority sites than for setting total traffic, which should be read as a range

Keywords *: Christmas and New Year; mobile data traffic; time series forecasting; uplift; ensemble learning; backtesting*

1. PENDAHULUAN

Prediksi trafik seluler menjadi dasar perencanaan kapasitas dan pengendalian kongesti jaringan [1], [2]. Kebutuhan ini meningkat pada periode khusus seperti Natal dan Tahun Baru (Nataru), ketika pola pemakaian menyimpang dari hari biasa [3]. Data operasional 38 kabupaten/kota di Jawa Timur pada empat musim Nataru (2022/23 hingga 2025/26) menunjukkan median kenaikan *payload* per *site* sebesar 6,4% hingga 9,7% terhadap *baseline* pra-Nataru, tetapi median per kabupaten berkisar dari −6,0% hingga +35,8%. Pada tiga wilayah penelitian, median rasio *payload* Nataru terhadap *baseline* berbeda tajam: 0,94–0,99 di Kota Surabaya, 1,04–1,10 di Jember, dan 1,12–1,18 di Jombang. Variasi ini menuntut prediksi pada tingkat *site* yang sudah tersedia sekitar tiga bulan sebelum Nataru, padahal setiap *site* hanya memiliki paling banyak empat musim Nataru sebagai riwayat.

Penelitian prediksi trafik seluler berkembang ke arah *deep learning*, antara lain *transformer* spasial-temporal untuk trafik tingkat kota [4], *Temporal Fusion Transformer* [2], dan model hierarkis untuk prediksi trafik 5G jangka panjang yang menangani ketidakstasioneran data [5], serta graf perhatian yang memodelkan hubungan spasial antar-stasiun pangkalan [6]. Di Indonesia, Syam dan Aziz menggabungkan ARIMA dan *Long Short-Term Memory* (LSTM) untuk memprediksi kebutuhan *bandwidth* tingkat *cell* [7], sedangkan Rasyadan dkk. melaporkan bahwa kinerja LSTM untuk trafik jam sibuk menurun pada horizon yang lebih panjang [8], dan Maliky dkk. melaporkan GRU paling akurat di antara SARIMA, *Decision Tree*, SVR, dan XGBoost pada trafik seluler [9]. Dari sisi metodologi peramalan, kompetisi M5 menunjukkan bahwa model berbasis pohon dan kombinasi prakiraan yang dilatih lintas banyak deret memberikan akurasi yang sangat baik [10], dan dominasi model pohon tersebut dibahas khusus oleh Januschowski dkk. [11]. Temuan ini sejalan dengan hasil bahwa model berbasis pohon tetap kompetitif pada data tabular berukuran sedang [12] dan bahwa kombinasi prakiraan mengurangi ketergantungan pada satu model terbaik [13].

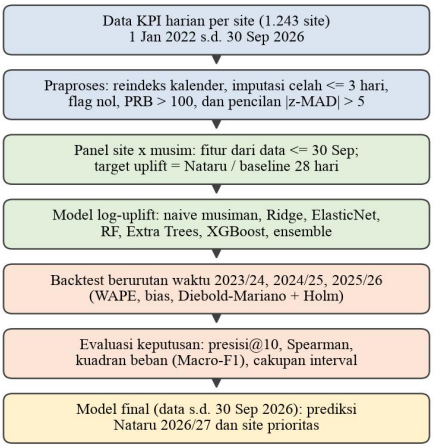
Pada literatur yang ditinjau, tiga hal belum banyak dijawab. Pertama, bagaimana memprediksi respons trafik setiap *site* terhadap kejadian tahunan yang jarang dengan horizon sekitar tiga bulan dan riwayat kejadian yang sangat terbatas. Kedua, apakah model *machine learning* benar-benar lebih baik daripada pembanding *naive* musiman setelah selisih galatnya diuji secara statistik, karena evaluasi yang kurang cermat dapat menghasilkan keunggulan semu [14]. Ketiga, apakah prediksi tersebut berguna untuk keputusan operasional, yaitu memilih *site* prioritas, dan bukan hanya akurat secara rata-rata.

Penelitian ini bertujuan memprediksi rata-rata *payload* harian per *site* selama Nataru 2026/27 di Kota Surabaya, Jember, dan Jombang serta mengevaluasi kegunaannya bagi perencanaan kapasitas. Pertanyaan penelitiannya adalah apakah pemodelan *uplift* dengan *machine learning* menurunkan galat secara signifikan dibandingkan *naive* musiman, dan seberapa baik prediksi tersebut memilih sepuluh *site* dengan beban terbesar serta

mengklasifikasikan kuadran beban. Kontribusinya adalah formulasi prediksi sebagai level *payload* terkini dikalikan *uplift* Nataru yang memungkinkan model belajar lintas lebih dari seribu *site*, protokol *backtest* berurutan waktu pada tiga musim dengan uji Diebold-Mariano terkoreksi Holm, dan evaluasi berbasis keputusan yang melaporkan bias agregat secara terbuka.

2. METODE PENELITIAN

Penelitian ini merupakan eksperimen prediktif kuantitatif pada data historis indikator kinerja (KPI) harian tingkat *site*. Alur penelitian ditunjukkan pada Gambar 1. Data dipraproses, disusun menjadi panel *site* × musim, dimodelkan sebagai *uplift* Nataru, dievaluasi melalui *backtest* berurutan waktu, lalu model final dipakai untuk memprediksi Nataru 2026/27.



Gambar 1. Alur penelitian

Pada Gambar 1, tahap praproses dan pembentukan panel memastikan setiap fitur hanya memakai data hingga 30 September, sekitar tiga bulan sebelum Nataru, sehingga skenario *backtest* meniru kondisi prediksi yang sebenarnya.

2.1. Data dan Praproses

Data berupa KPI harian per *site* dari lima berkas tahunan (1 Januari 2022 hingga 30 September 2026) untuk Kota Surabaya, Jember, dan Jombang. Setelah *site dummy* dikeluarkan, data berisi 2.053.792 baris dari 1.243 *site* (684 di Kota Surabaya, 371 di Jember, dan 188 di Jombang). Variabel yang dipakai adalah *payload* (MB), utilisasi *physical resource block* arah *downlink* (PRB, %), jumlah pengguna, koordinat *site*, kabupaten, dan tanggal. Praproses tidak membuang baris, tetapi memberi penanda: (i) kalender harian setiap *site* direindeks sehingga 8.696 baris hari hilang ditambahkan, termasuk 4–8 Januari 2024 yang kosong untuk semua *site*; (ii) *payload* atau pengguna bernilai nol (563 baris) dianggap hilang; (iii) PRB di atas 100% (31.627 baris, maksimum 128,6) ditandai dan dipotong pada 100% untuk analisis; (iv) celah hingga tiga hari berturut-turut (3.991 baris) diisi dengan interpolasi linier, sedangkan celah yang lebih panjang tidak diisi; dan (v) pencilan ditandai (4.081 baris) bila skor-z robust pada (1) melebihi 5, tanpa dihapus.

$$z_t = \frac{y_t - med_{28}(y)_t}{1{,}4826 MAD_{28}(y)_t} \qquad (1)$$

dengan *y* adalah *payload* harian, med_{28} dan MAD_{28} adalah median dan simpangan absolut median pada jendela bergulir 28 hari.

Nataru didefinisikan sebagai 24 Desember hingga 1 Januari (9 hari), sedangkan *baseline* adalah 28 hari penuh dari 20 November hingga 17 Desember pada musim yang sama. Empat musim memiliki data lengkap (2022/23 hingga 2025/26), dan musim 2026/27 menjadi target prediksi. Rata-rata Nataru dan *baseline* dihitung bila tersedia masing-masing minimal 7 dan 21 hari valid.

2.2. Formulasi Prediksi *Uplift*

Untuk *site i* pada musim *s*, misalkan *Y* adalah rata-rata *payload* harian selama Nataru dan *B* rata-rata *payload* pada *baseline*. *Uplift* τ didefinisikan sebagai

$$\tau_{i,s} = \frac{Y_{i,s}}{B_{i,s}} \qquad (2)$$

sehingga τ > 1 berarti Nataru menaikkan beban *site*. Prediksi harus dibuat pada 30 September, maka level terkini diukur sebagai *m*, yaitu rata-rata *payload* 28 hari terakhir hingga tanggal tersebut. *Payload* Nataru dimodelkan sebagai

$$\widehat{Y}_{i,s} = m_{i,s} \cdot exp\left(\widehat{g}\left(\boldsymbol{x}_{i,s}\right)\right) \qquad (3)$$

dengan *g* adalah fungsi yang memprediksi log-*uplift* dari vektor fitur x yang hanya dihitung dari data hingga 30 September. Pemodelan pada skala logaritma menjaga prediksi tetap positif dan membuat galat bersifat multiplikatif. Pendekatan ini memungkinkan satu model dipelajari lintas semua *site* meskipun setiap *site* hanya memiliki beberapa musim Nataru. Karena *m* diukur pada September sedangkan *B* pada November–Desember, formulasi ini mengasumsikan level September mewakili *baseline*. Target alternatif log(*Y*/*m*) diuji sebagai analisis sensitivitas, tetapi tidak konsisten antarmusim sehingga target *uplift* dipertahankan.

2.3. Fitur

Fitur dibentuk untuk setiap pasangan *site*–musim dan dikelompokkan seperti pada Tabel 1.

Tabel 1. Kelompok fitur model

Kelompok	Fitur	Jumlah
Level dan tren	*m* 28 dan 90 hari dan sejak 1 Januari; log rasio 28/90 hari dan 28 hari/sejak 1 Januari	5
PRB dan pengguna	Rata-rata PRB dan pengguna 28 hari	2
Volatilitas	Koefisien variasi 28 hari; rasio akhir pekan/hari kerja 90 hari	2
Spasial	Jarak ke pusat kabupaten; kode kabupaten	2
Site baru	Penanda *site* mulai setelah Maret 2022	1
Riwayat musim lalu	τ, rasio *Y*/*m*, dan selisih PRB Nataru terhadap PRB 28 hari musim lalu	3
Total		**15**

Tabel 1 memuat 15 fitur dalam enam kelompok. Fitur riwayat musim lalu tidak tersedia pada lipatan pertama sehingga lipatan tersebut memakai 12 fitur. Nilai hilang diisi median data latih untuk model linier dan model pohon acak, sedangkan XGBoost menangani nilai hilang secara internal. Fitur Ridge dan ElasticNet distandardisasi dengan statistik data latih

saja. Seluruh fitur dihitung dari data hingga 30 September musim yang sama, dan diperiksa bahwa keluarannya identik bila data setelah tanggal tersebut dihapus.

2.4. Model Prediksi

Tujuh model dibandingkan pada target log-*uplift*, ditambah satu pembanding tanpa *uplift*. Pembanding tersebut adalah *naive* ($\hat{Y} = m$). Model pertama adalah *naive* musiman, yaitu *m* dikalikan *uplift* musim lalu *site* yang sama. Model pembelajaran mesin meliputi Ridge dan ElasticNet (parameter regularisasi dipilih dengan validasi silang pada data latih), *Random Forest* dan *Extra Trees* (masing-masing 300 pohon dengan daun minimal 10 sampel), serta XGBoost [15] dengan 300 pohon, kedalaman 3, laju belajar 0,03, *subsampling* 0,8, *min_child_weight* 10, dan regularisasi L2 sebesar 5. Varian XGBoost yang turut memakai musim tanpa fitur riwayat juga diuji. Model ketujuh adalah *ensemble* rerata aritmetika *naive* musiman, Ridge, dan XGBoost,

$$\hat{Y}_{i,s}^{ens} = \frac{1}{3}\left(\hat{Y}_{i,s}^{xgb}\right) \qquad (4)$$

yang dipilih karena kombinasi prakiraan sederhana mengurangi ketergantungan pada satu model terbaik [13]. Penyetelan hiperparameter XGBoost (18 konfigurasi dengan validasi expanding *window*) menghasilkan konfigurasi yang sama dengan pengaturan awal, sehingga pengaturan awal dipakai. PRB Nataru diprediksi dengan XGBoost pada target selisih PRB Nataru terhadap rata-rata PRB 28 hari, dengan hasil dipotong pada 0–100%.

2.5. Validasi dan Evaluasi

Validasi memakai *backtest* berurutan waktu tiga lipatan dengan titik potong 30 September (Tabel 2). Model dilatih hanya dari musim sebelumnya dan diuji pada musim berikutnya tanpa pengacakan [14]. Titik potong tersebut meniru jarak prediksi sekitar tiga bulan. Musim 2022/23 tidak dapat diuji karena tidak memiliki musim sebelumnya.

Tabel 2. Konfigurasi *backtest*

Lipatan	Musim uji	Musim latih	n latih	n uji
1	2023/24	2022/23	1.141	1.146
2	2024/25	2023/24	1.146	1.168
3	2025/26	2023/24 dan 2024/25	2.314	1.195

Tabel 2 menunjukkan jumlah pasangan *site*–musim pada tiap lipatan. Lipatan 1 hanya memiliki satu musim latih tanpa fitur riwayat sehingga hasilnya paling terbatas. Model final dilatih pada musim 2023/24 hingga 2025/26, yaitu musim yang memiliki fitur riwayat, dengan data hingga 30 September 2026, lalu memprediksi 1.242 *site* untuk Nataru 2026/27. Model final dipilih dengan aturan rata-rata WAPE terendah pada tiga lipatan. Karena lipatan yang sama dipakai untuk memilih dan mengevaluasi, estimasi kinerja dapat sedikit terlalu optimistis.

Galat diukur dengan *Weighted Absolute Percentage Error* (WAPE) dan bias agregat,

$$WAPE = \frac{\sum_{i=1}^{n} |\hat{Y}_i|}{\sum_{i=1}^{n} Y_i} \qquad (5)$$

$$Bias = \frac{\sum_{i=1}^{n} (Y_i)}{\sum_{i=1}^{n} Y_i} \qquad (6)$$

dengan bias positif berarti prediksi melebihi nilai aktual. Selang kepercayaan 95% WAPE diperoleh dengan *bootstrap* 1.000 resampel *site*. Perbedaan galat antarmodel diuji dengan uji

Diebold–Mariano [16] dengan koreksi sampel kecil Harvey dkk. [17] pada galat absolut per *site*, terhadap *naive* musiman, dan seluruh 24 uji (8 model × 3 lipatan) dikoreksi dengan prosedur Holm. Statistik uji dihitung sebagai

$$DM = \frac{\bar{d}}{\sqrt{\frac{\hat{\gamma}_0}{n}}}, d_j = |e_{A,j}| - |e_{B,j}| \qquad (7)$$

dengan d adalah selisih galat absolut model A dan B pada *site* ke-j, dan γ_0 adalah ragamnya. Karena *site* diamati pada musim yang sama, uji ini mengandaikan galat antar-*site* saling bebas, sehingga ketergantungan spasial dapat membuat nilai p terlalu kecil. Perbedaan *uplift* terhadap 1 diuji dengan uji Wilcoxon bertanda peringkat per kabupaten dan musim (satuan analisis: *site*), disertai koreksi Holm.

Kegunaan keputusan dinilai dengan dua cara. Pertama, kualitas peringkat *site* per kabupaten dan musim melalui presisi10,

$$P10 = \frac{|\hat{T}_{10}|}{10} \qquad (8)$$

dengan T_{10} dan T_{10} adalah himpunan sepuluh *site* dengan *payload* aktual dan prediksi terbesar, serta korelasi Spearman antara peringkat aktual dan prediksi. Nilai rata-rata atas sembilan pasangan musim × kabupaten dibandingkan dengan pembanding level saja (m) memakai uji Wilcoxon bertanda peringkat. Kedua, klasifikasi kuadran *Traffic-Load Matrix*: dalam setiap kabupaten, $\log_{10}$ *payload* dan PRB dinormalkan menjadi skor-z, lalu ambang nol (rata-rata kabupaten) membagi *site* ke empat kuadran, yaitu *Critical* (*payload* dan PRB tinggi), *Efficiency Star* (*payload* tinggi, PRB rendah), *Congestion Risk* (*payload* rendah, PRB tinggi), dan *Under-optimized* (keduanya rendah). Kuadran prediksi XGBoost dibandingkan dengan kuadran aktual memakai Macro-F1,

$$Macro - F1 = \frac{1}{K}\sum_{k=1}^{K} \frac{2P_k R_k}{P_k + R_k} \qquad (9)$$

dengan P dan R adalah presisi dan *recall* kuadran ke-k dari K = 4 kuadran, serta koefisien kappa Cohen. Pembandingnya adalah kuadran musim lalu dan kuadran dari level terkini m tanpa model; evaluasi dilakukan pada musim uji 2024/25 dan 2025/26.

Selang prediksi 80% dibentuk dari kuantil ke-10 dan ke-90 residual log per kabupaten dengan gagasan kalibrasi *conformal* [18], dan diperlebar 1,5 kali untuk *site* baru. Pada evaluasi, selang dikalibrasi dari residual lipatan sebelumnya, sedangkan pada model final dari residual seluruh lipatan *backtest*. Cakupan selang (*prediction interval coverage probability*, PICP) diuji pada musim uji.

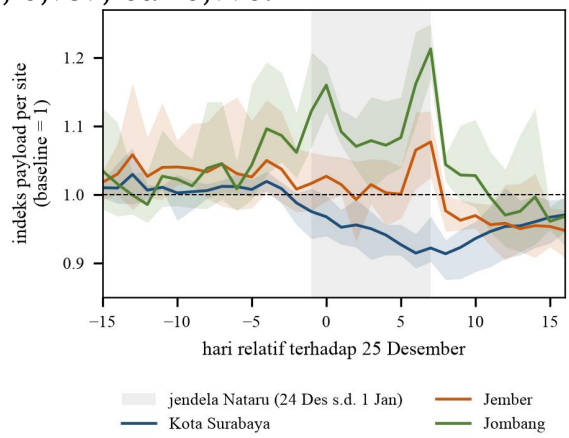
3. HASIL DAN PEMBAHASAN

Bab ini menyajikan hasil penelitian mengikuti urutan pertanyaan penelitian. Subbab 3.1 menggambarkan pola payload di sekitar Nataru pada ketiga kabupaten sebagai dasar empiris pemodelan uplift. Subbab 3.2 dan 3.3 menilai akurasi backtest pada tiga musim uji (2023/24 hingga 2025/26) beserta kontribusi fitur dan pilihan target, sedangkan subbab 3.4 menilai kegunaan prediksi untuk memilih site prioritas dan mengklasifikasikan kuadran beban. Subbab 3.5 menyajikan prediksi Nataru 2026/27, dan subbab 3.6 membahas temuan tersebut terhadap penelitian terdahulu beserta keterbatasannya

3.1. Pola *Payload* di Sekitar Nataru

Gambar 2 menunjukkan indeks *payload* harian per *site* terhadap *baseline* dari 15 hari sebelum hingga 16 hari sesudah 25 Desember. Ketiga kabupaten bereaksi berbeda. Rata-rata

indeks pada jendela Nataru (24 Desember hingga 1 Januari) adalah 0,945 di Kota Surabaya, 1,023 di Jember, dan 1,117 di Jombang. Jombang berada di atas *baseline* sepanjang jendela dan memuncak pada 1,21 di 1 Januari, Jember memuncak pada 1,08 di hari yang sama, sedangkan Kota Surabaya berada di bawah *baseline* selama seluruh jendela dengan indeks tertinggi hanya 0,975 dan terendah 0,915 pada 31 Desember. Pada 2–10 Januari, rata-rata indeks turun menjadi 0,947, 0,959, dan 0,996.



Gambar 2. Indeks *payload* harian di sekitar Nataru

Pita pada Gambar 2 cukup lebar sehingga besaran kenaikan berbeda antarmusim, tetapi urutan antarkabupaten konsisten. Median *uplift* per musim berkisar 0,94–0,99 di Kota Surabaya, 1,04–1,10 di Jember, dan 1,12–1,18 di Jombang, dan seluruhnya berbeda dari 1 (uji Wilcoxon, p Holm maksimum 5×10^{-5}). Urutan ini sama dengan urutan rasio *payload* akhir pekan terhadap hari kerja, yaitu 1,113 di Jombang, 1,036 di Jember, dan 0,976 di Kota Surabaya. Pada tingkat *site*, korelasi Spearman antara *uplift* dan rasio tersebut mencapai 0,68 (n = 4.732 pasangan *site*-musim, p Holm < 0,001). Ini adalah asosiasi, bukan hubungan sebab-akibat; dugaan bahwa wilayah dengan pemakaian akhir pekan lebih tinggi juga naik pada Nataru masuk akal, tetapi tidak dapat diuji dengan data yang tersedia.

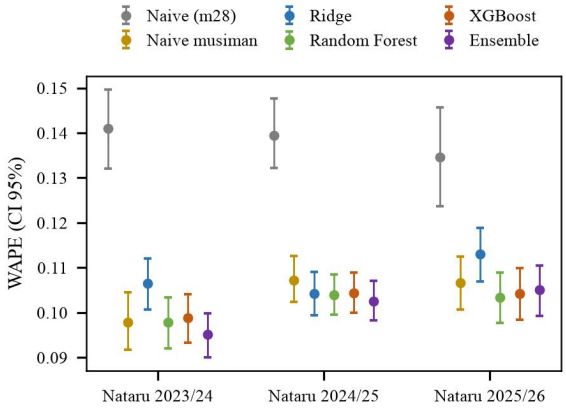
3.2. Akurasi *Backtest*

Tabel 3 memuat WAPE delapan model pada tiga musim uji. Semua model berbasis *uplift* memiliki WAPE rata-rata 0,101 hingga 0,108, yaitu 22% hingga 27% lebih rendah daripada *naive* tanpa *uplift* (0,1384). *Ensemble* menghasilkan rata-rata terendah (0,1009), disusul *Random Forest* (0,1017) dan XGBoost (0,1025), sedangkan *naive* musiman mencapai 0,1039.

Tabel 3. WAPE per model dan musim uji

Model	23/24	24/25	25/26	Rata-rata
Naive (*m*)	0,1410	0,1394	0,1347	0,1384
Naive musiman	0,0978	0,1072	0,1066	0,1039
Ridge	0,1064	0,1042	0,1131	0,1079
ElasticNet	0,1055	0,1057	0,1127	0,1079
Random Forest	0,0978	0,1040	0,1034	0,1017
Extra Trees	0,1025	0,1042	0,1041	0,1036
XGBoost	0,0989	0,1044	0,1043	0,1025
Ensemble	0,0952	0,1025	0,1051	0,1009

Ensemble terendah pada 2023/24 dan 2024/25, tetapi pada 2025/26 *Random Forest*, *Extra Trees*, dan XGBoost lebih rendah daripadanya (0,1034 untuk *Random Forest* dibanding 0,1051). Selisih antarmodel *uplift* kecil dibandingkan ketidakpastiannya. Gambar 3 memperlihatkan bahwa selang kepercayaan 95% *ensemble* dan *naive* musiman tumpang tindih di setiap musim, sedangkan selang *naive* tanpa *uplift* terpisah jelas dari semua model lain.



Gambar 3. WAPE dan selang kepercayaan 95%

Uji Diebold–Mariano dengan koreksi Holm menegaskan pembacaan tersebut. *Naive* tanpa *uplift* secara signifikan lebih buruk daripada *naive* musiman di ketiga musim (p Holm maksimum 2×10^{-7}), sehingga informasi *uplift* terbukti penting. Sebaliknya, *ensemble* hanya signifikan lebih baik daripada *naive* musiman pada 2024/25, dengan selisih rata-rata galat absolut 4.922 MB per *site* per hari (p Holm < 0,001). Pada 2023/24 dan 2025/26 selisihnya 2.361 dan 1.475 MB tidak signifikan (p Holm 1). Ridge dan ElasticNet justru lebih buruk daripada *naive* musiman pada 2025/26 (p Holm 0,015 dan 0,042), sedangkan *Random Forest*, *Extra Trees*, dan XGBoost tidak berbeda nyata di semua musim. Dengan demikian, keunggulan pembelajaran mesin atas *naive* musiman kecil dan tidak konsisten.

Bias agregat *ensemble* berganti tanda antarmusim, yaitu −3,3%, −6,6%, dan +6,6% pada 2023/24 hingga 2025/26. *Naive* musiman menunjukkan pola yang hampir sama (−2,7%, −6,9%, dan +6,1%), sehingga bias ini berasal dari pergeseran level seluruh wilayah, bukan dari kekhasan model. Per kabupaten, bias berkisar dari −8,8% hingga +13,6%. Dekomposisi galat log XGBoost menunjukkan bahwa guncangan bersama per kabupaten dan musim menyumbang 3,2%, 7,6%, dan 8,4% ragam galat. Jika guncangan itu dibuang, WAPE turun dari 0,104 menjadi 0,081 pada 2024/25 dan dari 0,104 menjadi 0,084 pada 2025/26. Komponen ini tidak teramati pada data hingga 30 September sehingga tidak dapat dikecilkan oleh model tingkat *site*. Dampak yang sama terlihat pada selang prediksi 80%: cakupannya hanya 0,72 (2024/25) dan 0,62 (2025/26), dan di Jember pada 2025/26 hanya 0,33.

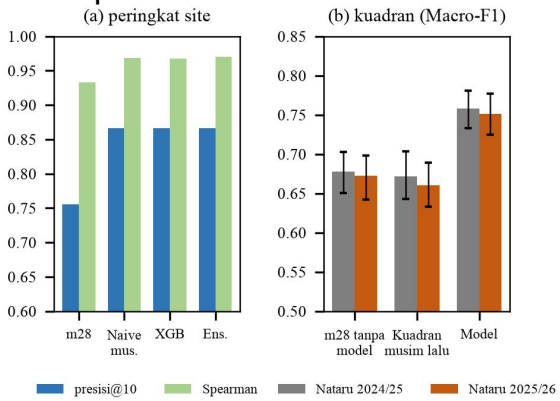
3.3. Kontribusi Fitur dan Target

Analisis ablasi per kelompok fitur menunjukkan bahwa penyumbang akurasi terbesar adalah volatilitas (kenaikan WAPE +0,0056 bila dibuang) dan level serta tren (+0,0026), sedangkan PRB dan pengguna (−0,0009) serta riwayat musim lalu (−0,0024) tidak memperbaiki WAPE *ensemble*. Seluruh efek berukuran di bawah 0,01. Temuan ini tampak berbeda dari *permutation importance* XGBoost yang menempatkan *uplift* musim lalu paling atas (kenaikan WAPE 0,023, dibanding 0,010 untuk rasio akhir pekan/hari kerja). Penjelasan yang mungkin, tetapi tidak diuji terpisah, adalah bahwa informasi riwayat sudah dibawa oleh komponen *naive* musiman dalam *ensemble*. Target alternatif $\log(Y/m)$ tidak konsisten: selisih WAPE *ensemble* terhadap target *uplift* adalah −0,0002, −0,0118, dan +0,0304 pada tiga musim, sehingga target *uplift* dipertahankan. Pergeseran tanda ini sejalan dengan ketidakstabilan level antarmusim.

3.4. Kegunaan untuk Keputusan

Meskipun selisih WAPE kecil, kualitas pemilihan *site* prioritas lebih informatif. Pada Gambar 4(a), *ensemble* menangkap rata-rata 8,7 dari 10 *site* teratas aktual (presisi@10 0,867) dengan Spearman 0,970, dibanding 0,756 dan 0,933 untuk pembanding level saja (m). Selisih ini signifikan pada uji Wilcoxon (p = 0,016 dan 0,004), tetapi hanya didasarkan pada sembilan

pasangan musim × kabupaten sehingga bersifat indikatif. *Naive* musiman dan XGBoost mencapai presisi@10 yang sama (0,867), sehingga keunggulan terutama berasal dari informasi *uplift* musim lalu, bukan dari kompleksitas model.



Gambar 4. Kualitas keputusan model

Pada Gambar 4(b), kuadran beban dari XGBoost mencapai Macro-F1 0,758 (selang 95% 0,733–0,781) pada 2024/25 dan 0,751 (0,725–0,777) pada 2025/26. Pembanding level saja memperoleh 0,677 dan 0,672, sedangkan kuadran musim lalu memperoleh 0,671 dan 0,660. Selang kepercayaan model tidak tumpang tindih dengan kedua pembanding, dan koefisien kappa Cohen naik dari 0,58 dan 0,58 menjadi 0,69 dan 0,70. Jadi, nilai praktis model terletak pada klasifikasi beban dan pemilihan *site* prioritas, bukan pada penurunan galat level.

3.5. Prediksi Nataru 2026/27

Model final memprediksi 1.242 *site*. Median *uplift* prediksi adalah 0,94 di Kota Surabaya, 1,03 di Jember, dan 1,13 di Jombang, mengikuti pola historis. Total rata-rata harian Nataru seluruh *site* diprediksi 1.092,4 TB (1 TB = 10^6 MB), sedangkan Tabel 4 hanya memuat *site* yang juga ada pada 2025/26 agar sebanding. Karena bias level agregat berganti tanda antarmusim, total ini dibandingkan dengan dua ekstrapolasi sederhana, yaitu total Nataru 2025/26 dikalikan (1 + pertumbuhan tahunan) dengan pertumbuhan dihitung dari Januari–September 2026 (Skenario A, kolom Sken. A) dan dari delapan minggu terakhir (Skenario B, kolom Sken. B), seperti pada Tabel 4, dengan Δ adalah selisih model terhadap aktual 2025/26.

Tabel 4. Total *payload* Nataru 2026/27 (TB/hari)

Wilayah	Aktual	Model	Δ (%)	Sken. A	Sken. B
Kota Surabaya	826,7	756,8	−8,5	771,4	765,0
Jember	227,5	220,0	−3,3	206,4	212,8
Jombang	119,2	113,6	−4,7	108,2	113,7
Total	**1.173,4**	**1.090,5**	**−7,1**	**1.086,0**	**1.091,4**

Model memprediksi perubahan total sebesar −8,5% di Kota Surabaya, −3,3% di Jember, dan −4,7% di Jombang terhadap Nataru 2025/26. Dibanding Skenario A, selisihnya −1,9%, +6,6%, dan +5,0%; dibanding Skenario B, −1,1%, +3,4%, dan 0,0%. Penurunan ini bukan artefak komposisi: pada *site* dengan cakupan data minimal 90%, *payload* Januari–September 2026 turun −6,7%, −9,3%, dan −9,2% dibanding tahun sebelumnya, dan 70% hingga 85% *site* mengalami penurunan. Penyebab pastinya tidak dapat dibuktikan karena data jaringan, tarif, dan perubahan definisi KPI tidak tersedia. *Site* dengan beban prediksi tertinggi adalah SBY348 (6,68 TB per hari) di Kota Surabaya, JBR036 (2,28 TB) di Jember, dan JMB113 (2,03 TB) di Jombang.

3.6. Pembahasan

Temuan utama penelitian ini adalah bahwa *uplift* musim lalu merupakan sumber informasi paling berharga, sedangkan kompleksitas model memberi tambahan kecil. Hasil ini

sejalan dengan bukti bahwa model berbasis pohon dan kombinasi prakiraan kompetitif pada data berukuran sedang [10], [12], [13], tetapi studi tersebut tidak menyiratkan bahwa model itu selalu mengalahkan pembanding musiman sederhana. Pada data ini, selisih WAPE 0,003 antara *ensemble* dan *naive* musiman tampak sebagai keunggulan jika hanya dilihat dari rata-rata, padahal uji terkoreksi hanya mendukungnya pada satu dari tiga musim. Hal ini sejalan dengan peringatan bahwa evaluasi tanpa uji statistik dapat menghasilkan keunggulan semu [14].

Bias yang berganti tanda dan cakupan selang di bawah 0,80 menunjukkan bahwa kesulitan utama bukan pada pemodelan antar-*site*, melainkan pada perubahan level antarmusim yang tidak terlihat pada 30 September. Kesulitan serupa dilaporkan pada prediksi trafik 5G jangka panjang, ketika ketidakstasioneran data menurunkan kinerja metode yang ada [5]. Karena target dan horizonnya berbeda, WAPE penelitian ini tidak dapat dibandingkan langsung dengan studi prediksi trafik kontinu. Model *deep learning* juga tidak diuji, dan dengan paling banyak tiga musim latih per *site* penelitian ini tidak dapat menyimpulkan perbandingannya. Implikasi operasionalnya, total prediksi sebaiknya dibaca sebagai rentang dengan ketidakpastian level sekitar −9% hingga +14% (kisaran bias historis), sedangkan peringkat dan kuadran *site* lebih dapat diandalkan untuk menentukan prioritas kapasitas.

Keterbatasan penelitian ini adalah sebagai berikut. Pertama, model dipilih dan dievaluasi pada tiga lipatan yang sama sehingga estimasi kinerja dapat sedikit optimistis. Kedua, uji Diebold–Mariano mengabaikan ketergantungan spasial antar-*site* sehingga nilai p cenderung terlalu kecil; karena kesimpulannya adalah keunggulan model kecil, kesimpulan itu justru menjadi lebih kuat. Ketiga, hanya tiga kabupaten di satu provinsi yang diteliti, sehingga pola pada Gambar 2 belum tentu berlaku di wilayah lain. Keempat, penurunan *payload* 2026 belum dapat dijelaskan dan memengaruhi keandalan prediksi level.

4. KESIMPULAN DAN SARAN

Penelitian ini memprediksi rata-rata *payload* harian per *site* selama Nataru 2026/27 di Kota Surabaya, Jember, dan Jombang dengan mengalikan level *payload* terkini dengan *uplift* Nataru yang diprediksi model. Pada *backtest* tiga musim, semua model berbasis *uplift* menurunkan WAPE sebesar 22% hingga 27% dibanding *naive* tanpa *uplift*, dan *ensemble* mencapai WAPE rata-rata terendah (0,1009). Keunggulannya atas *naive* musiman (0,1039) kecil dan hanya signifikan pada satu dari tiga musim. Nilai tambah model lebih jelas pada keputusan: kuadran beban dari XGBoost mencapai Macro-F1 0,758 dan 0,751, dibanding 0,677 dan 0,672 untuk pembanding level saja, sedangkan presisi@10 *ensemble* (0,867) setara *naive* musiman dan lebih tinggi daripada pembanding level (0,756). Bias agregat berganti tanda antarmusim dan cakupan selang 80% hanya 0,72 dan 0,62, sehingga total prediksi 2026/27 sebesar 1.092,4 TB per hari (perubahan −7,1% terhadap aktual 2025/26 pada *site* yang sebanding) perlu dibaca sebagai perkiraan dengan ketidakpastian level.

Kontribusi penelitian ini adalah formulasi prediksi sebagai level *payload* dikalikan *uplift* yang memungkinkan satu model dipelajari lintas lebih dari seribu *site*, protokol *backtest* berurutan waktu dengan uji Diebold–Mariano terkoreksi Holm, serta evaluasi berbasis keputusan yang melaporkan bias secara terbuka. Implikasinya, perencana kapasitas dapat memakai peringkat dan kuadran *site* untuk menentukan prioritas, sedangkan total prediksi sebaiknya diperlakukan sebagai rentang. Penelitian lanjutan disarankan untuk: (i) menguji model pada 38 kabupaten dan musim yang lebih panjang agar generalisasinya dapat dinilai; (ii) memasukkan data jaringan, tarif, dan kalender untuk menjelaskan pergeseran level antarmusim, termasuk penurunan *payload* 2026; (iii) memperbarui prediksi level dengan data Oktober–November sebelum Nataru; dan (iv) membandingkan model *deep learning* serta uji yang memperhitungkan ketergantungan spasial bila riwayat sudah memadai.

DAFTAR PUSTAKA

W. Jiang, "Cellular traffic prediction with machine learning: A survey," *Expert Systems with Applications*, vol. 201, Art. no. 117163, 2022, doi: 10.1016/j.eswa.2022.117163.

G. Kougioumtzidis, V. K. Poulkov, P. I. Lazaridis, and Z. D. Zaharis, "Mobile network traffic prediction using temporal fusion transformer," *IEEE Transactions on Artificial Intelligence*, vol. 6, no. 10, pp. 2685–2699, 2025, doi: 10.1109/TAI.2025.3556627.

E. Obiodu *et al.*, "How special is New Year eve traffic? Insights from four years 3G/4G/5G user measurements," in *Proc. 45th Int. Conf. Telecommunications and Signal Processing (TSP)*, 2022, pp. 349–354, doi: 10.1109/TSP55681.2022.9851299.

B. Gu, J. Zhan, S. Gong, W. Liu, Z. Su, and M. Guizani, "A spatial-temporal transformer network for city-level cellular traffic analysis and prediction," *IEEE Transactions on Wireless Communications*, vol. 22, no. 12, pp. 9412–9423, 2023, doi: 10.1109/TWC.2023.3270441.

Y. Yang *et al.*, "Long term 5G network traffic forecasting via modeling non-stationarity with deep learning," *Communications Engineering*, vol. 2, no. 1, Art. no. 33, 2023, doi: 10.1038/s44172-023-00081-4.

X. Chen, G. Chuai, K. Zhang, and W. Gao, "Spatial-temporal cellular traffic prediction: A novel method based on causality and graph attention network," in *Proc. IEEE Wireless Communications and Networking Conf. (WCNC)*, 2023, pp. 1–6, doi: 10.1109/WCNC55385.2023.10118616.

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.

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, 2026, doi: 10.22441/incomtech.v16i1.27716.

A. Maliky, Iskandar, and H. Yuliana, "Cellular network traffic prediction using machine learning and gate recurrent unit algorithm," in *Proc. 11th Int. Conf. Wireless and Telematics (ICWT)*, 2025, pp. 1–5, doi: 10.1109/ICWT66752.2025.11181848.

S. Makridakis, E. Spiliotis, and V. Assimakopoulos, "M5 accuracy competition: Results, findings, and conclusions," *International Journal of Forecasting*, vol. 38, no. 4, pp. 1346–1364, 2022, doi: 10.1016/j.ijforecast.2021.11.013.

T. Januschowski, Y. Wang, K. Torkkola, T. Erkkilä, H. Hasson, and J. Gasthaus, "Forecasting with trees," *International Journal of Forecasting*, vol. 38, no. 4, pp. 1473–1481, 2022, doi: 10.1016/j.ijforecast.2021.10.004.

L. Grinsztajn, E. Oyallon, and G. Varoquaux, "Why do tree-based models still outperform deep learning on typical tabular data?," in *Advances in Neural Information Processing Systems 35 (NeurIPS 2022)*, 2022, pp. 507–520, doi: 10.52202/068431-0037.

X. Wang, R. J. Hyndman, F. Li, and Y. Kang, "Forecast combinations: An over 50-year review," *International Journal of Forecasting*, vol. 39, no. 4, pp. 1518–1547, 2023, doi: 10.1016/j.ijforecast.2022.11.005.

H. Hewamalage, K. Ackermann, and C. Bergmeir, "Forecast evaluation for data scientists: Common pitfalls and best practices," *Data Mining and Knowledge Discovery*, vol. 37, no. 2, pp. 788–832, 2023, doi: 10.1007/s10618-022-00894-5.

T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in *Proc. 22nd ACM SIGKDD Int. Conf. Knowledge Discovery and Data Mining*, 2016, pp. 785–794, doi: 10.1145/2939672.2939785.

F. X. Diebold and R. S. Mariano, "Comparing predictive accuracy," *Journal of Business & Economic Statistics*, vol. 13, no. 3, pp. 253–263, 1995, doi: 10.1080/07350015.1995.10524599.

D. Harvey, S. Leybourne, and P. Newbold, “Testing the equality of prediction mean squared errors,” *International Journal of Forecasting*, vol. 13, no. 2, pp. 281–291, 1997, doi: 10.1016/S0169-2070(96)00719-4.

A. N. Angelopoulos and S. Bates, “Conformal prediction: A gentle introduction,” *Foundations and Trends in Machine Learning*, vol. 16, no. 4, pp. 494–591, 2023, doi: 10.1561/2200000101.