

Penerapan Algoritma XGBoost untuk Prediksi Harga Sewa Bus Pariwisata pada PT. Setio Budi Luhur Trans

Daniel Prasetyo Budiman¹, Retno Mumpuni^{2*}, Afina Lina Nurlaili³

^{1,2,3} Program Studi Informatika, Fakultas Ilmu Komputer, Universitas Pembangunan Nasional Jawa Timur

¹122081010304@student.upnjatim.ac.id

³afina.lina.if@upnjatim.ac.id

*Corresponding author email: retnomumpuni.if@upnjatim.ac.id

Abstrak— Penentuan harga sewa bus pariwisata di PT. Setio Budi Luhur Trans selama ini dilakukan secara subjektif tanpa standar perhitungan yang konsisten, sehingga berpotensi menimbulkan ketidaktepatan harga dan menurunkan daya saing perusahaan. Penelitian-penelitian sebelumnya mengenai prediksi harga pada sektor transportasi umumnya berfokus pada tarif taksi atau ride-hailing dengan dataset bervolume besar dan fitur yang beragam, sementara penelitian tentang prediksi harga sewa bus pariwisata dengan dataset terbatas pada konteks perusahaan lokal masih belum banyak dieksplorasi. Kesenjangan inilah yang menjadi landasan penelitian ini. Penelitian ini mengusulkan penerapan algoritma Extreme Gradient Boosting (XGBoost) untuk memprediksi harga sewa bus secara otomatis berdasarkan tiga variabel utama, yaitu jarak perjalanan, jenis bus, dan kondisi musim atau hari libur. Dataset yang digunakan berjumlah 196 data historis penyewaan bus yang diperoleh langsung dari PT. Setio Budi Luhur Trans. Tahapan penelitian meliputi pengumpulan data, praproses data (transformasi atribut kategorikal menjadi numerik dan penghitungan jarak), penerapan algoritma XGBoost secara iteratif melalui mekanisme boosting, serta evaluasi model menggunakan metrik Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), Root Mean Square Error (RMSE), dan koefisien determinasi (R^2). Hasil penelitian menunjukkan bahwa algoritma XGBoost mampu menghasilkan model prediksi harga sewa bus dengan akurasi rata-rata sebesar 97,08%, MAPE rata-rata 2,92%, dan nilai R^2 rata-rata 0,9422, sehingga dapat diintegrasikan ke dalam sistem informasi manajemen bus pariwisata berbasis web. Implementasi ini diharapkan dapat meningkatkan efisiensi operasional dan transparansi penetapan harga bagi perusahaan maupun pelanggan.

Kata Kunci— XGBoost, Prediksi Harga, Bus Pariwisata, Machine Learning, Sistem Informasi Manajemen.

I. PENDAHULUAN

P Perkembangan teknologi informasi saat ini telah mencapai kemajuan yang sangat pesat dan signifikan di berbagai sektor, termasuk industri transportasi pariwisata [1]. Hal ini ditandai dengan pemanfaatan teknologi komputer yang kini menjadi kebutuhan mendasar di hampir segala bidang sebagai alat bantu utama untuk meningkatkan efisiensi operasional [2]. PT. Setio Budi Luhur Trans merupakan perusahaan penyedia layanan penyewaan bus pariwisata yang menghadapi tantangan dalam penetapan harga sewa secara konsisten dan objektif. Selama ini, penentuan harga dilakukan secara manual berdasarkan pengalaman subjektif administrasi atau kesepakatan langsung

dengan pelanggan, tanpa mempertimbangkan faktor-faktor yang berpengaruh secara sistematis [3].

Permasalahan utama yang dihadapi adalah tidak adanya standar perhitungan harga yang mempertimbangkan variabel seperti jarak perjalanan, jenis bus, dan kondisi musim atau hari libur secara bersamaan. Kondisi ini berpotensi menyebabkan harga yang ditetapkan tidak kompetitif, merugikan perusahaan, atau tidak memuaskan pelanggan [4]. Lebih jauh, belum tersedianya model prediktif yang dapat diintegrasikan ke dalam sistem manajemen membuat proses penetapan harga tetap bergantung pada faktor subjektif dan tidak terstandarisasi.

Penerapan machine learning, khususnya algoritma Extreme Gradient Boosting (XGBoost), menjadi solusi yang relevan untuk permasalahan prediksi harga. XGBoost dikembangkan oleh Chen dan Guestrin [5] sebagai penyempurnaan dari Gradient Boosting Machine (GBM) dengan penambahan mekanisme regularisasi yang efektif. Dibandingkan dengan algoritma lain seperti Random Forest dan LightGBM, XGBoost menawarkan keseimbangan optimal antara kecepatan komputasi, akurasi prediksi, dan kemampuan menangani data berukuran kecil tanpa risiko overfitting yang tinggi [6].

Meskipun sistem penyewaan berbasis web telah terbukti efektif dalam manajemen operasional [7][8][9], mayoritas penelitian tersebut berfokus pada fungsionalitas sistem tanpa menyertakan mekanisme otomatisasi penetapan harga berbasis data. Selain itu, penelitian prediksi harga di sektor transportasi yang ada umumnya menggunakan dataset berskala besar dengan fitur yang kompleks, sehingga kurang relevan diterapkan langsung pada perusahaan bus pariwisata lokal dengan skala data yang lebih terbatas. Hal ini menciptakan kesenjangan (research gap) yang signifikan: belum adanya model prediksi harga sewa bus pariwisata yang secara spesifik dirancang untuk kondisi perusahaan lokal di Indonesia dengan keterbatasan data historis. Berdasarkan uraian tersebut, rumusan masalah dalam penelitian ini adalah: bagaimana menerapkan algoritma XGBoost untuk memprediksi harga sewa bus pariwisata pada PT. Setio Budi Luhur Trans secara akurat berdasarkan fitur-fitur yang tersedia, dan seberapa baik performa model yang dihasilkan? Penelitian ini berfokus pada penerapan model XGBoost untuk memprediksi harga sewa bus, yang kemudian diintegrasikan langsung ke dalam arsitektur sistem informasi manajemen berbasis web, guna mencapai tingkat efisiensi manajemen yang terotomatisasi secara menyeluruh.

II. TINJAUAN PUSTAKA

A. Prediksi Harga

Prediksi harga merupakan proses memperkirakan nilai harga di masa mendatang berdasarkan data historis dan variabel-variabel yang memengaruhinya. Dalam konteks bisnis jasa transportasi, prediksi harga yang akurat membantu perusahaan menetapkan tarif yang kompetitif dan berbasis data, sekaligus memberikan transparansi bagi pelanggan [10]. Berbagai pendekatan machine learning telah digunakan dalam prediksi harga, mulai dari regresi linear, Support Vector Regression (SVR), hingga metode ensemble seperti Random Forest dan Gradient Boosting [11]. Setiap metode memiliki keunggulan dan keterbatasan tersendiri tergantung pada karakteristik data yang digunakan [5]. Pada sektor transportasi khususnya, faktor-faktor seperti jarak tempuh, tipe kendaraan, jumlah penumpang, dan kondisi waktu perjalanan (musim liburan atau hari biasa) secara signifikan mempengaruhi penentuan harga sewa. Oleh karena itu, model prediksi yang mampu menangkap hubungan non-linear antara variabel-variabel tersebut dengan harga sewa menjadi sangat penting untuk menghasilkan estimasi yang akurat dan dapat diandalkan oleh perusahaan transportasi pariwisata.

B. Algoritma XGBoost

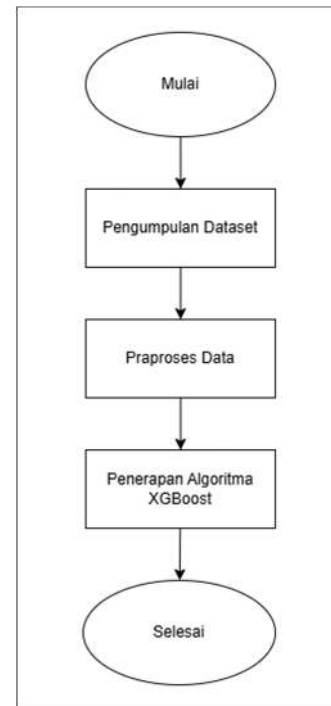
Extreme Gradient Boosting (XGBoost) adalah metode machine learning berbasis ensemble yang menggabungkan beberapa model Decision Tree secara sekuensial [5]. Setiap pohon keputusan dibangun untuk mengoreksi kesalahan dari pohon sebelumnya menggunakan pendekatan gradient descent. Keunggulan XGBoost antara lain mekanisme regularisasi L1 (Lasso) dan L2 (Ridge) yang mencegah overfitting, efisiensi komputasi yang tinggi melalui teknik approximate greedy algorithm dan weighted quantile sketch, serta kemampuan analisis feature importance yang membantu mengidentifikasi variabel paling berpengaruh terhadap prediksi [6]. XGBoost juga mendukung pemrosesan paralel pada level pembangunan pohon, menjadikannya jauh lebih cepat dibandingkan implementasi Gradient Boosting konvensional. Selain itu, XGBoost memiliki mekanisme penanganan missing values secara otomatis, sehingga lebih robust terhadap ketidaklengkapan data. Dibandingkan metode ensemble lainnya seperti Random Forest yang membangun pohon secara paralel dan independen, XGBoost membangun pohon secara sekuensial sehingga setiap pohon baru berfokus pada kesalahan yang belum terkoreksi oleh pohon-pohon sebelumnya. Dibandingkan LightGBM, XGBoost lebih stabil pada dataset berukuran kecil karena menggunakan level-wise tree growth, sementara LightGBM menggunakan leaf-wise growth yang cenderung overfit pada data kecil. Karakteristik ini menjadikan XGBoost pilihan yang lebih tepat untuk penelitian ini yang menggunakan dataset sebanyak 196 sampel.

Fungsi objektif XGBoost terdiri dari fungsi loss yang mengukur kesalahan prediksi dan istilah regularisasi yang

mengontrol kompleksitas model. Untuk kasus regresi dengan Mean Squared Error (MSE), turunan pertama (gradient) dan turunan kedua (hessian).

III. METODOLOGI PENELITIAN

Metodologi penelitian yang digunakan dalam penelitian ini terdiri dari tiga tahapan utama yang dilakukan secara sistematis dan terstruktur. Setiap tahapan dirancang untuk mendukung proses penelitian secara komprehensif sehingga tujuan penelitian dapat tercapai dengan baik. Berikut ditampilkan metodologi yang digunakan dalam penelitian ini:



Gbr. 1 Diagram Alur Metodologi Penelitian.

A. Dataset

Dataset yang digunakan merupakan data historis harga penyewaan bus pariwisata dari PT. Setio Budi Luhur Trans, berjumlah 196 baris data. Data mentah mencakup lokasi penjemputan, destinasi tujuan, jenis bus, kondisi musim, dan harga sewa. Contoh data awal ditunjukkan pada Tabel I.

TABEL I. CONTOH DATASET HARGA PENYEWAAN BUS

Penjemputan	Destinasi	Jenis Bus	Musim	Harga (Rp)
Surabaya	Tuban	Medium	Normal	2.300.000
Surabaya	Tuban	Medium	Libur	2.500.000
Surabaya	Tuban	Big	Normal	3.300.000
Surabaya	Gresik	Medium	Normal	1.600.000
Surabaya	Gresik	Big	Libur	2.700.000

B. Praproses Data

Tahapan praproses dilakukan untuk mengubah data mentah menjadi format yang siap diproses oleh XGBoost. Langkah-langkah yang dilakukan adalah:

1) Seleksi atribut: Atribut lokasi penjemputan dan destinasi dieliminasi dan digantikan dengan atribut jarak perjalanan dalam satuan kilometer, yang dihitung menggunakan layanan pemetaan digital.

2) Transformasi kategorikal: Variabel jenis bus diubah menjadi numerik (0 = Medium, 1 = Big) dan variabel musim/libur diubah menjadi (0 = Normal, 1 = Libur).

Dataset final memiliki empat atribut: jarak (integer, km), jenis_bus (integer, 0/1), musim_libur (integer, 0/1), dan harga (integer, Rupiah). Contoh dataset final ditunjukkan pada Tabel II.

TABEL II. CONTOH DATASET FINAL

Jarak (km)	Jenis Bus	Musim	Harga (Rp)
200	0	0	2.300.000
200	0	1	2.500.000
200	1	0	3.300.000
40	0	0	1.600.000
120	0	1	2.200.000

C. Penerapan Algoritma XGBoost

Algoritma XGBoost menerapkan mekanisme boosting secara iteratif, di mana setiap pohon keputusan yang dibentuk bertujuan untuk mengoreksi kesalahan prediksi yang dihasilkan oleh model sebelumnya. Pada contoh penerapan ini, algoritma XGBoost akan memiliki parameter input sebagai berikut:

- Jarak : 150
- Jenis bus : 0 (medium bus)
- Musim : 0 (tidak libur)

Dengan contoh dataset implementasi sebagai berikut ini:

TABEL III. CONTOH DATASET IMPLEMENTASI

Jarak (km)	Jenis Bus	Musim	Harga (Rp)
200	0	0	2.300.000
200	0	1	2.500.000
200	1	0	3.300.000
200	1	1	3.500.000
40	0	0	1.600.000
40	0	1	1.800.000
40	1	0	2.500.000
40	1	1	2.700.000
120	0	0	2.000.000
120	0	1	2.200.000

Proses pelatihan model ini berlangsung melalui enam tahapan utama sebagai berikut.

1) Inisialisasi prediksi awal: Model memberikan prediksi awal berupa rata-rata seluruh nilai target. Dari 10 sampel data contoh diperoleh :

$$\hat{y}^{(0)} = \frac{1}{n} \sum_{i=1}^n y_i$$

$$\hat{y}^{(0)} = 2.440.000$$

2) Perhitungan Gradient dan Hessian

Setelah prediksi awal diperoleh, langkah selanjutnya adalah menghitung nilai gradient dan hessian untuk setiap data. Untuk fungsi loss MSE, rumus yang digunakan adalah:

$$g_i = \hat{y}_i - y_i \quad (\text{gradient / turunan pertama})$$

$$h_i = 1 \quad (\text{hessian / turunan kedua})$$

Nilai gradient menggambarkan arah dan besaran kesalahan prediksi terhadap nilai aktual. Hasil ini digunakan sebagai simpul induk (parent node). Score parent dihitung sebagai:

$$Score_{parent} = \frac{(\sum g_i)^2}{\sum h_i + \lambda}$$

$$Score_{parent} = \frac{(0)^2}{10 + 1} = 0$$

Nilai Score_parent = 0 karena jumlah seluruh gradient saling menghapus ($\sum g_i = 0$). Selanjutnya nilai ini menjadi acuan perhitungan gain.

TABEL V. CONTOH PERHITUNGAN GRADIENT DAN HESSIAN

No	y_i	$\hat{y}_i^{(0)}$	$g_i = \hat{y}_i - y_i$	$h_i^{(0)}$
1	2.300.000	2.440.000	+140.000	1
2	2.500.000	2.440.000	-60.000	1
3	3.300.000	2.440.000	-860.000	1
4	3.500.000	2.440.000	-1.060.000	1
5	1.600.000	2.440.000	+840.000	1
6	1.800.000	2.440.000	+640.000	1
7	2.500.000	2.440.000	-60.000	1
8	2.700.000	2.440.000	-260.000	1
9	2.000.000	2.440.000	+440.000	1
10	2.200.000	2.440.000	+240.000	1

3) Menghitung Score Split

Pada setiap node, score dihitung dengan formula:

$$Score = \frac{(\sum g_i)^2}{(\sum h_i + \lambda)}$$

di mana λ adalah parameter regularisasi L2. XGBoost mengevaluasi tiga kandidat split berikut:

a) Split pada Jarak < 150

- Node kiri (data no. 5,6,7,8,9,10) $\rightarrow \Sigma gi = +1.840.000$

$$Score_{left} = \frac{(1.840.000)^2}{(6 + 1)} = 483.657.143$$

- Node kanan (data no. 1,2,3,4) $\rightarrow \Sigma gi = -1.780.000$

$$Score_{right} = \frac{(-1.780.000)^2}{(4 + 1)} = 634.480.$$

b) Split pada Jenis Bus

- Node kiri (data no. 1,2,5,6,9,10) $\rightarrow \Sigma gi = +1.840.000$

$$Score_{left} = \frac{(1.840.000)^2}{(6 + 1)} = 716.800.000$$

- Node kanan (data no. 3,4,7,8) $\rightarrow \Sigma gi = -2.240.000$

$$Score_{right} = \frac{(-2.240.000)^2}{(4 + 1)} = 1.003.520.000$$

c) Split pada Musim Libur

- Node kiri (data no. 1,3,5,7,9) $\rightarrow \Sigma gi = +500.000$

$$Score_{left} = \frac{(500.000)^2}{(5 + 1)} = 41.666.667$$

- Node kanan (data no. 2,4,6,8,10) $\rightarrow \Sigma gi = -500.000$

$$Score_{right} = \frac{(-500.000)^2}{(5 + 1)} = 41.666.667$$

4) Menentukan Split Optimal (Gain Tertinggi)

Nilai Gain dihitung untuk setiap kandidat split dengan rumus:

$$Gain = Score_{left} + Score_{right} - Score_{parent}$$

a) Gain split Jarak < 150:

$$Gain = 483.657.143 + 677.120.000 - 0 = 1.160.777.143$$

b) Gain split Jenis Bus:

$$Gain = 716.800.000 + 1.003.520.000 - 0 = 1.720.320.000$$

c) Gain split Musim Libur:

$$Gain = 41.666.667 + 41.666.667 - 0 = 83.333.334$$

Kesimpulannya, split pada atribut jenis_bus dipilih karena menghasilkan Gain tertinggi (1.720.320.000), sehingga pohon keputusan pertama membagi data berdasarkan atribut tersebut.

5) Membangun Pohon Keputusan

Nilai output setiap daun (leaf weight) dihitung dengan rumus:

$$w = \frac{-\Sigma gi}{(\Sigma hi + \lambda)}$$

a) Daun Kiri (jenis_bus = 0, data no. 1, 2, 5, 6, 9, 10):

$$\begin{aligned} \Sigma gi &= 140.000 - 60.000 + 840.000 + 640.000 \\ &+ 440.000 + 240.000 = +2.240.000 \end{aligned}$$

$$w_{kiri} = \frac{-2.240.000}{(6 + 1)} = -320.000$$

b) Daun Kanan (jenis_bus = 1, data no. 3, 4, 7, 8):

$$\begin{aligned} \Sigma gi &= -860.000 - 1.060.000 - 60.000 - 260.000 \\ &= -2.240.000 \end{aligned}$$

$$w_{kanan} = \frac{-(-2.240.000)}{(4 + 1)} = +448.000$$

6) Pembaruan Prediksi

Prediksi diperbarui pada setiap iterasi dengan rumus:

$$\hat{y}(t) = \hat{y}(t-1) + \eta \times ft(xi)$$

di mana η = learning rate = 0,3

a) Daun Kiri (jenis_bus = 0):

$$\hat{y}(1) = 2.440.000 + 0,3 \times (-320.000) = Rp 2.344.000$$

b) Daun Kanan (jenis_bus = 1):

$$\hat{y}(1) = 2.440.000 + 0,3 \times (448.000) = Rp 2.574.400$$

Hasil pembaruan prediksi seluruh 10 data ditampilkan pada Tabel IV.

IV. TABEL IV. CONTOH HASIL PREDIKSI ITERASI PERTAMA

No	Jenis Bus	$\hat{y}(0)$ (Rp)	Leaf	$\hat{y}(1)$ (Rp)
1-2,5-6,9-10	0 (Medium)	2.440.000	-320.000	2.344.000
3-4,7-8	1 (Big)	2.440.000	+448.000	2.574.400

Berdasarkan parameter masukan yang digunakan, yaitu jarak = 150, jenis bus = 0, dan musim libur = 0, hasil prediksi diarahkan ke daun kiri (left leaf) dengan kondisi jenis bus = 0. Pada daun tersebut, nilai prediksi harga mengalami pembaruan dari nilai

awal sebesar Rp 2.440.000 yang diperoleh dari rata-rata seluruh target, menjadi Rp 2.344.000.

7) Pengulangan proses

Tahap akhir dalam proses ini dilakukan dengan mengulangi seluruh tahapan mulai dari tahap kedua, yaitu menghitung kembali nilai gradient dan hessian berdasarkan sisa kesalahan dari hasil prediksi sebelumnya. Nilai-nilai tersebut kemudian digunakan untuk membangun pohon keputusan selanjutnya. Pada setiap iterasi, dilakukan perhitungan nilai gain untuk menentukan titik pemisahan (split) yang paling optimal. Proses ini berlangsung secara berulang hingga jumlah pohon yang telah ditentukan tercapai, sesuai dengan prinsip boosting di mana setiap pohon baru berperan dalam memperbaiki kesalahan dari model sebelumnya. Dengan mekanisme iteratif ini, XGBoost secara bertahap memperkecil residual error sehingga menghasilkan model prediksi yang semakin akurat pada setiap iterasi berikutnya.

D. Parameter Model

Dataset yang digunakan merupakan data historis harga penyewaan bus pariwisata dari PT. Setio Budi Luhur Trans, berjumlah 196 baris data. Data mentah mencakup lokasi penjemputan, destinasi tujuan, jenis bus, kondisi musim, dan harga sewa. Contoh data awal ditunjukkan pada Tabel I

IV. HASIL DAN PEMBAHASAN

A. Evaluasi Model XGBoost

Evaluasi model XGBoost dilakukan terpisah untuk setiap kombinasi kategori bus dan kondisi musim menggunakan data uji (testing set) sebesar 15% dari total 196 data, dengan pembagian menggunakan $\text{random_state}=21$. Metrik yang digunakan meliputi Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), Root Mean Square Error (RMSE), dan koefisien determinasi (R^2). Hasil evaluasi pada data uji disajikan pada Tabel V.

TABEL VI. HASIL EVALUASI MODEL XGBOOST PADA DATA UJI

Kategori	N Train	N Test	MAE Test (Rp)	MAPE Test	RMSE Test (Rp)	R^2	Akurasi
Medium - Normal	41	8	87.500	2,81%	145.774	0,9655	97,19%
Medium - Libur	41	8	120.212	3,47%	208.620	0,9294	96,53%
Big - Normal	41	8	126.463	2,76%	232.338	0,9369	97,24%
Big - Libur	41	8	126.463	2,64%	232.338	0,9369	97,36%
Rata-rata	-	-	115.160	2,92%	-	0,9422	97,08%

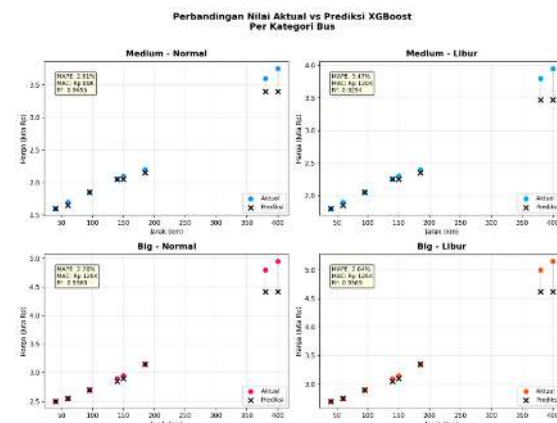
Berdasarkan Tabel VI, seluruh model XGBoost per kategori menunjukkan performa yang sangat baik pada data uji. Nilai MAPE berkisar antara 2,64% hingga 3,47%, jauh di bawah ambang batas 10% yang umum dijadikan standar akurasi prediksi yang baik. Rata-rata akurasi keseluruhan mencapai

97,08% dengan rata-rata MAPE sebesar 2,92% dan rata-rata R^2 sebesar 0,9422.

Nilai R^2 rata-rata sebesar 0,9422 menunjukkan bahwa model mampu menjelaskan sekitar 94,22% variansi harga sewa bus berdasarkan tiga variabel input. Kategori Medium - Normal menghasilkan R^2 tertinggi (0,9655), sementara Medium - Libur memiliki R^2 terendah (0,9294), yang dapat disebabkan oleh variabilitas harga yang lebih tinggi pada periode liburan. Nilai RMSE pada semua kategori masih proporsional terhadap skala harga sewa bus yang berkisar antara Rp 1.600.000 hingga Rp 7.200.000.

B. Perbandingan Nilai Aktual dan Prediksi

Gambar 2 menampilkan perbandingan nilai harga aktual dengan hasil prediksi XGBoost untuk masing-masing kategori pada data uji. Titik berwarna menunjukkan nilai aktual, sedangkan tanda silang (x) menunjukkan nilai prediksi.



Gbr. 2 Perbandingan Nilai Aktual dan Prediksi XGBoost

Dari Gambar 2 terlihat bahwa prediksi model cenderung mengikuti pola aktual dengan baik pada semua kategori. Sebagian besar titik prediksi berimpit atau sangat dekat dengan nilai aktual, dengan error terbesar terjadi pada rentang jarak 350–400 km. Hal ini disebabkan oleh sedikitnya jumlah data training pada rentang jarak tersebut, sehingga model kurang optimal dalam mempelajari pola di area itu. Meskipun demikian, nilai MAPE data uji pada semua kategori tetap berada di bawah 3,5%, yang menunjukkan bahwa model XGBoost masih mampu memberikan estimasi harga yang akurat secara keseluruhan.

V. KESIMPULAN

Penelitian ini berhasil menerapkan algoritma Extreme Gradient Boosting (XGBoost) untuk memprediksi harga sewa bus pariwisata pada PT. Setio Budi Luhur Trans. Kesimpulan dari penelitian ini adalah: (1) XGBoost mampu membangun model prediksi harga sewa bus berdasarkan tiga variabel utama, yaitu jarak perjalanan, jenis bus, dan kondisi musim/libur, dengan akurasi rata-rata sebesar 97,08% dan MAPE rata-rata 2,92%

pada data uji; (2) Nilai R^2 rata-rata sebesar 0,9422 menunjukkan model mampu menjelaskan 94,22% variansi harga; (3) Dari keempat kategori, model Medium - Normal menghasilkan performa terbaik dengan MAPE 2,81% dan R^2 0,9655; (4) Algoritma XGBoost terbukti cocok diterapkan pada dataset berukuran kecil dengan fitur yang relevan untuk kasus prediksi harga jasa transportasi pariwisata.

Untuk pengembangan selanjutnya, disarankan untuk memperluas jumlah data dan variabel prediksi seperti kapasitas bus dan durasi perjalanan, serta mengeksplorasi teknik hyperparameter tuning yang lebih komprehensif untuk mengurangi gap antara performa training dan testing.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada PT. Setio Budi Luhur Trans atas dukungan data dalam penelitian ini, serta kepada Dosen Pembimbing yang telah memberikan arahan selama proses penelitian berlangsung.

REFERENSI

- [1] A. Rahman, A. Yusnita, and H. Ekawati, "Penerapan Algoritma XGBoost Dalam Prediksi Harga Sewa Kos Di Kota Samarinda," *Bulletin of Information Technology (BIT)*, vol. 6, no. 4, pp. 379-390, Dec. 2025, doi: 10.47065/bit.v7i1.2304.
- [2] Alatas, A. T., Mumpuni, R., & Nurlaili, A. L. (2021). Spk penilaian kinerja untuk kenaikan jabatan pegawai menggunakan metode moora. *JIFOSI*, 2(2), 171-180, 2024.
- [3] J. E. T. Tampubolon et al., "Pengembangan Sistem Kasir Berbasis Web dengan Model Agile pada CV. Permata Garden," *AKSIOMA J. Sains Ekon. Dan Edukasi*, vol. 2, no. 8, pp. 1732-1745, 2025.
- [4] D. Aqsha, "Perbandingan Kinerja Algoritma Extreme Gradient Boosting dan Random Forest untuk Prediksi Harga Rumah di Jabodetabek," *Jurnal Ilmu Komputer dan Sistem Informasi*, vol. 13, no. 1, pp. 1-7, 2025, doi: 10.24912/jiksi.v13i1.32863.
- [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.
- [6] D. J. Y. Yantono, N. R. Puspaningrum, S. A. Zein, A. H. A. W. R. Marani, and G. J. Namat, "Analisis Komparasi Algoritma Machine Learning untuk Prediksi Harga Sewa Properti Komersial," *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, vol. 6, no. 1, pp. 149-155, Jan. 2026, doi: 10.57152/malcom.v6i1.2412.
- [7] H. Ardiana et al., "Perancangan Sistem Aplikasi Penyewaan Bus Pariwisata Berbasis Java pada PT Djess Berkah Wisata," *J. JRAMI*, vol. 2, no. 01, Jan. 2021.
- [8] D. T. Al Fatah et al., "Perancangan Sistem Penyewaan Bus Berbasis Laravel Framework pada PT Sumber Jaya Trans Tangerang," 2021.
- [9] T. F. Miraza, "Pengembangan Sistem Informasi Manajemen Transportasi Berbasis Web di PT Surveyor Indonesia Menggunakan Framework Laravel," *UIN Maulana Malik Ibrahim*, 2023.
- [10] J. Friedman, "Greedy Function Approximation: A Gradient Boosting Machine," *Ann. Stat.*, vol. 29, Nov. 2000.
- [11] D. J. Y. Yantono et al., "Perbandingan Model Ensemble Learning dalam Memprediksi Harga Sewa Indekos di Jakarta," *Jurnal Informatika dan Teknik Elektro Terapan (JITET)*, Oct. 2025, doi: 10.23960/jitet.v13i3.7737.