

Analisis Regresi *Linear* dan *Ensemble Learning* Berbasis Kontribusi Fitur dalam Prediksi Harga Mobil Listrik

¹Nur Oktavin Idris, ²Fuad Pontooyo

^{1,2}Universitas Negeri Gorontalo, Indonesia

¹nur.oktavin@ung.ac.id; ²fuad.pontooyo@ung.ac.id

Article Info

Article history:

Received, 2026-01-12

Revised, 2026-01-28

Accepted, 2026-01-30

Kata Kunci:

mobil listrik
prediksi harga
machine learning
ensemble learning
regresi linear

Keywords:

electric vehicles
price prediction
machine learning
ensemble learning
linear regression

ABSTRAK

Penelitian ini bertujuan untuk menganalisis kinerja regresi linear dan metode *ensemble learning* dalam memprediksi harga mobil listrik berdasarkan spesifikasi teknis kendaraan, serta mengkaji kontribusi fitur utama terhadap hasil prediksi. Permasalahan utama dalam prediksi harga mobil listrik adalah tingginya variasi harga yang dipengaruhi oleh hubungan *nonlinier* antar atribut teknis, sehingga sulit dimodelkan menggunakan pendekatan linier sederhana. Penelitian ini menggunakan regresi linear sebagai model dasar, serta *Random Forest* dan *Gradient Boosting* sebagai model ensemble pembanding. Data penelitian diperoleh dari dataset spesifikasi mobil listrik pada platform *Kaggle* yang diproses melalui tahap pembersihan data, pengodean variabel kategorikal, normalisasi, serta pembagian data latih dan data uji dengan rasio 80:20. Evaluasi kinerja model dilakukan menggunakan *mean squared error* (MSE) dan *koefisien determinasi* (R^2). Hasil pengujian menunjukkan bahwa model *Gradient Boosting* memberikan performa terbaik dengan nilai MSE sebesar 8.63 dan R^2 sebesar 0.891, lebih baik dibandingkan *Random Forest* dan regresi linear yang menghasilkan nilai MSE lebih tinggi dan R^2 lebih rendah. Analisis kontribusi fitur menunjukkan bahwa waktu akselerasi kendaraan merupakan faktor yang paling dominan dalam memengaruhi harga mobil listrik. Temuan ini menegaskan bahwa pendekatan *ensemble learning* tidak hanya meningkatkan akurasi prediksi, tetapi juga memberikan pemahaman analitis terhadap faktor teknis utama dalam pembentukan harga mobil listrik.

ABSTRACT

This study aims to analyze the performance of linear regression and ensemble learning methods in predicting electric vehicle prices based on technical specifications, as well as to examine the contribution of key features to the prediction results. The main challenge in electric vehicle price prediction lies in the high price variability driven by nonlinear relationships among technical attributes, which are difficult to capture using simple linear models. Linear regression was employed as a baseline model, while *Random Forest* and *Gradient Boosting* were used as ensemble learning approaches. The dataset was obtained from *Kaggle* and processed through data cleaning, categorical encoding, normalization, and an 80:20 train-test split. Model performance was evaluated using mean squared error (MSE) and the coefficient of determination (R^2). The results indicate that the *Gradient Boosting* model achieved the best performance, with an MSE of 8.63 and an R^2 of 0.891, outperforming both *Random Forest* and linear regression models. Feature contribution analysis reveals that vehicle acceleration time is the most influential factor in determining electric vehicle prices. These findings demonstrate that ensemble learning not only improves predictive accuracy but also provides analytical insights into the key technical factors shaping electric vehicle pricing.

This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-nc-nd/4.0/) license.



Penulis Korespondensi:

Nur Oktavin Idris,
Program Studi Teknologi Rekayasa Perangkat Lunak,
Universitas Negeri Gorontalo,
Email: nur.oktavin@ung.ac.id

1. PENDAHULUAN

Sektor transportasi merupakan salah satu penyumbang utama emisi gas rumah kaca dan konsumsi energi global. Dalam beberapa dekade terakhir, kendaraan listrik (*electric vehicles*) khususnya mobil listrik, dikembangkan sebagai alternatif transportasi yang lebih ramah lingkungan karena memiliki efisiensi energi yang lebih tinggi dan emisi operasional yang lebih rendah dibandingkan kendaraan berbahan bakar fosil. Studi *life cycle assessment* menunjukkan bahwa mobil listrik memiliki potensi pengurangan emisi yang signifikan dibandingkan kendaraan konvensional, terutama ketika didukung oleh sistem energi yang lebih bersih [1]. Sejalan dengan hal tersebut, laporan International Energy Agency menunjukkan bahwa pertumbuhan mobil listrik secara global meningkat pesat seiring dengan kebijakan dekarbonisasi dan transisi energi berkelanjutan di berbagai negara [2].

Meskipun demikian, adopsi mobil listrik masih menghadapi berbagai tantangan, terutama dari sisi ekonomi. Harga mobil listrik, khususnya di Indonesia relatif lebih tinggi dan bervariasi dibandingkan kendaraan konvensional, sehingga menjadi salah satu faktor penghambat utama bagi konsumen dalam pengambilan keputusan [3, 4]. Variasi harga tersebut dipengaruhi oleh berbagai spesifikasi teknis, seperti kapasitas baterai, jarak tempuh maksimum, konsumsi energi, dan daya motor listrik, yang saling berinteraksi secara kompleks dan sulit dimodelkan menggunakan pendekatan linier sederhana.

Sejumlah penelitian sebelumnya telah mencoba memodelkan harga kendaraan, termasuk mobil listrik dan mobil konvensional, menggunakan pendekatan statistik dan *machine learning*. Metode regresi linear banyak digunakan karena kesederhanaannya, namun memiliki keterbatasan dalam menangkap hubungan nonlinier antar variabel harga dan spesifikasi teknis kendaraan [5]. Untuk mengatasi keterbatasan tersebut, pendekatan *ensemble learning* seperti *Random Forest* dan *Gradient Boosting* diperkenalkan dan terbukti mampu meningkatkan akurasi prediksi pada berbagai permasalahan regresi berbasis data kendaraan [6, 7]

Penelitian terkini menunjukkan bahwa pendekatan *machine learning* mampu memberikan performa yang lebih baik dalam prediksi harga mobil listrik dibandingkan metode konvensional, khususnya ketika menggunakan algoritma *ensemble* [8, 9]. Namun, sebagian besar studi masih berfokus pada satu algoritma utama dan belum mengkaji kontribusi fitur secara sistematis.

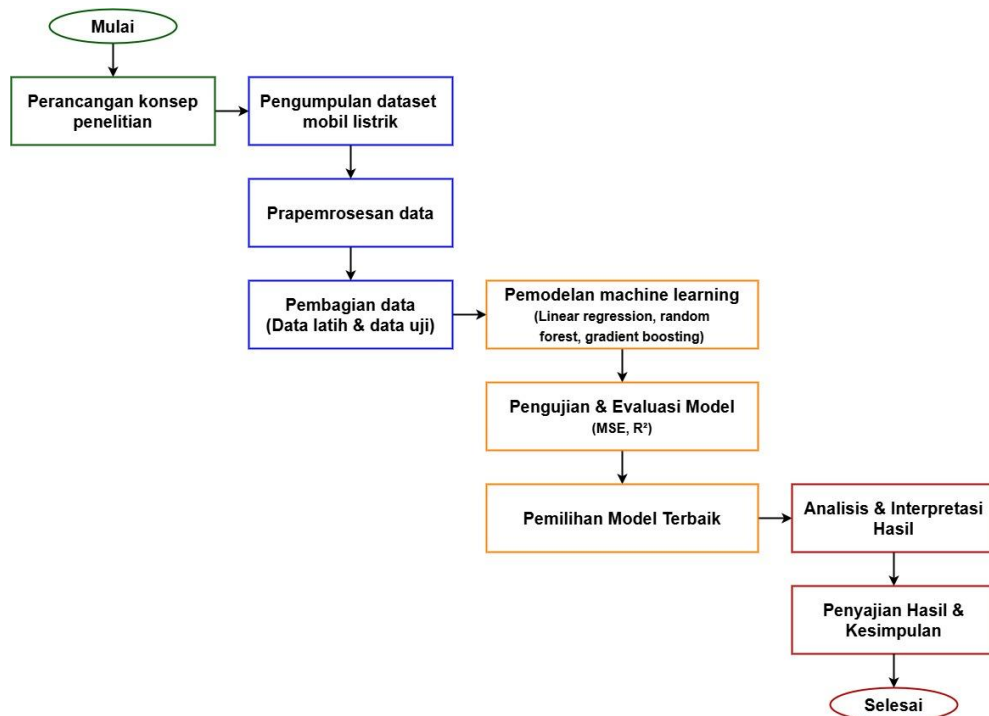
Berdasarkan permasalahan tersebut, pendekatan *machine learning* diusulkan sebagai solusi untuk memodelkan hubungan kompleks antar spesifikasi teknis mobil listrik dan menghasilkan prediksi harga. Kerangka kerja *machine learning* modern yang didukung pustaka terbuka seperti *scikit-learn* memungkinkan proses pemodelan, evaluasi, dan replikasi dilakukan secara sistematis dan terstandar [10]. Inovasi penelitian ini terletak pada penerapan dan perbandingan beberapa algoritma regresi *machine learning* menggunakan dataset spesifikasi mobil listrik global. Selain itu, penelitian ini juga menganalisis kontribusi fitur utama terhadap harga mobil listrik.

Berdasarkan latar belakang tersebut, permasalahan penelitian ini adalah bagaimana membangun model prediksi harga mobil listrik yang akurat menggunakan *machine learning* serta menentukan algoritma yang memiliki performa terbaik. Tujuan penelitian ini adalah mengembangkan dan mengevaluasi model prediksi harga mobil listrik berbasis *machine learning*. Manfaat penelitian ini diharapkan dapat memberikan kontribusi akademik dalam bidang informatika terapan, khususnya pada pemodelan prediktif berbasis data kendaraan, serta mendukung pengambilan keputusan bagi konsumen, produsen, dan pemangku kebijakan dalam pengembangan ekosistem mobil listrik.

Berdasarkan telaah terhadap penelitian-penelitian sebelumnya, dapat diidentifikasi bahwa masih terdapat celah penelitian dalam pemodelan harga mobil listrik. Sebagian besar studi lebih menekankan pada peningkatan akurasi prediksi melalui penerapan algoritma tertentu, tanpa melakukan analisis perbandingan kinerja model secara komprehensif serta tanpa mengkaji kontribusi fitur teknis secara eksplisit. Oleh karena itu, penelitian ini menawarkan kontribusi berupa analisis kinerja regresi linear dan *ensemble learning* yang dilengkapi dengan analisis kontribusi fitur utama dalam prediksi harga mobil listrik. Kontribusi penelitian ini meliputi: (1) penyajian perbandingan performa model regresi linear dan metode *ensemble learning* berbasis metrik kuantitatif, (2) analisis kontribusi fitur teknis utama terhadap pembentukan harga mobil listrik, serta (3) penyediaan dasar analitis bagi pengembangan sistem pendukung keputusan berbasis data dalam konteks kendaraan listrik.

2. METODE PENELITIAN

Metode penelitian ini disusun secara kronologis untuk membangun model prediksi harga mobil listrik berbasis *machine learning*. Desain penelitian mengikuti pendekatan pemodelan prediktif yang umum digunakan dalam studi prediksi harga kendaraan terutama mobil. Tahapan penelitian meliputi perancangan konsep, pengumpulan dan prapemrosesan data, pemodelan algoritma, pengujian kinerja model, serta analisis hasil. Alur penelitian secara keseluruhan disajikan pada Gambar 1.



Gambar 1. Diagram alir penelitian

Tahap awal penelitian adalah perancangan konsep dengan menentukan variabel input dan output berdasarkan telaah pustaka. Variabel input berupa spesifikasi teknis mobil listrik, seperti kapasitas baterai, jarak tempuh maksimum, konsumsi energi, performa kendaraan, serta karakteristik fungsional lainnya, sedangkan variabel output yang digunakan adalah harga mobil listrik. Pemilihan variabel ini didasarkan pada penelitian sebelumnya yang menunjukkan bahwa atribut teknis kendaraan terutama mobil berpengaruh signifikan terhadap harga [9, 11].

Data penelitian ini menggunakan dataset dari platform Kaggle, yaitu *EVs – One Electric Vehicle Dataset*, yang dikembangkan oleh Geoff Neil. Dataset ini berisi informasi spesifikasi teknis mobil listrik, seperti ditunjukkan pada Tabel 1. Secara keseluruhan, dataset terdiri dari 103 baris data dengan 14 kolom atau fitur, yang mencakup data numerik dan kategorikal. Seluruh fitur digunakan sebagai variabel input, sedangkan variabel *PriceEuro* digunakan sebagai variabel target/output dalam pemodelan prediksi harga mobil listrik.

Tabel 1. Deskripsi dataset

Nama Kolom	Deskripsi	Tipe Data
<i>Brand</i>	Merek mobil listrik	<i>Object</i>
<i>Model</i>	Model mobil listrik	<i>Object</i>
<i>AccelSec</i>	Waktu akselerasi (0–100 km/jam)	<i>Float</i>
<i>TopSpeed_KmH</i>	Kecepatan maksimum kendaraan	<i>Integer</i>
<i>Range_Km</i>	Jarak tempuh maksimum	<i>Integer</i>
<i>Efficiency_WhKm</i>	Konsumsi energi per km	<i>Integer</i>
<i>FastCharge_KmH</i>	Kecepatan pengisian daya cepat	<i>Integer</i>
<i>RapidCharge</i>	Dukungan pengisian cepat	<i>Object</i>
<i>PowerTrain</i>	Konfigurasi penggerak (FWD/RWD/AWD)	<i>Object</i>
<i>PlugType</i>	Jenis konektor pengisian daya	<i>Object</i>
<i>BodyStyle</i>	Bentuk bodi kendaraan	<i>Object</i>
<i>Segment</i>	Segmen kendaraan	<i>Object</i>
<i>Seats</i>	Jumlah kursi	<i>Integer</i>
<i>PriceEuro</i>	Harga mobil listrik (Euro)	<i>Integer</i>

Dataset yang digunakan dalam penelitian ini terdiri dari 103 data mobil listrik dengan 14 fitur input. Ukuran dataset tersebut dinilai memadai untuk tujuan penelitian komparatif antar model regresi, khususnya dalam mengevaluasi perbedaan kinerja antara regresi linear dan metode *ensemble learning*. Fokus utama penelitian ini bukan pada pengembangan model berskala besar, melainkan pada analisis perbandingan performa model dan kontribusi fitur. Selain itu, penggunaan dataset dengan ukuran terbatas memungkinkan interpretasi hasil model dan analisis kontribusi fitur dilakukan secara lebih terkontrol dan transparan.

Dataset tersebut kemudian diproses melalui tahap prapemrosesan data untuk meningkatkan kualitas dan konsistensi data. Tahapan prapemrosesan meliputi penghapusan data duplikat, penanganan nilai hilang, encoding variabel kategorikal, serta normalisasi fitur numerik. Proses ini bertujuan untuk memastikan data sesuai dengan kebutuhan algoritma *machine learning* serta meningkatkan stabilitas dan akurasi model prediksi [9].

Perlu dicatat bahwa dataset yang digunakan bersumber dari platform Kaggle dan merepresentasikan spesifikasi teknis mobil listrik yang tersedia secara publik. Dataset ini memiliki keterbatasan, antara lain jumlah data yang relatif terbatas serta kemungkinan bias representasi terhadap merek, segmen kendaraan, dan wilayah pasar tertentu. Oleh karena itu, hasil penelitian ini lebih menekankan pada evaluasi metodologis dan analisis pola hubungan antar fitur, bukan pada generalisasi mutlak terhadap seluruh pasar mobil listrik global.

Setelah prapemrosesan, data dibagi menjadi data latih dan data uji dengan rasio 80:20 yang umum digunakan dalam penelitian prediksi harga kendaraan dan mobil untuk menyeimbangkan kebutuhan pelatihan model dan evaluasi model. Tahap pemodelan dilakukan dengan menerapkan beberapa algoritma regresi *machine learning*, yaitu Linear Regression sebagai model dasar, serta Random Forest regressor dan Gradient Boosting regressor sebagai metode *ensemble learning*. Linear Regression digunakan sebagai pembanding awal untuk mengevaluasi hubungan linier antara spesifikasi teknis dan harga mobil. Pemilihan Random Forest didasarkan pada kemampuannya dalam menangani hubungan nonlinier dan interaksi antar fitur [7], sedangkan Gradient Boosting dipilih karena efektivitasnya dalam meminimalkan kesalahan prediksi secara bertahap melalui mekanisme *boosting* [8]. Pada penelitian ini, setiap algoritma diterapkan menggunakan parameter default sebagaimana direkomendasikan dalam pustaka *scikit-learn*. Pendekatan ini dipilih untuk menjaga konsistensi dan objektivitas dalam membandingkan kinerja antar model, sehingga perbedaan performa yang dihasilkan mencerminkan karakteristik algoritma, bukan hasil optimasi parameter tertentu. Pendekatan serupa juga banyak digunakan dalam penelitian komparatif berbasis *machine learning*.

Pengujian dan evaluasi model dilakukan menggunakan data uji untuk menilai kemampuan generalisasi masing-masing algoritma. Kinerja model dievaluasi menggunakan metrik Mean Squared Error (MSE), dan koefisien determinasi (R^2), yang secara luas digunakan dalam penelitian prediksi harga kendaraan terutama mobil untuk membandingkan performa model regresi secara objektif [12, 13]. Mean Squared Error dirumuskan sebagaimana ditunjukkan pada Persamaan (1) [14].

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (1)$$

Sedangkan koefisien determinasi dirumuskan sebagaimana ditunjukkan pada Persamaan (2) [15].

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (2)$$

di mana y_i merupakan nilai aktual, $\hat{y}_i$ adalah nilai prediksi, $\bar{y}$ adalah nilai rata-rata data aktual, dan n adalah jumlah data uji. Model dengan nilai kesalahan prediksi terendah dan nilai koefisien determinasi tertinggi ditetapkan sebagai model terbaik.

Tahap akhir penelitian meliputi analisis dan interpretasi hasil prediksi harga mobil listrik serta penyajian hasil penelitian dalam bentuk tabel dan grafik. Hasil analisis ini digunakan sebagai dasar dalam penarikan kesimpulan dan pembahasan implikasi model terhadap pemodelan harga mobil listrik.

3. HASIL DAN ANALISIS

Hasil evaluasi kinerja model prediksi harga mobil listrik disajikan pada Tabel 2. Evaluasi dilakukan terhadap tiga algoritma regresi *machine learning*, yaitu Linear Regression, Random Forest regressor, dan Gradient Boosting regressor, menggunakan metrik Mean Squared Error (MSE) dan koefisien determinasi (R^2). Berdasarkan hasil pengujian, model Linear Regression menghasilkan nilai MSE tertinggi dan nilai R^2 terendah dibandingkan model lainnya. Temuan ini menunjukkan bahwa hubungan antara spesifikasi teknis mobil listrik dan harga kendaraan tidak sepenuhnya bersifat linier, sehingga pendekatan regresi linier memiliki keterbatasan dalam menangkap kompleksitas hubungan antar variabel. Hasil ini sejalan dengan penelitian sebelumnya yang menyatakan bahwa model linier kurang efektif ketika data memiliki karakteristik nonlinier serta interaksi fitur yang kompleks [5].

Tabel 2. Hasil evaluasi kinerja model

Model	MSE	R ² Score
Random Forest	149124195.393752	0.811958
Gradient Boosting	86308893.468189	0.891167
Linear Regression	307617360.296505	0.612102

Berdasarkan hasil evaluasi pada Tabel 2, model *Gradient Boosting* menghasilkan nilai *Mean Squared Error* (MSE) sebesar 8.63, lebih rendah dibandingkan *Random Forest* (14.91) dan *Linear Regression* (30.76). Secara kuantitatif, *Gradient Boosting* mampu menurunkan nilai MSE sebesar sekitar 42% dibandingkan *Random Forest* dan sekitar 72% dibandingkan *Linear Regression*. Selain itu, nilai *koefisien determinasi* (R^2) yang dicapai oleh *Gradient Boosting* sebesar 0.891 menunjukkan peningkatan sekitar 9.7% dibandingkan *Random Forest* (0.812) dan sekitar 45.5% dibandingkan *Linear Regression* (0.612). Hasil ini mengindikasikan bahwa pendekatan *ensemble learning*, khususnya *Gradient Boosting*, memberikan peningkatan performa prediksi yang signifikan dibandingkan regresi linier.

Temuan tersebut sejalan dengan penelitian sebelumnya yang menunjukkan bahwa metode *ensemble learning* lebih efektif dalam memodelkan hubungan nonlinier pada data kendaraan berbasis spesifikasi teknis. Breiman [6] menjelaskan bahwa *Random Forest* unggul dalam menangkap interaksi antar fitur melalui mekanisme pembentukan pohon keputusan secara acak, yang menjelaskan peningkatan performa *Random Forest* dibandingkan *Linear Regression* pada penelitian ini. Sementara itu, *Gradient Boosting* dilaporkan memberikan performa yang lebih tinggi pada permasalahan regresi dengan kompleksitas fitur yang tinggi karena mekanisme pembelajaran bertahap yang meminimalkan kesalahan residual [7]. Konsistensi hasil penelitian ini dengan studi-studi terdahulu memperkuat validitas pendekatan *ensemble learning* dalam konteks prediksi harga mobil listrik. Secara keseluruhan, hasil perbandingan kinerja antar model menunjukkan bahwa pendekatan *ensemble learning* mampu mengatasi keterbatasan regresi linier dalam memodelkan hubungan nonlinier pada data spesifikasi mobil listrik. Di antara model yang diuji, *Gradient Boosting* memberikan performa terbaik karena mekanisme pembelajaran bertahap yang efektif dalam meminimalkan kesalahan prediksi.

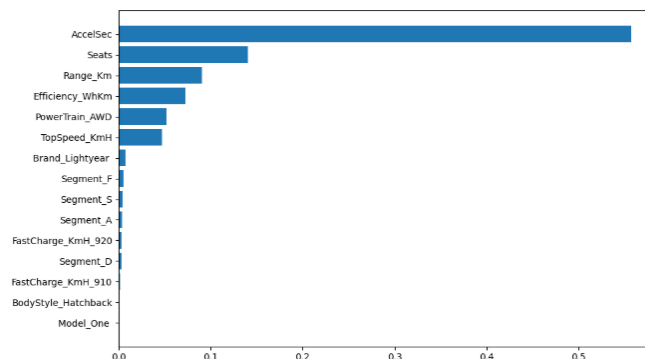
Perbandingan antara nilai harga aktual dan nilai hasil prediksi menggunakan model terbaik yaitu *Gradient Boosting regressor*, disajikan pada Tabel 3.

Tabel 3. Perbandingan harga aktual dan prediksi harga

Brand	Model	PriceEuro	Predicted Price
Audi	e-tron 50 quattro	67358	69522.2
Tesla	Cybertruck Dual Motor	55000	51249.2
Nissan	Ariya 87kWh	50000	47004.6
Porsche	Taycan 4S	102945	133576.3
Audi	e-tron Sportback 55 quattro	81639	78383.4

Nilai prediksi yang mendekati nilai aktual pada Tabel 3 mengindikasikan bahwa model memiliki tingkat akurasi yang baik serta mampu melakukan generalisasi terhadap data uji yang belum pernah dilihat sebelumnya. Hal ini menunjukkan bahwa model tidak hanya fit terhadap data latih, tetapi juga memiliki kemampuan prediktif yang memadai pada data baru.

Selanjutnya, analisis *feature importance* dilakukan untuk mengidentifikasi fitur yang paling berpengaruh terhadap prediksi harga mobil listrik, sebagaimana ditunjukkan pada Gambar 2.



Gambar 2. Feature importance

Hasil analisis *feature importance* pada Gambar 2 menunjukkan bahwa waktu akselerasi kendaraan (AccelSec) merupakan variabel dengan kontribusi terbesar terhadap prediksi harga mobil listrik, yang menegaskan dominasi aspek performa kendaraan dalam pembentukan harga. Dominasi fitur akselerasi mencerminkan posisi performa sebagai indikator utama dalam segmentasi pasar mobil listrik, di mana kendaraan dengan akselerasi

lebih cepat umumnya diposisikan sebagai produk dengan nilai jual yang lebih tinggi. Meskipun variabel jarak tempuh (*Range_Km*) dan efisiensi energy (*Efficiency_WhKm*) tetap memberikan kontribusi signifikan, pengaruhnya relatif lebih kecil dibandingkan fitur performa dan kapasitas fungsional kendaraan. Selain itu, faktor performa dan kapasitas lain seperti jumlah kursi (*Seats*), konfigurasi penggerak (AWD), dan kecepatan maksimum (*TopSpeed_KmH*) juga menunjukkan pengaruh yang berarti terhadap harga mobil listrik. Temuan ini mengindikasikan bahwa harga mobil listrik dibentuk oleh kombinasi antara performa, kapasitas fungsional, dan efisiensi teknologi, sementara fitur kategorikal seperti merek dan segmen kendaraan memberikan kontribusi yang relatif lebih kecil dalam model prediksi. Secara keseluruhan, hasil penelitian ini menunjukkan bahwa pendekatan *machine learning* berbasis algoritma *ensemble*, khususnya *Gradient Boosting regressor*, memberikan performa yang lebih baik dibandingkan pendekatan regresi linier dalam memprediksi harga mobil listrik. Integrasi hasil evaluasi kuantitatif, analisis visual, analisis perbandingan prediksi, dan analisis *feature importance* menunjukkan bahwa model yang dikembangkan memiliki validitas empiris dan relevansi praktis untuk mendukung pengembangan sistem pendukung keputusan berbasis data dalam penentuan harga dan segmentasi pasar mobil listrik.

4. KESIMPULAN

Penelitian ini berhasil mencapai tujuan penelitian, yaitu membangun dan mengevaluasi model prediksi harga mobil listrik berbasis spesifikasi teknis kendaraan menggunakan pendekatan *machine learning*. Hasil penelitian menunjukkan bahwa metode *ensemble learning*, khususnya *Gradient Boosting regressor*, memberikan performa terbaik dengan nilai Mean Squared Error (MSE) sebesar 8.63 dan koefisien determinasi (R^2) sebesar 0.891, yang secara signifikan lebih baik dibandingkan *Random Forest* dan regresi linier. Temuan ini menegaskan bahwa hubungan antara spesifikasi teknis kendaraan dan harga mobil listrik bersifat nonlinier dan lebih efektif dimodelkan menggunakan algoritma *ensemble*. Analisis kontribusi fitur menunjukkan bahwa waktu akselerasi kendaraan (*AccelSec*) merupakan faktor paling dominan dalam memengaruhi harga mobil listrik, diikuti oleh kapasitas fungsional dan efisiensi teknologi, sementara fitur kategorikal seperti merek dan segmen memberikan kontribusi yang relatif lebih kecil. Kontribusi utama penelitian ini terletak pada perbandingan kinerja model regresi linier dan *ensemble learning* yang dilengkapi dengan analisis *feature importance*, sehingga memberikan pemahaman analitis terhadap faktor teknis pembentuk harga mobil listrik. Secara praktis, model yang dikembangkan berpotensi digunakan sebagai dasar sistem pendukung keputusan berbasis data untuk membantu konsumen dalam membandingkan harga kendaraan, produsen dalam menentukan strategi penetapan harga, serta pembuat kebijakan dalam analisis pasar mobil listrik. Penelitian selanjutnya disarankan untuk menggunakan dataset yang lebih besar dan beragam, serta menerapkan optimasi parameter dan *cross validation* guna meningkatkan generalisasi dan keandalan model.

REFERENSI

- [1] T. R. Hawkins, B. Singh, G. Majeau-bettez, and A. H. Strømman, "Comparative Environmental Life Cycle Assessment of Conventional and Electric Vehicles," *J. Ind. Ecol.*, vol. 17, no. 1, pp. 53–64, 2012, doi: 10.1111/j.1530-9290.2012.00532.x.
- [2] International Energy Agency, *Global EV Outlook 2023 Catching up with climate ambition*. Paris: IEA, 2023. [Online]. Available: <https://www.iea.org/reports/global-ev-outlook-2023>
- [3] S. Li, L. Tong, J. Xing, and Y. Zhou, "The Market for Electric Vehicles: Indirect Network Effects and Policy Design," *J. Assoc. Environ. Resour. Econ.*, vol. 4, no. 1, pp. 89–133, 2017, doi: 10.1086/689702.
- [4] V. Tulus and P. Sidabutar, "Kajian pengembangan kendaraan listrik di Indonesia : prospek dan hambatannya," vol. 15, no. 1, pp. 21–38, 2020.
- [5] K. Noor and S. Jan, "Vehicle Price Prediction System using Machine Learning Techniques," *Int. J. Comput. Appl.*, vol. 167, no. 9, pp. 27–31, 2017.
- [6] L. E. O. Breiman, "Random Forests," *Mach. Learn.*, vol. 45, pp. 5–32, 2001, doi: 10.1023/A:1010933404324.
- [7] J. H. Friedman, "Greedy function approximation: A gradient boosting machine," *Ann. Stat.*, vol. 29, no. 5, pp. 1189–1232, 2001, doi: 10.1214/aos/1013203451.
- [8] A. Rakul, V. Santhosh, A. Kabilish, J. M. Raj, and T. Arachelvi, "Electric Vehicle Price Prediction Using Machine Learning Techniques," *Int. J. Innov. Res. Technol.*, vol. 11, no. 12, pp. 5217–5222, 2025.
- [9] A. R. Chitra and B. C. Arjun, "Performance Analysis of Regression Algorithms for Used Car Price Prediction : KNIME Analytics Platform," *International J. Res. Appl. Sci. Eng. Technol.*, vol. 11, no. 11, pp. 1324–1331, 2023, doi: 10.22214/ijraset.2023.49180.

- [10] A. Geron, *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems*, 2nd ed. New York: O'Reilly Media, 2019.
- [11] S. Dubey, A. Renavikar, and S. Kanungo, "Machine Learning Model for Prediction of Electric Vehicle Prices," *J. Inf. Syst. Eng. Manag.*, vol. 10, no. 50s, pp. 2468–4376, 2025.
- [12] X. V. Nguyen, N. T. Le, and T. D. Tran, "Forecasting Resale Value Of Cars : A Comparative Analysis Of Linear Regression , Random Forest , And Feedforward Neural Networks," in *Proceedings of the International Conferences (ISC) 2024 "Green Values for Sustainable Development,"* 2024, pp. 704–713.
- [13] M. D. H. Kusuma and H. Syarif, "Penerapan Model Regresi Linier dalam Prediksi Harga Mobil Bekas di India dan Visualisasi dengan Menggunakan Power BI," *J. Indones. Manaj. Inform. dan Komun.*, vol. 5, no. 2, pp. 1097–1110, 2024, doi: 10.35870/jimik.v5i2.629.
- [14] B. Al Qohar, A. U. Dullah, P. Utami, and Jumanto, "Vehicle CO₂ Emission Predictive Analytics Using HistGradientBoosting Regression Algorithms," *J. Media Inf. Teknol.*, vol. 2, no. 1, pp. 1–12, 2025, doi: 10.69616/mit.v2i1.198.
- [15] M. R. Herlando, S. H. Wijoyo, and M. A. Akbar, "Analisis Perbandingan Machine Learning Untuk Memprediksi Harga Minyak dengan Regresi Linear dan Support Vector Regression," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 9, no. 6, pp. 1–10, 2025.