



Prediksi Konsumsi Bahan Bakar Menggunakan *Machine Learning* Berdasarkan Data ECU

Faiq Aryandi Sofyan^{1*}, Pratomo Budi Santosa²

^{1,2}Program Studi Teknik Elektro, Universitas Muhammadiyah Surakarta, Indonesia

*Penulis Korespondensi: d400220095@student.ums.ac.id

Abstract. Fuel consumption efficiency is a key indicator in the development of modern vehicles, as it directly affects operating costs and exhaust emissions. This study aimed to analyze the characteristics of Electronic Control Unit (ECU) data as predictive parameters for fuel consumption and to develop a regression-based Machine Learning model for predicting vehicle fuel consumption. The dataset consisted of 597 vehicle records, with Engine Size, Cylinders, CO₂ Emissions, and Fuel Type serving as predictor variables, while Combined Fuel Consumption was used as the target variable. Data were analyzed using multiple linear regression, and model performance was evaluated based on the coefficient of determination (R^2), Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and Standard Error of Estimate (SEE). The results demonstrated that the proposed model achieved excellent predictive performance, with an R value of 0.989, an R^2 of 0.977, an adjusted R^2 of 0.977, an SEE of 0.432 L/100 km, an MAE of 0.24 L/100 km, and an RMSE of 0.430 L/100 km. Among the predictor variables, CO₂ Emissions was identified as the most influential factor affecting fuel consumption. These findings indicate that ECU data can provide accurate fuel consumption predictions and have significant potential to support the development of vehicle fuel-efficiency monitoring systems and data-driven decision-making in the automotive industry.

Keywords: ECU; Fuel Consumption; Machine Learning; Prediction; Regression.

Abstrak. Efisiensi konsumsi bahan bakar merupakan salah satu indikator penting dalam pengembangan kendaraan modern karena berpengaruh terhadap biaya operasional dan emisi gas buang. Penelitian ini bertujuan menganalisis karakteristik data *Electronic Control Unit* (ECU) sebagai parameter prediksi konsumsi bahan bakar serta membangun model *Machine Learning* berbasis regresi untuk memprediksi konsumsi bahan bakar kendaraan. Dataset yang digunakan terdiri atas 597 data kendaraan dengan variabel *Engine Size*, *Cylinders*, *CO₂ Emissions*, dan *Fuel Type* sebagai prediktor, sedangkan konsumsi bahan bakar gabungan (*Combined Fuel Consumption*) sebagai variabel target. Analisis dilakukan menggunakan regresi linear berganda dengan evaluasi kinerja model berdasarkan koefisien determinasi (R^2), *Mean Absolute Error* (MAE), *Root Mean Square Error* (RMSE), dan *Standard Error of Estimate* (SEE). Hasil penelitian menunjukkan bahwa model memiliki performa yang sangat baik dengan nilai R sebesar 0,989, R^2 sebesar 0,977, *Adjusted R²* sebesar 0,977, SEE sebesar 0,432 L/100 km, MAE sebesar 0,24 L/100 km, dan RMSE sebesar 0,430 L/100 km. Variabel *CO₂ Emissions* menjadi prediktor yang paling dominan dalam memengaruhi konsumsi bahan bakar. Temuan ini menunjukkan bahwa data ECU mampu menghasilkan prediksi konsumsi bahan bakar yang akurat sehingga berpotensi mendukung pengembangan sistem pemantauan efisiensi kendaraan dan pengambilan keputusan berbasis data pada industri otomotif.

Kata Kunci: ECU; Konsumsi Bahan Bakar; *Machine Learning*; Prediksi; Regresi.

1. LATAR BELAKANG

Perkembangan teknologi otomotif telah mengalami transformasi yang sangat pesat seiring meningkatnya penerapan sistem elektronika dan teknologi digital pada kendaraan modern. Hampir seluruh sistem utama kendaraan saat ini dikendalikan oleh *Electronic Control Unit* (ECU) yang berfungsi sebagai pusat pengendali sistem injeksi bahan bakar, pengapian, kontrol emisi, transmisi, serta berbagai sistem pendukung lainnya (Keskin et al., 2025). Selain menjalankan fungsi pengendalian, ECU juga mampu merekam berbagai parameter operasional kendaraan secara *real-time*, seperti putaran mesin (*Engine Revolutions Per Minute/RPM*),

kecepatan kendaraan (*Vehicle Speed*), posisi throttle (*Throttle Position Sensor/TPS*), beban mesin (*Engine Load*), tekanan udara masuk (*Manifold Absolute Pressure/MAP*), laju aliran udara (*Mass Air Flow/MAF*), suhu mesin (*Engine Coolant Temperature/ECT*), suhu udara masuk (*Intake Air Temperature/IAT*), hingga laju konsumsi bahan bakar (*Fuel Rate*) (Canal et al., 2024).

Salah satu pemanfaatan data ECU yang semakin berkembang adalah untuk memprediksi konsumsi bahan bakar kendaraan. Konsumsi bahan bakar merupakan indikator penting dalam mengevaluasi efisiensi operasional kendaraan karena berkaitan langsung dengan biaya penggunaan kendaraan, konsumsi energi fosil, serta emisi gas rumah kaca. Meningkatnya jumlah kendaraan bermotor menyebabkan kebutuhan akan sistem prediksi konsumsi bahan bakar yang cepat, akurat, dan mampu bekerja secara *real-time* menjadi semakin penting (Topić et al., 2022). Informasi tersebut dapat dimanfaatkan oleh pengguna kendaraan, perusahaan transportasi, maupun industri otomotif dalam meningkatkan efisiensi energi, mengurangi biaya operasional, serta mendukung strategi pengurangan emisi karbon (Canal et al., 2024).

Beberapa penelitian menunjukkan bahwa algoritma berbasis *Machine Learning*, terutama metode *ensemble learning* seperti *Random Forest* dan *Gradient Boosting*, memiliki akurasi prediksi konsumsi bahan bakar yang lebih tinggi dibandingkan algoritma konvensional karena mampu menangkap pola hubungan yang kompleks dari berbagai parameter kendaraan (Valladolid & Ortiz, 2026). Meskipun demikian, sebagian besar penelitian terdahulu masih menggunakan dataset simulasi, dataset publik, atau data hasil pengujian laboratorium yang belum sepenuhnya mencerminkan kondisi operasional kendaraan di lapangan. Selain itu, penelitian yang memanfaatkan data aktual hasil pembacaan ECU melalui OBD-II pada kendaraan yang beroperasi di Indonesia masih sangat terbatas (Siswanto et al., 2025). Belum banyak penelitian yang melakukan perbandingan performa beberapa algoritma *Machine Learning* menggunakan data ECU aktual untuk menentukan model prediksi yang paling sesuai dengan karakteristik kendaraan, pola berkendara, dan kondisi jalan di Indonesia. Kondisi tersebut menunjukkan adanya kesenjangan penelitian (*research gap*) yang masih perlu dikaji lebih lanjut.

Berdasarkan kesenjangan tersebut, penelitian ini menawarkan pendekatan berbasis data aktual ECU kendaraan sebagai sumber data utama dalam membangun model prediksi konsumsi bahan bakar menggunakan beberapa algoritma *Machine Learning*. Kebaruan penelitian ini terletak pada pemanfaatan data operasional kendaraan yang diperoleh secara langsung melalui ECU pada kondisi penggunaan nyata serta evaluasi komparatif terhadap algoritma *Machine Learning* untuk mengidentifikasi model prediksi dengan tingkat akurasi terbaik. Penelitian ini

bertujuan untuk mengetahui Prediksi Konsumsi Bahan Bakar Menggunakan *Machine Learning* Berdasarkan Data ECU.

2. KAJIAN TEORITIS

Electronic Control Unit (ECU)

Electronic Control Unit (ECU) merupakan sistem komputer tertanam yang mengendalikan berbagai sistem elektronik kendaraan, seperti injeksi bahan bakar, pengapian, dan kontrol emisi. ECU juga menghasilkan data operasional kendaraan secara *real-time* melalui sistem *On-Board Diagnostics II* (OBD-II), meliputi *Engine RPM*, *Vehicle Speed*, *Engine Load*, *Throttle Position*, *Mass Air Flow (MAF)*, *Manifold Absolute Pressure (MAP)*, *Engine Coolant Temperature (ECT)*, *Intake Air Temperature (IAT)*, *Fuel Rate*, dan *Fuel Trim*. Data ini menjadi sumber informasi pengembangan model prediksi konsumsi bahan bakar berbasis *Machine Learning* (Canal et al., 2024).

Konsumsi Bahan Bakar Kendaraan

Konsumsi bahan bakar menunjukkan jumlah bahan bakar kendaraan dalam satuan jarak atau waktu, seperti L/100 km, km/L, atau L/jam. Nilainya dipengaruhi oleh berbagai faktor, antara lain putaran mesin, kecepatan, beban mesin, kondisi jalan, efisiensi pembakaran, dan gaya mengemudi. Metode konvensional berbasis perbandingan volume bahan bakar dan jarak tempuh belum mampu memberikan estimasi *real-time*, sehingga pendekatan berbasis data ECU dan *Machine Learning* menjadi alternatif yang lebih akurat (Jessica et al., 2025; Ibrahim & Zakzouk, 2026).

Machine Learning

Machine Learning (ML) merupakan salah satu cabang dari *Artificial Intelligence* (AI) yang memungkinkan komputer mempelajari pola, hubungan, dan karakteristik data secara otomatis tanpa harus diprogram secara eksplisit untuk setiap kondisi (Deepthi et al., 2026). Dalam bidang otomotif, *Machine Learning* dimanfaatkan untuk menganalisis data yang dihasilkan oleh berbagai sensor kendaraan guna memprediksi kondisi kendaraan, mendeteksi kerusakan mesin, menganalisis perilaku berkendara (*driving behavior*), serta memprediksi konsumsi bahan bakar (Canal et al., 2024).

Algoritma Machine Learning untuk Prediksi Konsumsi Bahan Bakar

Algoritma yang umum digunakan meliputi *Linear Regression*, *Decision Tree*, *Random Forest*, *Support Vector Regression* (SVR), dan *Extreme Gradient Boosting* (XGBoost). *Linear Regression* memiliki proses komputasi yang sederhana namun kurang optimal pada hubungan nonlinier (Mariprasath et al., 2026). *Decision Tree* mampu memodelkan hubungan kompleks

tetapi rentan mengalami *overfitting* (Shankar, 2026). *Random Forest* menggabungkan banyak pohon keputusan sehingga menghasilkan prediksi yang lebih stabil dan akurat (Li et al., 2022). SVR efektif untuk data berukuran kecil dengan hubungan nonlinier (Arda et al., 2026), sedangkan XGBoost dikenal memiliki kecepatan pelatihan tinggi, kemampuan menangani *missing value*, dan akurasi prediksi yang sangat baik (Deepthi et al., 2026).

Data ECU sebagai Feature Machine Learning

Data ECU merupakan sumber fitur utama dalam pengembangan model prediksi konsumsi bahan bakar. Parameter yang umum digunakan meliputi *RPM*, *Vehicle Speed*, *Engine Load*, *Throttle Position*, *MAP*, *MAF*, *Coolant Temperature*, *Intake Air Temperature*, dan *Fuel Rate*. Data tersebut bersifat objektif, tersedia secara *real-time*, serta mampu merepresentasikan kondisi operasional kendaraan sehingga meningkatkan akurasi model prediksi dibandingkan data simulasi (Canal et al., 2024).

Tahapan Prediksi Menggunakan Machine Learning

Pengembangan model prediksi dilakukan melalui tahapan akuisisi data ECU, *data cleaning*, normalisasi atau standarisasi, seleksi fitur, pembagian data menjadi *training* dan *testing set*, pelatihan model, serta evaluasi performa. Tahapan tersebut merupakan prosedur standar dalam pengembangan model prediksi berbasis *Machine Learning* (Pazhouh et al., 2026). Tahapan ini meliputi;

Preprocessing Data

Tahap ini untuk meningkatkan kualitas data sebelum digunakan dalam proses pemodelan. Proses ini meliputi:

- Pembersihan data (*data cleaning*) dengan memeriksa data yang hilang (*missing values*), data duplikat, serta ketidaksesuaian format atau nilai yang tidak valid sehingga diperoleh dataset yang konsisten.
- One-Hot Encoding*, yaitu mengubah variabel kategorik seperti *Transmission* dan *Fuel Type* menjadi representasi numerik dalam bentuk variabel biner agar dapat diproses oleh algoritma *Linear Regression*.
- Normalisasi menggunakan *Min-Max Scaling*, yaitu mentransformasikan nilai setiap variabel numerik ke dalam rentang 0–1 menggunakan persamaan:

$$X' = \frac{X - X_{min}}{X_{max} - X_{min}} \dots\dots\dots(i)$$

Normalisasi dilakukan untuk menyamakan skala antarvariabel sehingga proses pelatihan model menjadi lebih stabil dan efisien. Setelah proses *preprocessing* selesai, dataset dibagi menjadi 80% data pelatihan (*training set*) dan 20% data pengujian (*testing set*) menggunakan metode *hold-out validation*. Data pelatihan untuk membangun model *Linear Regression*, sedangkan data pengujian digunakan untuk mengevaluasi kemampuan model dalam melakukan prediksi terhadap data yang belum pernah dipelajari sebelumnya.

Model Training

Tahap ini menggunakan algoritma *Linear Regression*. Pada tahap ini model mempelajari hubungan linier antara variabel input dengan variabel target melalui proses estimasi koefisien regresi menggunakan metode *Ordinary Least Squares (OLS)* sehingga diperoleh persamaan regresi yang mampu memprediksi konsumsi bahan bakar kendaraan.

Evaluasi Model Machine Learning

Evaluasi model menggunakan indikator utama, yaitu *Mean Absolute Error (MAE)*, *Root Mean Square Error (RMSE)*, serta *Coefficient of Determination (R²)* (Akram, 2026). Model yang baik ditunjukkan nilai MAE, RMSE yang rendah serta nilai R² yang mendekati satu (Ibrahim & Zakzouk, 2026).

Coefficient of Determination (R²)

Koefisien determinasi (R²) merupakan ukuran statistik yang menunjukkan seberapa besar proporsi variasi pada variabel dependen yang dijelaskan oleh variabel independen dalam model regresi. Rumus R² adalah sebagai berikut:

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

Keterangan:

R² = Koefisien Determinasi

y_i = nilai aktual pada pengamatan ke-i

$\hat{y}_i$ = nilai prediksi pada pengamatan ke-i

$\bar{y}$ = rata-rata nilai aktual

n = jumlah pengamatan

Evaluasi model prediksi juga dinilai dari *Standard Error of the Estimate (SEE)* yaitu ukuran yang menunjukkan rata-rata besarnya kesalahan prediksi model regresi terhadap nilai aktual. Rumusnya sebagai berikut;

$$SEE = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n - k - 1}} \dots\dots\dots(iii)$$

Keterangan:

SEE = Standard Error of the Estimate

y_i = nilai aktual

$\hat{y}_i$ = nilai prediksi

n = jumlah observasi

k = jumlah variabel independen dalam model

Evaluasi model prediksi juga menggunakan statistik Durbin-Watson untuk mendeteksi autokorelasi pada residual model regresi. Rumusnya;

$$DW = \frac{\sum_{t=2}^n (e_t - e_{t-1})^2}{\sum_{t=1}^n e_t^2} \dots\dots\dots (iv)$$

Keterangan:

DW = Statistik Durbin-Watson

e_t = residual pada observasi ke-ttt

e_{t-1} = residual pada observasi sebelumnya

n = jumlah observasi

Mean Absolute Error (MAE)

Mean Absolute Error (MAE) merupakan metrik evaluasi untuk mengukur rata-rata nilai absolut selisih antara nilai aktual dan nilai prediksi yang dihasilkan model. MAE menunjukkan besarnya kesalahan prediksi tanpa memperhatikan arah kesalahan (positif atau negatif). Semakin kecil nilai MAE, semakin baik kemampuan model dalam melakukan prediksi. Rumusnya sebagai berikut;

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \dots\dots\dots (v)$$

Keterangan:

MAE = Mean Absolute Error

n = jumlah data pengamatan

y_i = nilai aktual pada data ke-i

$\hat{y}_i$ = nilai prediksi pada data ke-i

$|y_i - \hat{y}_i|$ = nilai absolut dari selisih antara nilai aktual dan prediksi

Root Mean Square Error (RMSE)

Root Mean Square Error (RMSE) merupakan ukuran evaluasi yang menghitung akar kuadrat dari rata-rata kuadrat selisih antara nilai aktual dan nilai prediksi. Rumus RMSE adalah sebagai berikut:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \dots\dots\dots (vi)$$

Keterangan:

RMSE = Root Mean Square Error

n = jumlah data pengamatan

y_i = nilai aktual pada data ke- i

$\hat{y}_i$ = nilai prediksi pada data ke- i

$(y_i - \hat{y}_i)^2$ = kuadrat selisih antara nilai aktual dan prediksi

Literatur Review

Penelitian tentang prediksi konsumsi bahan bakar kendaraan berbasis *Machine Learning* menunjukkan hasil yang menjanjikan. Ko (2022) membandingkan beberapa algoritma, yaitu *Linear Regression*, *K-Nearest Neighbor (KNN)*, *Decision Tree*, *Random Forest*, dan *Gradient Boosting*, menggunakan data uji kendaraan di jalan raya. Hasilnya menunjukkan bahwa algoritma berbasis *ensemble*, khususnya *Random Forest* dan *Gradient Boosting*, menghasilkan akurasi prediksi yang lebih tinggi dibandingkan regresi linier (Ko, 2022). Penelitian dengan memanfaatkan data aktual *Electronic Control Unit (ECU)* seperti *RPM*, *throttle position*, *engine load*, dan *vehicle speed* untuk memprediksi konsumsi bahan bakar serta mengklasifikasikan gaya mengemudi dengan akurasi tinggi (Canal et al., 2024). Model *Machine Learning* lebih akurat dibandingkan pendekatan berbasis model fisika dalam memprediksi konsumsi bahan bakar dan emisi CO_2 (Ibrahim & Zakzouk, 2026). Pentingnya pembaruan model secara berkala untuk mengatasi *concept drift* pada data kendaraan (Kusnawi et al., 2026).

3. METODE PENELITIAN

Penelitian ini menggunakan pendekatan kuantitatif dengan desain eksperimen komparatif (*Comparative Experimental Design*) untuk membandingkan kinerja beberapa algoritma *Machine Learning* dalam memprediksi konsumsi bahan bakar kendaraan berdasarkan parameter teknis kendaraan. Penelitian dilaksanakan di Program Studi Teknik Elektro, Kota Surakarta, pada Juni–Juli 2026 dengan memanfaatkan dataset terbuka dari *Natural Resources Canada* (NRCan) yang mencakup variabel *Engine Size*, *Cylinders*, *Transmission*, *Fuel Type*, dan *CO₂ Emissions* sebagai variabel bebas, serta *Fuel Consumption Combined (L/100 km)* sebagai variabel terikat. Data dapat diakses melalui; <https://open.canada.ca/data/en/dataset/98f1a129-f628-4ce4-b24d-6f16bf24dd64>

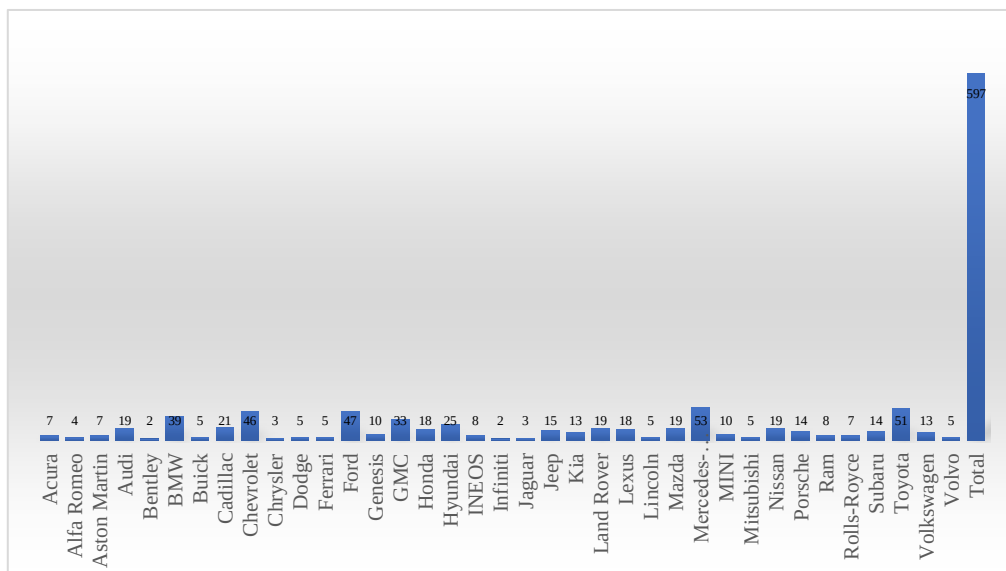
Tahapan penelitian meliputi studi literatur, pengumpulan dataset, *preprocessing* data (pembersihan data, *one-hot encoding*, dan normalisasi *Min-Max Scaling*), pembagian data menjadi 80% data pelatihan dan 20% data pengujian menggunakan metode *hold-out validation*, pelatihan model menggunakan algoritma *Linear Regression*, dilanjutkan dengan proses

prediksi serta evaluasi performa menggunakan metrik *Mean Absolute Error (MAE)*, *Root Mean Square Error (RMSE)*, dan koefisien determinasi (R^2) untuk menentukan algoritma dengan akurasi prediksi terbaik.

4. HASIL DAN PEMBAHASAN

Deskripsi Dataset Penelitian

Dataset penelitian merupakan data spesifikasi kendaraan tahun 2026 yang diperoleh dari *Natural Resources Canada (NRCAN)* yang diakses 14 Juli 2026 dengan data 597 jenis kendaraan. Dataset terdiri dari berbagai informasi teknis kendaraan yang meliputi *Model Year*, *Make*, *Model*, *Vehicle Class*, kapasitas mesin (*Engine Size*), jumlah silinder (*Cylinders*), jenis transmisi (*Transmission*), jenis bahan bakar (*Fuel Type*), konsumsi bahan bakar kondisi perkotaan (*City Fuel Consumption*), konsumsi bahan bakar jalan raya (*Highway Fuel Consumption*), konsumsi bahan bakar gabungan (*Combined Fuel Consumption*), emisi CO₂ (*CO₂ Emissions*), serta nilai efisiensi lingkungan kendaraan.



Gambar 1. Jenis Dataset Kendaraan.

Dataset penelitian yang ditunjukkan Gambar 1 berasal dari *Natural Resources Canada (NRCAN)* dan terdiri atas 597 data kendaraan. Jenis merk kendaraan paling banyak adalah Mercedes-Benz 53 unit (8,9%), diikuti Toyota 51 unit (8,5%), Ford 47 unit (7,9%), dan Chevrolet 46 unit (7,7%). Selanjutnya, BMW 39 unit (6,5%), sedangkan GMC 33 unit (5,5%). Merek lainnya memiliki proporsi yang lebih kecil, seperti Hyundai (4,2%), Cadillac (3,5%), Audi, Land Rover, Mazda, dan Nissan yang masing-masing sebesar 3,2%, serta Honda dan Lexus masing-masing 3,0%. Sementara itu, merek dengan frekuensi paling sedikit adalah Bentley dan Infiniti, masing-masing 2 unit (0,3%).

Preprocessing Data

Tahap *preprocessing* data, dilakukan pemeriksaan terhadap karakteristik statistik setiap variabel numerik untuk memastikan kualitas data sebelum digunakan dalam proses pemodelan *machine learning*. Variabel yang digunakan sebagai fitur (*feature*) adalah *Engine Size* (L), *Cylinders*, *Fuel Type*, dan *CO₂ Emissions* (g/km), sedangkan variabel target (*target variable*) yang diprediksi adalah *Combined Fuel Consumption* (L/100 km). Pemilihan variabel tersebut didasarkan pada karakteristik teknis kendaraan yang memiliki hubungan terhadap kebutuhan energi mesin dan efisiensi konsumsi bahan bakar. Hasil analisis diskriptif didapatkan sebagai berikut;

Tabel 1. Analisis Diskriptif *Engine Size* (L), *Cylinders*, *Fuel Type*, *CO₂ Emissions* (g/km) dan *Combined Fuel Consumption* (L/100 km).

Variabel	Mean	Median	SD	Min–Maks
Engine Size (L)	3.00	2.70	1.23	1.20–6.70
Cylinders	5.42	4.00	1.84	3–12
CO ₂ Emissions (g/km)	254.12	253.00	66.56	104.00–460.00
Combined Fuel Consumption (L/100 km)	10.84	10.70	2.87	4.40–19.70

Hasil analisis deskriptif Tabel 1, menunjukkan rata-rata kapasitas mesin (*Engine Size*) sebesar 3,00 L dengan median 2,70 L, standar deviasi 1,23 L, serta rentang nilai 1,20–6,70 L. Variabel *Cylinders* memiliki rata-rata 5,42 silinder, median 4 silinder, standar deviasi 1,84, dan berkisar antara 3 hingga 12 silinder, yang menunjukkan variasi konfigurasi mesin pada kendaraan dalam dataset. Rata-rata *CO₂ Emissions* sebesar 254,12 g/km dengan median 253,00 g/km, standar deviasi 66,56 g/km, serta rentang nilai 104,00–460,00 g/km. Variabel *Combined Fuel Consumption* memiliki rata-rata 10,84 L/100 km, median 10,70 L/100 km, standar deviasi 2,87 L/100 km, dan berkisar antara 4,40–19,70 L/100 km. Hasil tersebut menunjukkan adanya variasi yang cukup besar pada karakteristik mesin, tingkat emisi karbon, dan konsumsi bahan bakar kendaraan, sehingga dataset dinilai cukup representatif evaluasi model *Machine Learning* dalam memprediksi konsumsi bahan bakar kendaraan.

Hasil Data Cleaning

Tahap pertama adalah pemeriksaan kualitas data. Hasil pemeriksaan menunjukkan hasil berikut:

Tabel 2. Hasil Data *Cleaning*.

Pemeriksaan	Hasil
Jumlah data awal	597
Missing value	0
Data duplikat	0
Data tidak valid	0
Jumlah data setelah <i>cleaning</i>	597

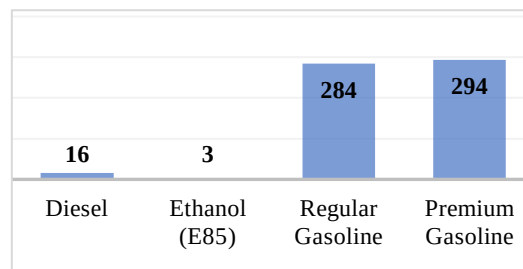
Berdasarkan analisis data *cleaning* Tabel 2 diketahui seluruh data memenuhi kriteria analisis sehingga tidak ada data yang dihapus. Dataset tetap berjumlah 597 observasi dan dapat digunakan pada tahap berikutnya.

Hasil One-Hot Encoding

Variabel kategorik berupa *Fuel Type* dan *Transmission* tidak dapat langsung diproses oleh algoritma *Linear Regression* sehingga dilakukan proses *One-Hot Encoding* dengan hasil berikut;

Tabel 3. Analisis Diskriptif Jenis Bahan Bakar (*Fuel Type*).

	Fuel Type	f	%
D	Diesel	16	2.7
E	Ethanol (E85)	3	.5
X	Regular Gasoline	284	47.6
Z	Premium Gasoline	294	49.2
	Total	597	100.0



Gambar 2. Grafik Jenis Bahan Bakar (*Fuel Type*).

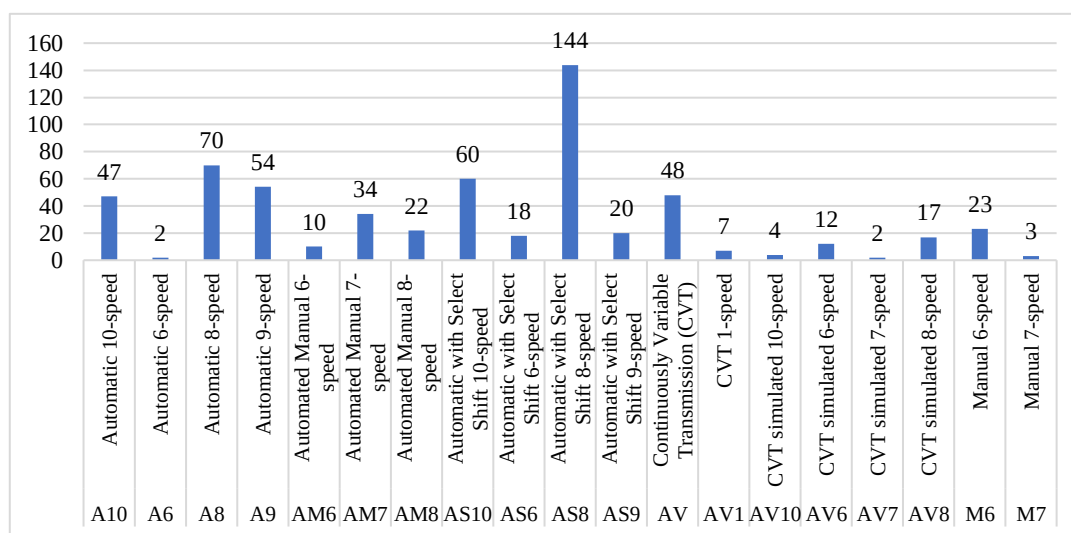
Hasil analisis jenis bahan bakar terhadap 597 kendaraan Tabel 3, diketahui paling banyak adalah Premium Gasoline (Z) 294 kendaraan (49,2%), diikuti Regular Gasoline (X) 284 kendaraan (47,6%), sementara itu kendaraan berbahan bakar Diesel (D) berjumlah 16 kendaraan (2,7%), dan Ethanol (E) merupakan jenis bahan bakar yang paling sedikit digunakan, yaitu 3 kendaraan (0,5%). Hasil tersebut menunjukkan bahwa sebagian besar kendaraan dalam dataset menggunakan bahan bakar bensin, baik bensin premium maupun reguler, yang secara keseluruhan 578 kendaraan (96,8%).

Kendaraan berbahan bakar diesel dan etanol hanya memiliki proporsi sangat kecil. Kondisi ini menunjukkan dataset didominasi kendaraan berbahan bakar bensin sehingga karakteristik konsumsi bahan bakar dan emisi yang dipelajari merepresentasikan kendaraan berbahan bakar bensin. Oleh karena itu, variabel *Fuel Type* menjadi salah satu fitur penting dalam pembangunan model *Machine Learning* karena setiap jenis bahan bakar memiliki karakteristik pembakaran, efisiensi energi, dan tingkat emisi berbeda sehingga dapat memengaruhi prediksi konsumsi bahan bakar kendaraan.

Tabel 4. Analisis Diskriptif *Transmission*.

Kode	Jenis Transmisi	f	%
A10	Automatic 10-speed	47	7,9
A6	Automatic 6-speed	2	0,3
A8	Automatic 8-speed	70	11,7
A9	Automatic 9-speed	54	9,0
AM6	Automated Manual 6-speed	10	1,7
AM7	Automated Manual 7-speed	34	5,7
AM8	Automated Manual 8-speed	22	3,7
AS10	Automatic with Select Shift 10-speed	60	10,1
AS6	Automatic with Select Shift 6-speed	18	3,0
AS8	Automatic with Select Shift 8-speed	144	24,1
AS9	Automatic with Select Shift 9-speed	20	3,4
AV	Continuously Variable Transmission (CVT)	48	8,0
AV1	CVT 1-speed	7	1,2
AV10	CVT simulated 10-speed	4	0,7
AV6	CVT simulated 6-speed	12	2,0
AV7	CVT simulated 7-speed	2	0,3
AV8	CVT simulated 8-speed	17	2,8
M6	Manual 6-speed	23	3,9
M7	Manual 7-speed	3	0,5
Total		597	100,0

Hasil analisis diskriptif jenis transmisi ditunjukkan Tabel 4 diketahui paling banyak adalah *Automatic with Select Shift 8-speed* (AS8) 144 kendaraan (24,1%), *Automatic 8-speed* (A8) 70 kendaraan (11,7%), *Automatic with Select Shift 10-speed* (AS10) 60 kendaraan (10,1%), dan *Automatic 9-speed* (A9) 54 kendaraan (9,0%). Jenis transmisi yang paling sedikit digunakan adalah *Automatic 6-speed* (A6) dan *CVT simulated 7-speed* (AV7) yang masing-masing hanya sebanyak 2 kendaraan (0,3%). Secara umum, hasil ini menunjukkan bahwa dataset didominasi oleh kendaraan dengan transmisi otomatis, khususnya tipe *Automatic with Select Shift* (AS) dan *Automatic* (A), sedangkan kendaraan dengan transmisi manual memiliki proporsi relatif kecil. Dominasi transmisi otomatis ini mencerminkan karakteristik kendaraan modern yang menjadi objek penelitian.

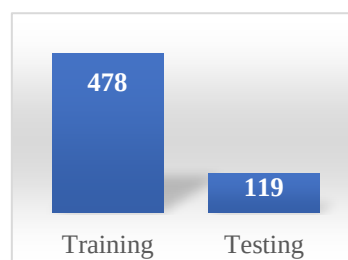
**Gambar 3.** Grafik Jenis Transmisi Kendaraan.

Hasil Normalisasi Min-Max Scaling

Variabel numerik dinormalisasi menggunakan metode *Min-Max Scaling* dengan penghitungan menggunakan persamaan rumus 1; variabel *Engine Size* (nilai minimum = 1,2 L, nilai maksimum = 6,8 L dan salah satu data = 2,4 L). Hasil penghitungan didapatkan nilai *Engine Size* 2,4 L berubah menjadi 0,214 setelah normalisasi. Seluruh variabel numerik berhasil ditransformasikan ke rentang 0–1, sehingga perbedaan skala antarvariabel dapat diminimalkan dan proses pembelajaran model menjadi lebih stabil.

Pembagian Dataset

Setelah tahap *preprocessing* selesai, dataset dibagi menggunakan metode *hold-out validation* dengan proporsi 80% data pelatihan dan 20% data pengujian. Perhitungan jumlah data adalah sebagai berikut;



Gambar 4. Pembagian Dataset Kendaraan.

Jumlah data keseluruhan (N) = 597
 Data pelatihan = $0.80 \times 597 = 477,6$ (478)
 Data pengujian = $0.20 \times 597 = 119,4$ (119)

Hasil penghitungan didapatkan sebanyak 478 data digunakan untuk membangun model *Linear Regression*, sedangkan 119 data digunakan untuk mengevaluasi kemampuan model terhadap data yang belum pernah dipelajari sebelumnya.

Evaluasi Model Prediksi

Coefficient of Determination (R²)

Penghitungan evaluasi model prediksi dengan indikator variabel penelitian dan *Coefficient of Determination (R²)* didapatkan hasil sebagai berikut:

Tabel 5. Evaluasi Model Prediksi Konsumsi Bahan Bakar dengan *Coefficient of Determination (R²)*.

Model	R				
	R	Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
1	0.989 ^a	0.977	0.977	.43190	1.855

a. Predictors: (Constant), Fuel_Type, Engine_Size, CO2_emissions, Cylinders
 b. Dependent Variable: Combined

Hasil analisis yang ditunjukkan Tabel 5 diketahui bahwa evaluasi model *Machine Learning* berbasis regresi yang dibangun menggunakan parameter *Electronic Control Unit* (ECU), yaitu *Engine Size*, *Cylinders*, *CO₂ Emissions*, dan *Fuel Type*, mampu memprediksi konsumsi bahan bakar gabungan (*Combined fuel consumption*) dengan tingkat akurasi yang sangat tinggi. Hasil penghitungan dengan rumus persamaan 2 didapatkan nilai koefisien korelasi ($R = 0,989$) yang menunjukkan hubungan sangat kuat antara nilai aktual dan nilai hasil prediksi. Tingginya nilai korelasi tersebut mengindikasikan pola konsumsi bahan bakar yang dihasilkan model hampir identik dengan pola konsumsi bahan bakar aktual, sehingga model mampu merepresentasikan hubungan antara karakteristik mesin dan konsumsi bahan bakar secara akurat. Penelitian menyebutkan bahwa parameter karakteristik kendaraan seperti kapasitas mesin, jumlah silinder, tipe transmisi, dan karakteristik mesin merupakan prediktor yang sangat berpengaruh terhadap konsumsi bahan bakar, di mana model regresi mampu menghasilkan tingkat akurasi yang tinggi ketika menggunakan variabel-variabel tersebut sebagai masukan (Guodong et al., 2025).

Kemampuan model dalam menjelaskan variasi data juga terlihat dari nilai koefisien determinasi (R^2) sebesar 0,977. Nilai tersebut menunjukkan bahwa 97,7% variasi konsumsi bahan bakar kendaraan dapat dijelaskan oleh kombinasi variabel *Engine Size*, *Cylinders*, *CO₂ Emissions*, dan *Fuel Type*, sedangkan hanya 2,3% dipengaruhi oleh faktor lain di luar model. Nilai Adjusted R^2 yang tetap sebesar 0,977 menunjukkan bahwa penambahan variabel prediktor tidak menyebabkan overfitting dan seluruh variabel yang digunakan memberikan kontribusi yang relevan terhadap kemampuan prediksi model. Dengan demikian, model yang dikembangkan memiliki kemampuan generalisasi yang baik terhadap data yang digunakan.

Hasil penghitungan dengan rumus persamaan 3 didapatkan nilai *Standard Error of Estimate* (SEE) sebesar 0,432 L/100 km, yang berarti rata-rata penyimpangan antara nilai aktual dan nilai prediksi relatif kecil. Nilai SEE yang hampir sama dengan RMSE mengindikasikan konsistensi model dalam menghasilkan prediksi pada seluruh rentang data. Dengan demikian, model tidak hanya memiliki kemampuan menjelaskan variasi data yang tinggi, tetapi juga menghasilkan estimasi yang stabil dan presisi. Sedangkan hasil penghitungan dengan rumus persamaan 4 didapatkan nilai *Durbin–Watson* sebesar 1,855 menunjukkan bahwa residual model tidak mengalami autokorelasi yang signifikan karena nilainya mendekati angka 2. Hal ini menunjukkan bahwa residual bersifat independen sehingga salah satu asumsi penting dalam regresi linear telah terpenuhi. Terpenuhinya asumsi tersebut meningkatkan keandalan model dalam menghasilkan estimasi parameter yang tidak bias dan dapat digunakan sebagai dasar pengambilan keputusan.

Hasil ini memperlihatkan bahwa parameter yang dihasilkan ECU merupakan indikator yang sangat representatif dalam menggambarkan karakteristik konsumsi bahan bakar kendaraan. *Engine Size* dan jumlah silinder (*Cylinders*) mencerminkan kapasitas dan karakteristik kerja mesin yang secara langsung memengaruhi kebutuhan energi selama proses pembakaran. Kendaraan dengan kapasitas mesin dan jumlah silinder yang lebih besar umumnya memerlukan suplai bahan bakar yang lebih tinggi untuk menghasilkan tenaga yang lebih besar (Akram, 2026). Sementara itu, CO₂ Emissions memiliki hubungan yang sangat erat dengan konsumsi bahan bakar karena jumlah karbon dioksida yang dihasilkan merupakan konsekuensi langsung dari proses pembakaran bahan bakar. Oleh karena itu, tidak mengherankan apabila variabel CO₂ Emissions memberikan kontribusi terbesar dalam model regresi yang dibangun (T. Liu, 2026).

Hasil analisis koefisien regresi diketahui bahwa CO₂ Emissions merupakan prediktor paling dominan dengan koefisien beta terstandarisasi ($\beta = 0,959$; $p < 0,001$). Temuan ini menunjukkan bahwa peningkatan emisi karbon dioksida akan diikuti oleh peningkatan konsumsi bahan bakar kendaraan. Secara teknis, hubungan tersebut sesuai dengan prinsip konservasi energi pada mesin pembakaran dalam, di mana semakin besar jumlah bahan bakar yang dibakar maka semakin besar pula emisi CO₂ yang dihasilkan. Emisi CO₂ merupakan variabel dengan kontribusi terbesar dalam model prediksi *fuel economy* karena memiliki hubungan langsung dengan efisiensi proses pembakaran mesin. Variabel (Hazra et al., 2026).

Sebaliknya, variabel *Engine Size* dan *Fuel Type* juga memberikan pengaruh signifikan terhadap konsumsi bahan bakar meskipun kontribusinya relatif lebih kecil dibandingkan CO₂ Emissions. Variabel *Cylinders* menunjukkan nilai signifikansi yang berada pada batas statistik ($p = 0,051$), sehingga pengaruhnya terhadap konsumsi bahan bakar menjadi relatif lebih kecil ketika dianalisis secara simultan bersama variabel lain. Kondisi tersebut mengindikasikan adanya hubungan yang kuat antara jumlah silinder dengan kapasitas mesin sehingga sebagian informasi yang diberikan kedua variabel saling tumpang tindih. Informasi mengenai jenis bahan bakar yang diperoleh dari sistem ECU mampu meningkatkan performa model prediksi karena berkaitan dengan strategi pembakaran mesin dan efisiensi energi (Pazhouh et al., 2026).

Mean Absolute Error (MAE)

Penghitungan *Mean Absolute Error* (MAE) menggunakan rumus persamaan 5 didapatkan hasil sebagai berikut;

$$MAE = \frac{148,28}{597} = 0,24 \text{ L/100 km}$$

Nilai *Mean Absolute Error* (MAE) sebesar 0,24 L/100 km menunjukkan bahwa rata-rata selisih absolut antara nilai aktual dan hasil prediksi hanya sekitar 0,24 liter per 100 km. Kesalahan prediksi yang relatif kecil tersebut menunjukkan bahwa model mampu menghasilkan estimasi yang mendekati kondisi sebenarnya sehingga layak digunakan sebagai alat prediksi konsumsi bahan bakar kendaraan. MAE juga merupakan metrik evaluasi yang mudah diinterpretasikan karena dinyatakan dalam satuan yang sama dengan variabel target, sehingga memberikan gambaran langsung mengenai besar kecilnya kesalahan prediksi (X. Liu & Jin, 2023).

MAE merupakan salah satu indikator evaluasi yang paling representatif karena secara langsung menggambarkan rata-rata kesalahan prediksi dalam satuan yang sama dengan variabel target sehingga lebih mudah diinterpretasikan dalam aplikasi otomotif. Model *Machine Learning* dengan nilai kesalahan prediksi yang rendah mampu memperkirakan konsumsi bahan bakar pada kondisi berkendara nyata dengan tingkat akurasi tinggi (Shahariar et al., 2023). Penelitian menyebutkan bahwa model dengan performa terbaik memperoleh MAE sebesar 0,66 L/100 km dan R^2 sebesar 0,984, yang dikategorikan memiliki akurasi sangat tinggi dalam memprediksi efisiensi bahan bakar kendaraan (Zahid & Jamil, 2026).

Root Mean Square Error (RMSE)

Hasil prediksi model terhadap 597 data, untuk mengetahui jumlah kuadrat kesalahan (*Sum of Squared Errors/SSE*) menggunakan rumus persamaan 6 didapatkan hasil sebagai berikut:

$$RMSE = \sqrt{\frac{\sum (y_i - \hat{y}_i)^2}{597}} = \sqrt{\frac{110,38}{597}} = 0,430 \text{ L/100 km}$$

Nilai *Root Mean Square Error* (RMSE) 0,430 L/100 km menunjukkan rata-rata besarnya kesalahan prediksi model terhadap nilai aktual adalah 0,430 liter per 100 km. Nilai RMSE relatif kecil mengindikasikan model *Machine Learning* memiliki tingkat akurasi yang baik dalam memprediksi konsumsi bahan bakar kendaraan. Selain itu, karena RMSE memberikan penalti yang lebih besar terhadap kesalahan prediksi yang tinggi, nilai sebesar 0,430 L/100 km menunjukkan bahwa model mampu menjaga kesalahan prediksi tetap rendah meskipun terdapat variasi pada data kendaraan yang digunakan.

Nilai RMSE yang hanya sedikit lebih besar daripada MAE menunjukkan bahwa sebagian besar residual memiliki distribusi yang relatif seragam dan model tidak banyak dipengaruhi oleh kesalahan prediksi yang sangat besar (*outlier*). Meskipun pada beberapa observasi ditemukan residual yang cukup tinggi, seperti pada kendaraan dengan karakteristik tertentu,

secara keseluruhan model tetap mampu mempertahankan tingkat akurasi yang sangat baik. Kondisi ini menunjukkan bahwa model cukup robust dalam menangani variasi karakteristik kendaraan yang terdapat pada dataset. RMSE banyak digunakan dalam evaluasi model regresi karena mampu memberikan gambaran yang lebih sensitif terhadap kesalahan prediksi yang besar. Model dengan nilai RMSE yang lebih rendah menunjukkan kemampuan prediksi yang lebih baik dibandingkan model dengan nilai RMSE yang lebih tinggi (Shahariar et al., 2023).

Penelitian menyebutkan bahwa nilai RMSE yang rendah mampu menghasilkan estimasi konsumsi bahan bakar yang lebih stabil dan akurat dibandingkan model regresi konvensional (Zhang et al., 2024). RMSE merupakan indikator evaluasi yang penting karena mampu menggambarkan sensitivitas model terhadap kesalahan prediksi yang besar, sehingga semakin kecil nilai RMSE menunjukkan semakin baik kemampuan model dalam melakukan generalisasi terhadap data baru (X. Liu & Jin, 2023). Rendahnya nilai RMSE merupakan indikator penting keberhasilan model dalam menangkap hubungan kompleks antara karakteristik mesin, kondisi operasi kendaraan, dan konsumsi bahan bakar (Shahariar et al., 2023). Temuan ini mengindikasikan bahwa data yang dihasilkan oleh ECU memiliki potensi besar untuk dimanfaatkan sebagai dasar pengembangan sistem prediksi konsumsi bahan bakar secara real-time, mendukung peningkatan efisiensi energi, pengurangan emisi karbon, serta pengembangan sistem transportasi cerdas berbasis *Machine Learning* (Szumska et al., 2025).

Hasil penelitian ini menunjukkan bahwa pendekatan *Machine Learning* berbasis regresi menggunakan data ECU memiliki potensi yang sangat besar dalam memprediksi konsumsi bahan bakar kendaraan. Tingginya nilai R^2 yang disertai nilai MAE, RMSE, dan SEE yang rendah menunjukkan bahwa informasi yang dihasilkan ECU mampu dimanfaatkan secara efektif untuk membangun model prediksi yang akurat. Model seperti ini dapat diterapkan pada sistem *on-board diagnostics* (OBD), sistem manajemen kendaraan cerdas (*intelligent transportation systems*), maupun aplikasi pemantauan efisiensi energi secara real-time. Implementasi model prediksi konsumsi bahan bakar berbasis ECU tidak hanya membantu pengguna dalam mengoptimalkan efisiensi penggunaan bahan bakar, tetapi juga berpotensi mendukung pengurangan emisi karbon melalui pengelolaan operasi kendaraan yang lebih efisien (Anitha & Babu, 2026).

Penelitian ini memiliki beberapa keterbatasan. Model hanya memanfaatkan empat variabel utama dari data ECU, sedangkan konsumsi bahan bakar pada kondisi nyata juga dipengaruhi oleh berbagai faktor lain, seperti kecepatan kendaraan, akselerasi, putaran mesin (RPM), beban kendaraan, suhu mesin, tekanan ban, kondisi jalan, kemiringan medan, serta perilaku pengemudi. Oleh karena itu, penelitian selanjutnya disarankan mengintegrasikan lebih

banyak parameter ECU maupun data sensor kendaraan secara *real-time* serta mengevaluasi algoritma Machine Learning lain, seperti *Random Forest*, *Extreme Gradient Boosting (XGBoost)*, *Support Vector Regression (SVR)*, *Artificial Neural Network (ANN)*, atau *Long Short-Term Memory (LSTM)*, sehingga dapat diperoleh model prediksi konsumsi bahan bakar yang memiliki akurasi dan kemampuan generalisasi yang lebih tinggi pada berbagai kondisi operasional kendaraan.

5. KESIMPULAN DAN SARAN

Penelitian ini menunjukkan data *Electronic Control Unit (ECU)*, yang meliputi *Engine Size*, *Cylinders*, *CO₂ Emissions*, dan *Fuel Type*, dapat digunakan sebagai parameter yang efektif dalam membangun model *Machine Learning* berbasis regresi untuk memprediksi konsumsi bahan bakar kendaraan (*Combined Fuel Consumption*). Model yang dikembangkan mampu memberikan performa prediksi yang sangat baik, ditunjukkan oleh nilai koefisien determinasi ($R^2 = 0,977$), koefisien korelasi ($R = 0,989$), serta tingkat kesalahan prediksi yang rendah dengan nilai MAE sebesar 0,24 L/100 km, RMSE sebesar 0,430 L/100 km, dan *Standard Error of Estimate* sebesar 0,432 L/100 km. Hasil tersebut mengindikasikan bahwa model memiliki tingkat akurasi yang tinggi dalam memperkirakan konsumsi bahan bakar berdasarkan karakteristik kendaraan yang diperoleh dari data ECU.

Penelitian ini memiliki keterbatasan karena model dikembangkan menggunakan satu dataset dengan jumlah variabel prediktor yang masih terbatas sehingga kemampuan generalisasinya terhadap kendaraan dengan karakteristik yang berbeda belum dapat dipastikan. Selain itu, adanya beberapa observasi dengan residual yang relatif besar menunjukkan bahwa model masih sensitif terhadap data ekstrem (*outlier*) atau kondisi operasi kendaraan yang tidak sepenuhnya terwakili dalam dataset.

Penelitian selanjutnya disarankan untuk menggunakan dataset yang lebih besar dan lebih beragam, menambahkan parameter ECU lainnya seperti putaran mesin (*engine RPM*), torsi mesin, kecepatan kendaraan, beban mesin, posisi throttle, temperatur mesin, serta parameter operasional lainnya. Selain itu, penelitian mendatang dapat membandingkan berbagai algoritma *Machine Learning*, seperti *Random Forest*, *Support Vector Regression (SVR)*, *Gradient Boosting*, *XGBoost*, maupun *Artificial Neural Network (ANN)*, serta menerapkan teknik *cross-validation* dan pengujian pada data eksternal (*external validation*) agar diperoleh model prediksi konsumsi bahan bakar yang lebih akurat, robust, dan memiliki kemampuan generalisasi yang lebih baik pada berbagai jenis kendaraan dan kondisi pengoperasian.

DAFTAR REFERENSI

- Akram, A. (2026). Fuel Consumption Prediction: A Comparative Analysis of Machine Learning Paradigms. *Computer Science*, 1(1), 1–29. <https://doi.org/10.48550/arXiv.2603.21034>
- Anitha, K., & Babu, R. (2026). Stochastic Modeling of Fuel Consumption : A Predictive Machine Learning Framework for Energy Analytics. *Journal of Computational Analysis and Applications*, 35(6), 10–18. <https://doi.org/10.48047/jocaaa.2026.35.06.02>
- Arda, E., Demir, M., Çelebi, S., Demir, Ü., & Seçgin, Ö. (2026). Leveraging Advanced Machine Learning Algorithms for Energy Efficiency and Carbon Footprint Reduction in Diesel Engines. *International Journal of Automotive Science and Technology*, 10(1), 241–259. <https://doi.org/10.30939/ijastech..1896122>
- Canal, R., Riffel, F. K., & Gracioli, G. (2024). Machine learning for real-time fuel consumption prediction and driving profile classification based on ECU data. *IEEE Access*, 12, 1–16. <https://doi.org/10.1109/ACCESS.2024.3400933>
- Deepthi, P., Nithya, M., Mounika, P., Mourya, K., Reddy, P. M. V., Ananth, C., & Rabbimovich, Y. O. (2026). Machine Learning for Fuel Consumption Prediction and Driving Profile Classification Based on ECU Data Next-Generation Engineering Aprendizaje automático para la predicción del consumo de combustible y la clasificación del perfil de conducción según datos. *Superintelligence Series*, 3, 275–290. https://doi.org/10.62486/978-9915-9851-0-7_202648
- Guodong, S., Mingmao, H., Aihong, G., Qingshan, G., Qinghe, G., & Hao, T. (2025). Vehicle fuel consumption optimization prediction model based on XGBoost-MSIWOA-LSTM. *Computer Integrated Manufacturing System*, 31(9), 3467–3484. <https://doi.org/10.13196/j.cims.2023.0232>
- Hazra, S., Tangadpalliwar, S., & Hazra, S. (2026). Data-Driven Prediction of Global Automotive Trends: Forecasting Fuel Economy and CO₂ Emissions Using Machine Learning. *Symposium on International Automotive Technology (2026)*, 1–20. <https://doi.org/10.4271/2026-26-0643>
- Ibrahim, R. A., & Zakzouk, N. E. (2026). A machine learning framework for predicting fuel consumption and CO₂ emissions in hybrid and combustion vehicles: comparative analysis and performance evaluation. *PLOS ONE*, 21(2), e0342418. <https://doi.org/10.1371/journal.pone.0342418>
- Jessica, F., Thomas, W. C., Zepani, C. L., Fadillah, M. E., & Annisa, R. (2025). Aplikasi Perkiraan Efisiensi Bahan Bakar Mobil dengan Machine Learning dan Streamlit. *JUKOMIKA (Jurnal Ilmu Komputer Dan Informatika)*, 8(2), 86–92. <https://doi.org/10.54650>
- Keskin, R., Belge, E., & Kutoglu, S. H. (2025). Optimizing Fuel Consumption Prediction Model Without an On-Board Diagnostic System in Deep Learning Frameworks. In *Sensors* (Vol. 25, Issue 22, p. 7031). <https://doi.org/10.3390/s25227031>
- Ko, K. (2022). Machine Learning Methods to Predict Vehicle Fuel Consumption. *Journal of The Korea Society of Computer and Information*, 27(9), 13–20. <https://doi.org/10.9708/jksci.2022.27.09.013>
- Kusnawi, Wibowo, M. A., & Sanjaya, R. (2026). Adaptive Gradient Boosting for Fuel Consumption Prediction in Mining Haul Trucks under Concept Drift Monitoring.

Jurnal Teknik Informatika (JUTIF), 7(2), 1758–1777.
<https://doi.org/10.52436/1.jutif.2026.7.2.5635>

- Li, Y., Zeng, I. Y., Niu, Z., Shi, J., Wang, Z., & Guan, Z. (2022). Predicting vehicle fuel consumption based on multi-view deep neural network. *Neurocomputing*, 502, 140–147. <https://doi.org/10.1016/j.neucom.2022.06.047>
- Liu, T. (2026). Advanced active learning and metaheuristic optimization framework for robust vehicle carbon footprint prediction using engine specifications and fuel efficiency metrics. *Journal of Cleaner Production*, 549, 147827. <https://doi.org/10.1016/j.jclepro.2026.147827>
- Liu, X., & Jin, H. (2023). High-precision transient fuel consumption model based on support vector regression. *Fuel*, 338, 127368. <https://doi.org/10.1016/j.fuel.2022.127368>
- Mariprasath, T., Srinivasan, P., Anbazhagan, G., Karuvelampalayam, P., Bustanji, S. Al, & Berezhnychenko, V. (2026). Vehicle-wise validated ensemble learning framework for robust electric vehicle energy consumption prediction using physics-guided features and statistical benchmarking IN IN. *Scientific Reports*, 1(1), 1–61. <https://doi.org/10.1038/s41598-026-55167-8>
- Pazhouh, S. M. M., Rastegar, N., Karami, M., Alizadehnia, S., & Amirmazlaghani, M. (2026). Fuel Consumption Prediction of Vehicles Using XGBoost Algorithm and ECU Performance Data in the New European Driving Cycle (NEDC). *Conference: 14th International Conference on Internal Combustion Engines & OilAt: Tehran*, 1–12. <https://www.researchgate.net/publication/407128558>
- Shahariar, G. M. H., Bodisco, T. A., Surawski, N., Komol, M. M. R., Sajjad, M., Chu-Van, T., Ristovski, Z., & Brown, R. J. (2023). Real-driving CO₂, NO_x and fuel consumption estimation using machine learning approaches. *Next Energy*, 1(4), 100060. <https://doi.org/10.1016/j.nxener.2023.100060>
- Shankar, G. (2026). A Comprehensive Survey on Fuel Consumption Optimization : Machine Learning , Energy Management , and Emerging Powertrain Technologies. *IRE Journals*, 9(11), 2963–2972. <https://doi.org/10.64388/IREV9I11-1718015>
- Siswanto, D., Farid, A., Sulistiarini, E. B., & Pahlevi, R. P. (2025). Tinjauan Sistematis Parameter Kinerja Mesin Bensin Yang Diukur Melalui Engine Control Unit (ECU). *The 8th Conference on Innovation and Application of Science and Technology (CIASTECH 2025) The*, 8(1), 396–405. <https://doi.org/10.31328/ciastech.v8i1.7498>
- Szumska, E. M., Pawlik, Ł., Frej, D., & Wilk-Jakubowski, J. Ł. (2025). Machine Learning Applications in Energy Consumption Forecasting and Management for Electric Vehicles: A Systematic Review. *Energies*, 18(20), 5420. <https://doi.org/10.3390/en18205420>
- Topić, J., Škugor, B., & Deur, J. (2022). Neural Network-Based Prediction of Vehicle Fuel Consumption Based on Driving Cycle Data. In *Sustainability* (Vol. 14, Issue 2, p. 744). <https://doi.org/10.3390/su14020744>
- Valladolid, J. D., & Ortiz, J. P. (2026). Energy Consumption Prediction for an Electric Vehicle Using Machine Learning: A Comparative Study of Regression, Ensemble, and LSTM-Based Models. *Vehicles*, 8(5), 1–9. <https://doi.org/10.3390/vehicles8050099>
- Zahid, M. S., & Jamil, U. (2026). Data-driven machine learning techniques for fuel economy prediction in sustainable transportation systems. *Green Energy and Intelligent Transportation*, 5(1), 100303. <https://doi.org/10.1016/j.geits.2025.100303>

Zhang, X., Zhang, Z., Liu, Y., Xu, Z., & Qu, X. (2024). A review of machine learning approaches for electric vehicle energy consumption modelling in urban transportation. *Renewable Energy*, 234, 121243. <https://doi.org/10.1016/j.renene.2024.121243>