

Predicción del desempeño del motor diésel usando estrategias de regresión en machine learning

Prediction of diesel engine performance using regression strategies in machine learning

Recibido: setiembre 17 de 2024 | Revisado: octubre 25 de 2024 | Aceptado: noviembre 12 de 2024

CÉSAR ELÍAS MENDOZA-SUÁREZ¹
MARGARITA CHEVARRIA MOSCOSO²

RESUMEN

La inteligencia artificial ha alcanzado gran potencial en el desarrollo tecnológico, especialmente en la optimización de motores de combustión interna. Esta investigación busca pronosticar el desempeño de motores diésel utilizando estrategias de regresión en machine learning. El estudio, de enfoque cuantitativo y tipo aplicado, recopila datos de un motor diésel Komatsu de 30 litros y 1200 Hp mediante pruebas dinamométricas. Se miden la potencia al freno, el torque y el consumo de combustible, monitoreando diversos parámetros operativos. Con los datos, se desarrolló un modelo de pronóstico usando regresión lineal múltiple en Python. Los resultados muestran una alta correlación entre los parámetros de entrada y salida, destacando la presión del múltiple de admisión como el más relevante. Las predicciones alcanzan altos valores de R^2 : torque (0.96), potencia al freno (0.97) y consumo instantáneo (0.98). Se determinan también los coeficientes del modelo de regresión aplicables a los parámetros de entrada. En conclusión, los algoritmos de machine learning, específicamente la regresión lineal múltiple, son efectivos para predecir el comportamiento de motores diésel en pruebas dinamométricas.

Palabras clave: Inteligencia artificial, Regresión lineal múltiple, Aprendizaje automático, machine learning, motor diésel, potencia, torque, consumo de combustible

ABSTRACT

Artificial intelligence has achieved great potential in technological development, especially in the optimization of internal combustion engines. This research seeks to forecast the performance of diesel engines using regression strategies in machine learning. The study, with a quantitative and applied approach, collects data from a 30-liter, 1200 HP Komatsu diesel engine through dynamometric tests. Brake power, torque and fuel consumption are measured, monitoring various operating parameters. Using the data, a forecasting model was developed using multiple linear regression in Python. The results show a high correlation between the input and output parameters, highlighting the intake manifold pressure as the most relevant. The predictions reach high R^2 values: torque (0.96), brake power (0.97) and instantaneous consumption (0.98). The coefficients of the regression model applicable to the input parameters are also determined. In conclusion, machine learning algorithms, specifically multiple linear regression, are effective in predicting the behavior of diesel engines in dynamometric tests.

Keywords: Artificial intelligence, Multiple linear regression, machine learning, diesel engine, power, torque, fuel consumption

- 1 Universidad Nacional de Ingeniería, Lima-Perú
- 2 Universidad Nacional Agraria La Molina, Lima-Perú

Correo electrónico de contacto:
cmendoza@uni.edu.pe

© Los autores. Este artículo es publicado por la Revista Campus de la Facultad de Ingeniería y Arquitectura de la Universidad de San Martín de Porres. Este artículo se distribuye en los términos de la Licencia Creative Commons Atribución No-Comercial – Compartir-Igual 4.0 Internacional (<https://creativecommons.org/licenses/by/4.0/>), que permite el uso no comercial, distribución y reproducción en cualquier medio siempre que la obra original sea debidamente citada. Para uso comercial contactar a: revistacampus@usmp.pe.

<https://>

Introducción

El desarrollo de la inteligencia artificial (IA) ha revolucionado numerosas áreas de la tecnología moderna, y su impacto en la industria automotriz ha sido especialmente significativo. En particular, los motores de combustión interna han experimentado una transformación en cuanto a su optimización y diagnóstico mediante el uso de algoritmos de machine learning. Estos avances han permitido a los ingenieros no solo mejorar el rendimiento de los motores, sino también predecir con mayor precisión su comportamiento, lo que se traduce en una mayor eficiencia operativa y reducción de emisiones contaminantes (Venkatesh & Murugesan, 2020).

El aprendizaje automático facilita el desarrollo de sistemas inteligentes capaces de operar de manera autónoma. Algoritmos como la regresión lineal, la regresión de vectores de soporte, y las redes neuronales han demostrado su capacidad para analizar grandes cantidades de datos históricos y generar modelos predictivos fiables. Estos algoritmos aprenden a partir de datos históricos, que son sometidos a análisis estadístico y reconocimiento de patrones. Una vez entrenado el modelo, puede predecir el comportamiento de motores diésel bajo diferentes condiciones operativas, optimizando su rendimiento y anticipando posibles fallos.

Los motores diésel son evaluados a partir de una serie de parámetros clave, entre los que se encuentran el torque de salida, la potencia al freno y el consumo de combustible. El torque de salida mide la capacidad del motor para generar fuerza rotacional, mientras que la potencia al freno es un indicador de la

eficiencia del motor para convertir energía en movimiento útil. El consumo de combustible, por otro lado, es crítico en la determinación de la eficiencia energética y el costo operativo de las máquinas. Estos parámetros son monitoreados continuamente, pero su predicción a través de modelos de machine learning puede ofrecer ventajas significativas para el mantenimiento predictivo y la mejora del rendimiento.

Diversos estudios han investigado la predicción del comportamiento de motores diésel utilizando algoritmos de machine learning. Shateri *et al.* (2024) emplearon técnicas de IA para predecir el consumo de combustible, utilizando algoritmos como redes neuronales y procesos gaussianos, logrando identificar los factores más influyentes en el consumo de diésel. Sunil *et al.* (2024) investigaron el comportamiento de motores alimentados con aceite de pirólisis, utilizando regresión lineal y demostrando que este modelo podía explicar el 93.6% de la variación del rendimiento del motor. En otros estudios, como el de Deheri y Acharya (2023), se utilizaron algoritmos como regresión de árboles de decisión (DTR) para predecir el rendimiento y las emisiones de motores con mezclas de biocombustibles, obteniendo un valor R² entre 0.9 y 0.99.

A pesar de los avances en el uso de machine learning para la predicción de parámetros de motores diésel, todavía existen desafíos relacionados con la precisión de las predicciones en condiciones operativas extremas o variables. Los modelos actuales, aunque efectivos, pueden beneficiarse de una mayor optimización y de la incorporación de nuevas técnicas de aprendizaje

automático para mejorar su rendimiento. El presente estudio tiene como objetivo desarrollar un modelo predictivo basado en machine learning que permita anticipar el comportamiento de motores diésel bajo diferentes condiciones operativas. El modelo propuesto integrará datos históricos y utilizará algoritmos de aprendizaje supervisado como la regresión lineal y las redes neuronales para generar predicciones precisas sobre parámetros clave como el torque, la potencia y el consumo de combustible, contribuyendo así a mejorar la eficiencia y reducir el impacto ambiental de estos motores.

Método

Esta investigación tuvo como objetivo predecir el comportamiento de un motor diésel a partir del análisis de una base de datos generada mediante

pruebas dinamométricas, utilizando algoritmos de machine learning, específicamente regresión lineal múltiple. Los datos recolectados se dividieron en dos conjuntos: uno para entrenamiento del algoritmo y otro para pruebas, donde se evaluará la calidad predictiva del modelo mediante métricas de regresión como el Error Absoluto Medio (MAE), Error Cuadrático Medio (MSE), Raíz del Error Cuadrático Medio (RMSE) y el coeficiente de determinación (R^2).

Se recopilaron datos de un motor diésel en V de 30 litros, 12 cilindros y sistema de inyección common rail controlado electrónicamente de la marca Komatsu (Tabla 1). Este motor fue sometido a pruebas dinamométricas utilizando un dinamómetro hidráulico, cuyas características se detallan en la Tabla 2.

Tabla 1
Especificaciones del motor ensayado

Parámetro	Especificaciones
Fabricante del motor	Komatsu
Denominación del motor	SAA12V140E-3
Modelo de máquina	HD785-7
Número de cilindros	12
Diámetro x carrera	140 mm X 165 mm
Desplazamiento total	30.5 l {30,480 cc}
Orden de inyección	R1-L1-R5-L5-R3-L3-R6-l6-R2-L2-R4-l4
Potencia nominal de salida	895kw {1,200 Hp} @1,900rpm
Torque Max	5,076 Nm {518 kgm} @1,350rpm
Velocidad máxima sin carga	2,250 +/-25 rpm
Velocidad mínima sin carga	650 +/-20 rpm
Consumo específico	207g/kW·h {154} {g/HP h}

Tabla 2*Características del dinamómetro empleado*

Parámetro	Especificaciones
Fabricante	Powertest
Modelo	35X
Tipo	Hidráulico
Potencia máxima	2100 Hp (1566 kW)
Torque máximo	7,500 lb-ft (10168 Nm)
Rpm máximo	4000 rpm

Para la lectura y registro de parámetros del motor, se utilizó una laptop equipada con el software INSITE y la interfaz electrónica INLINE 6 para la comunicación entre la laptop y el controlador electrónico del motor (ECM) desarrollado por Cummins Inc.

El banco dinamométrico se configuró para monitorear y registrar el comportamiento del motor diésel bajo diferentes niveles de carga: 20%, 50%, 80%, 90% y 100%. Se generaron 1,916 registros que incluyen parámetros de salida como potencia, torque y consumo de combustible. A través del controlador del motor, se capturaron 27 parámetros de entrada, exportables en formato *.csv y convertidos a *.xls para su análisis. Se verificó la integridad de los datos, sin encontrar valores perdidos o repetidos.

Desarrollo de la investigación

Preprocesamiento

Para mejorar la eficacia del modelo de *machine learning*, los datos fueron estandarizados y se realizó un análisis de correlación para seleccionar los parámetros de entrada más relevantes. Se utilizó Python (versión 3.8) y las librerías Pandas, Numpy, Matplotlib y Scikit-Learn. La Figura 1 muestra el mapa de calor de la matriz de correlación. Las Tablas 3, 4 y 5 presentan las correlaciones entre los parámetros de entrada y las variables de salida: Torque, Potencia y Consumo de Combustible Instantáneo, respectivamente. Asimismo, en las Tablas 6 y 7 se presentan las características del dinamómetro empleado de entrada y salida.

Figura 1
Mapa de calor de la matriz de correlación

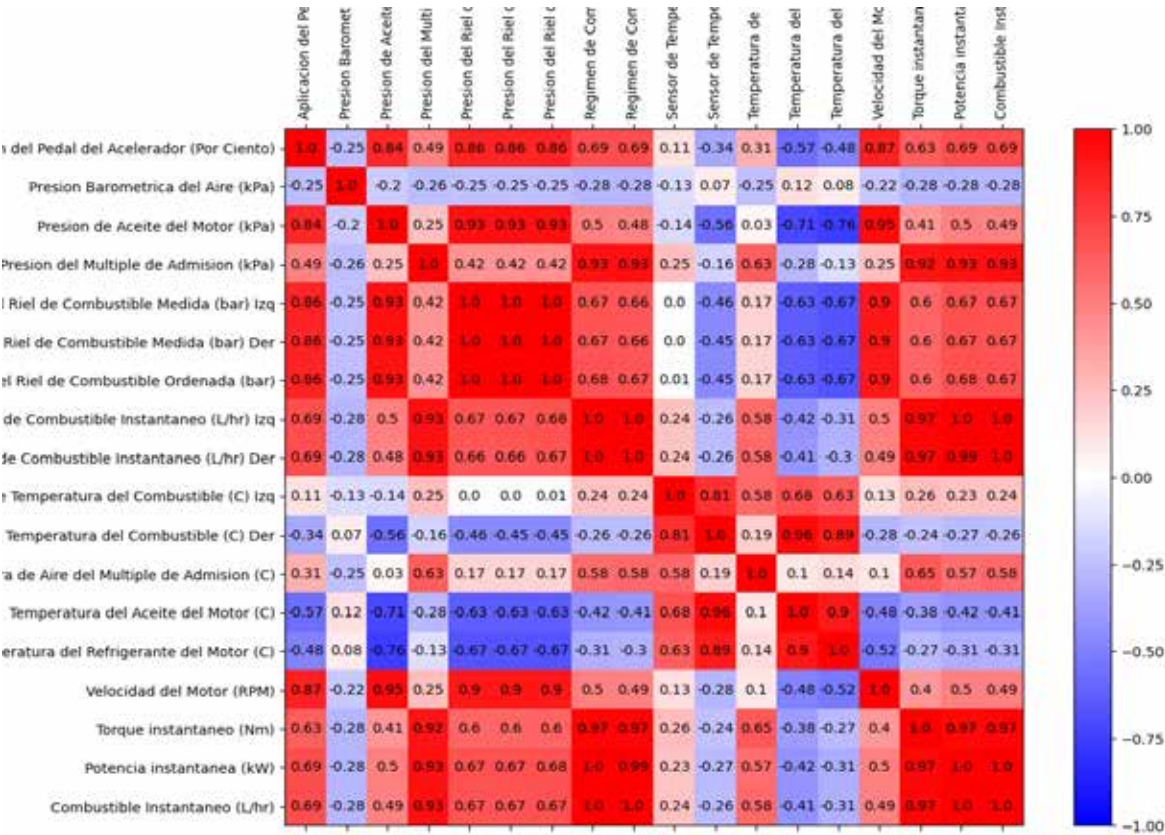


Tabla 3
Correlación para la variable Torque

Parámetro	Correlación
Presión del múltiple de admisión (kPa)	0.86052241
Presión del riel de combustible ordenada (bar)	0.64978312
Aplicación del pedal del acelerador (%)	0.60764519
Presión del riel de combustible medida (bar) der	0.57356629
Presión del riel de combustible medida (bar) izq	0.57344536
Temperatura de aire del múltiple de admisión (C)	0.49295114
Sensor de temperatura del combustible (C) izq	0.24061019
Presión de aceite del motor (kPa)	0.20552868
Velocidad del motor (RPM)	-0.06675378
Temperatura del refrigerante del motor (C)	-0.26405147
Presión barométrica del aire (kPa)	-0.27041841
Sensor de temperatura del combustible (C) der	-0.27602853

Tabla 4*Correlación para la variable potencia*

Parámetro	Correlación
Presión del múltiple de admisión (kPa)	0.91204206
Presión del riel de combustible ordenada (bar)	0.6761268
Aplicación del pedal del acelerador (%)	0.63280018
Presión del riel de combustible medida (bar) der	0.60078135
Presión del riel de combustible medida (bar) izq	0.60043003
Temperatura de aire del múltiple de admisión (C)	0.45569834
Presión de aceite del motor (kPa)	0.25854461
Sensor de temperatura del combustible (C) izq	0.20395494
Velocidad del motor (RPM)	-0.02867599
Presión barométrica del aire (kPa)	-0.26408779
Temperatura del refrigerante del motor (C)	-0.28895135
Sensor de temperatura del combustible (C) der	-0.28962989

Tabla 5*Correlación para la variable consumo de combustible instantáneo*

Parámetro	Correlación
Presión del múltiple de admisión (kPa)	0.90764256
Presión del riel de combustible ordenada (bar)	0.67824214
Aplicación del pedal del acelerador (%)	0.63510687
Presión del riel de combustible medida (bar) der	0.60177419
Presión del riel de combustible medida (bar) izq	0.60134151
Temperatura de aire del múltiple de admisión (C)	0.46372836
Presión de aceite del motor (kPa)	0.25616438
Sensor de temperatura del combustible (C) izq	0.21092699
Velocidad del motor (RPM)	-0.02913477
Presión barométrica del aire (kPa)	-0.26675232
Temperatura del refrigerante del motor (C)	-0.2838341
Sensor de temperatura del combustible (C) der	-0.28578525

Tabla 6

Características del dinamómetro empleado

Parámetro de entrada	Unidad	Instrumento
Presión del múltiple de admisión	kPa	Software Insite
Presión del riel de combustible ordenada	Bar	Software Insite
Aplicación del pedal del acelerador	%	Software Insite
Presión del riel de combustible medida der	Bar	Software Insite
Presión del riel de combustible medida izq	Bar	Software Insite
Temperatura de aire del múltiple de admisión	°C	Software Insite
Presión de aceite del motor	kPa	Software Insite
Velocidad del motor	RPM	Software Insite
Presión barométrica del aire	kPa	Software Insite
Temperatura del refrigerante del motor	°C	Software Insite
Temperatura del combustible der	°C	Software Insite
Temperatura del combustible izq	°C	Software Insite

Tabla 7

Características del dinamómetro empleado

Parámetro de salida	Unidad	Instrumento
Torque	Nm	Banco de pruebas
Potencia al freno	W	Banco de pruebas
Consumo de combustible instantáneo	L/hr	Software Insite

Modelamiento con Machine Learning

El modelado del comportamiento del motor se llevó a cabo utilizando la regresión lineal múltiple, un método

ampliamente utilizado para ajustar un plano (o hiperplano en dimensiones superiores) a los datos observados. La ecuación general del modelo es la siguiente:

$$y = a_0 + a_1x_1 + a_2x_2 + \dots + a_px_p \quad (1)$$

Donde, x_1 a x_p son las características de entrada, a_0 es el término independiente y a_1 a a_p son los coeficientes de las características. El objetivo del modelo es ajustar estos

coeficientes para minimizar los errores de predicción (Géron, 2019). La naturaleza multidimensional de tales regresiones las hace difícil de visualizar (Müller & Guido, 2017),

Desarrollo del modelo de Machine Learning

El flujo de trabajo para el desarrollo

del modelo se muestra en la Figura 2. Posteriormente, la codificación se realizó en Python, como se ilustra en la Figura 3.

Figura 2

Diagrama de flujo para el desarrollo del modelo de machine learning

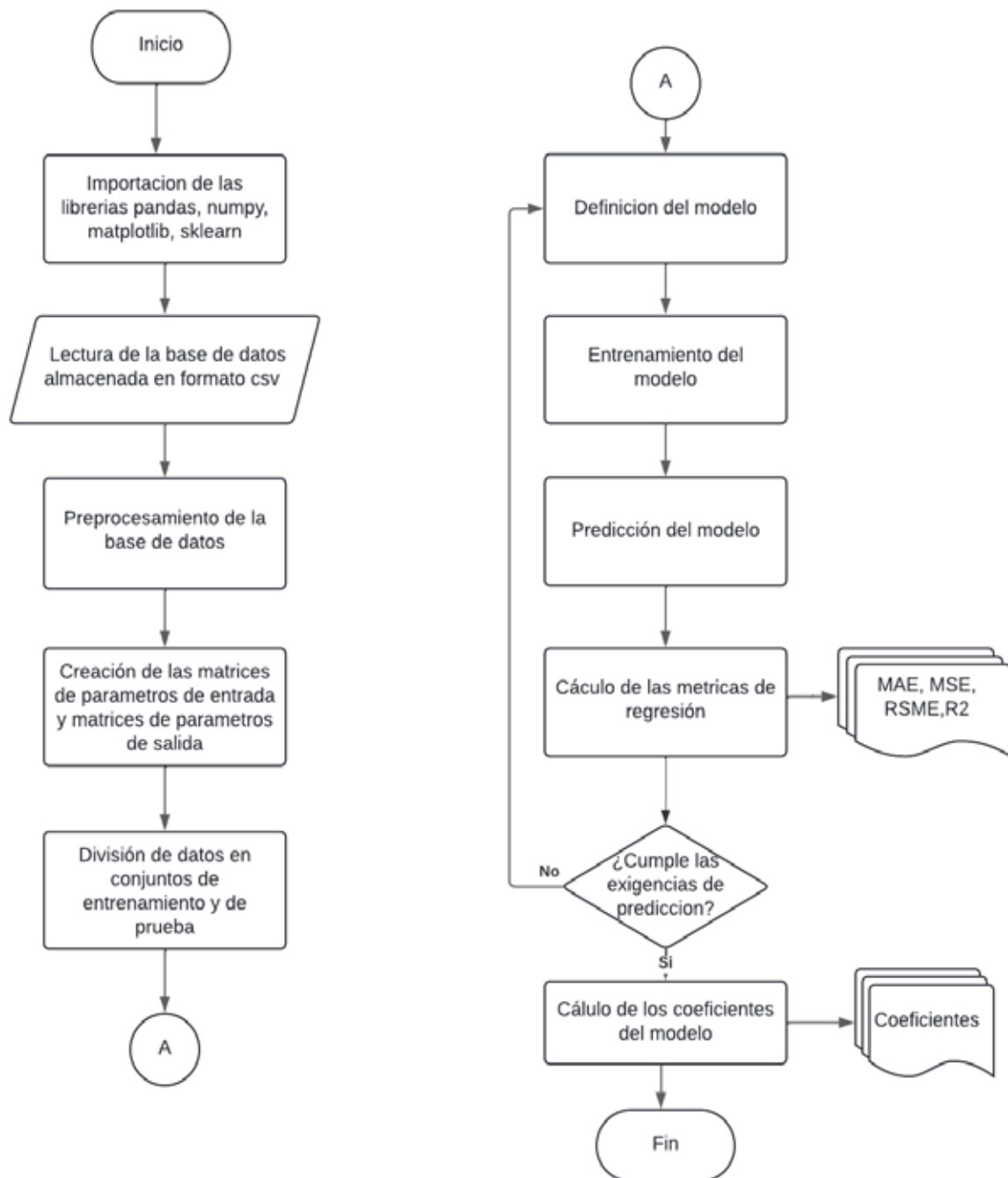


Figura 3*Extracto de la codificación en python*

```
#Regresion Lineal Multiple_01

import pandas as pd
import numpy as np
import matplotlib.pyplot as plt

datos=pd.read_csv("Datos_No_estandarizados.csv")|
Acel=datos["Aplicacion del Pedal del Acelerador (Por Ciento)"].values
P_aceite=datos["Presion de Aceite del Motor (kPa)"].values
P_bar=datos["Presion Barometrica del Aire (kPa)"].values
P_adm=datos["Presion del Multiple de Admision (kPa)"].values
P_riel_ordenada=datos["Presion del Riel de Combustible Ordenada (bar)"].values
P_riel_sensada_I=datos["Presion del Riel de Combustible Medida (bar) Izq"].values
P_riel_sensada_D=datos["Presion del Riel de Combustible Medida (bar) Der"].values
Tcomb_I=datos["Temperatura del Combustible (°C) Izq"].values
Tcomb_D=datos["Temperatura del Combustible (°C) Der"].values
T_aceite=datos["Temperatura del Aceite del Motor (°C)"].values
T_adm=datos["Temperatura de Aire del Multiple de Admision (°C)"].values
T_refrig=datos["Temperatura del Refrigerante del Motor (°C)"].values
rpm=datos["Velocidad del Motor (RPM)"].values

Par=datos["Torque instantaneo (Nm)"].values
Pot=datos["Potencia instantanea (kW)"].values
Consumo=datos["Combustible Instantaneo (L/hr)"].values
```

Resultados y Discusión

Para asegurar un desempeño satisfactorio del modelo, se estableció como criterio de validación obtener un coeficiente de determinación (R^2) superior al 95% para los parámetros de salida que describen el comportamiento del motor, definidos en la Tabla 7. El análisis de los datos preprocesados indicó una fuerte correlación entre los parámetros de entrada, como la Presión del Múltiple de Admisión (kPa), la Presión del Riel de Combustible Ordenada (bar), la Aplicación del Pedal del Acelerador (%), y otras variables, con respecto al Torque, la Potencia y el Consumo de Combustible, como se puede observar en las Tablas 3, 4 y 5. Estos resultados permitieron definir los parámetros de entrada (Tabla 6).

En este estudio, se utilizaron 1996 registros obtenidos durante las pruebas dinámicas del motor. El 80% de estos datos (1,597 registros) se utilizó para el entrenamiento del modelo, y el 20% restante (399 registros) para la validación del mismo.

Una vez ejecutado el entrenamiento del algoritmo de regresión lineal múltiple, se compararon los valores predichos con los medidos, evaluando el rendimiento del modelo a través de métricas como el Error Absoluto Medio (MAE), el Error Cuadrático Medio (MSE), la Raíz del Error Cuadrático Medio (RMSE) y el coeficiente de determinación (R^2), cuyos resultados se presentan en la Tabla 8.

Tabla 8
Métricas para evaluación de modelos de regresión

Salida	MAE	MSE	RSME	R2
Torque	0.1447	0.0447	0.2115	0.96060498
Potencia	0.1173	0.0313	0.1771	0.97057275
Consumo	0.1239	0.0373	0.1930	0.97613513

Los resultados obtenidos muestran que el modelo alcanzó un R^2 de 0.96 para el torque, 0.97 para la potencia y 0.98 para el consumo de combustible, lo que indica una excelente precisión en las predicciones.

Las gráficas correspondientes al torque, la potencia y el consumo predichos versus los valores medidos se muestran en las Figuras 4, 5 y 6, respectivamente.

Figura 4
Torque medido vs el torque predicho

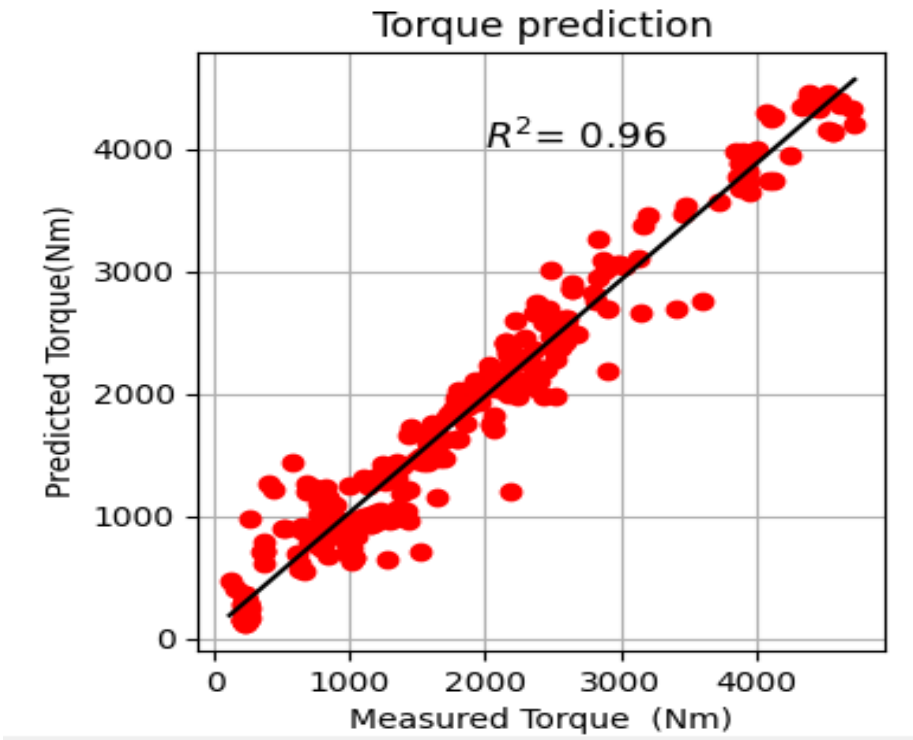


Figura 5

Potencia medida vs potencia predicha

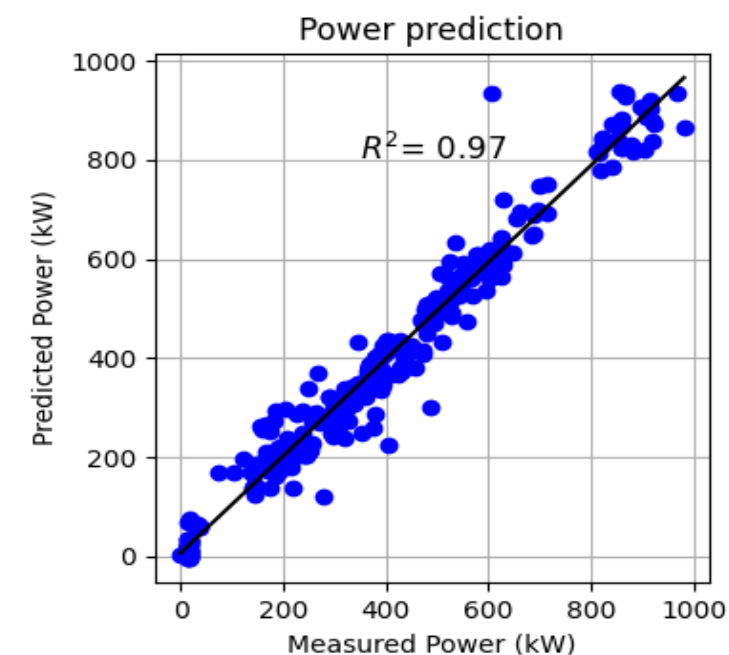
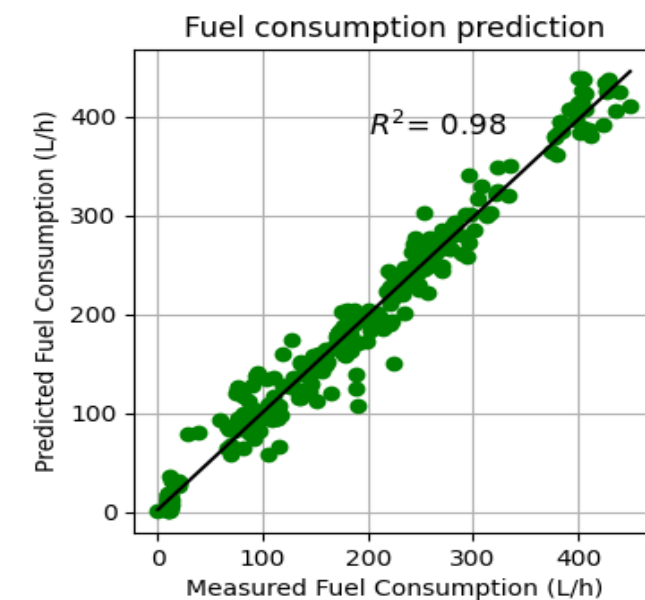


Figura 6

Consumo de combustible medido vs Consumo de combustible predicho



Del comportamiento analizado. Se calcularon los coeficientes asociados a los parámetros de entrada para cada una

de las variables de salida, como se muestra en la Tabla 9.

Tabla 9*Coefficientes de los parámetros de salida*

Parámetros de salida	Var	Torque	Potencia	Consumo
Presión del múltiple de admisión	X1	17.97	4.56	2.09
Presión del riel de combustible ordenada	X2	2.4	0.63	0.32
Aplicación del pedal del acelerador	X3	9.32	1.44	0.69
Presión del riel de combustible medida der	X4	-0.48	-0.21	-0.09
Presión del riel de combustible medida izq	X5	0.37	-0.01	-0.03
Temperatura de aire del múltiple de admisión	X6	-13.18	-3.68	-1.55
Presión de aceite del motor	X7	6.09	-0.62	-0.31
Velocidad del motor	X8	-2.56	-0.03	-0.02
Presión barométrica del aire	X9	124.29	11.51	8.73
Temperatura del refrigerante del motor	X10	5.85	-2.14	-1
Temperatura del combustible der	X11	-50.3	-7.35	-3.69
Temperatura del combustible izq	X12	88.89	10.24	4.98
Bias	-	-14706.77	-1009.43	-808.62

De tal forma que para las pruebas dinamométricas ejecutas la ecuación multilíneal será:

$$\begin{aligned}
 \text{Torque} = & 17.97x_1 + 2.40x_2 + 9.32x_3 - 0.48x_4 + 0.37x_5 - 13.18x_6 + 6.09x_7 - 2.56x_8 \\
 & + 124.29x_9 + 5.85x_{10} - 50.30x_{11} + 88.89x_{12} \\
 & - 14706.77
 \end{aligned} \quad (3)$$

$$\begin{aligned}
 \text{Power} = & 4.56 + 0.63x_2 + 1.44x_3 - 0.21x_4 + 0.37x_5 - 0.01x_6 - 0.62x_7 - 0.003x_8 \\
 & + 11.51x_9 - 2.14x_{10} - 7.35x_{11} - 10.24x_{12} - 1009.43
 \end{aligned} \quad (4)$$

Fuel consumption

$$\begin{aligned}
 = & 2.09x_1 + 0.32x_2 + 0.69x_3 - 0.09x_4 - 0.03x_5 - 1.55x_6 - 0.31x_7 \\
 & - 0.02x_8 + 18.73x_9 - x_{10} - 3.69x_{11} + 4.98x_{12} - 808.62
 \end{aligned} \quad (5)$$

Conclusiones

Los algoritmos de machine learning tienen una aplicación efectiva en la predicción del comportamiento de motores de combustión interna, como se ha demostrado tanto en la revisión de la literatura como en los resultados obtenidos en este estudio.

En particular, el algoritmo de regresión lineal múltiple ha mostrado una alta capacidad para predecir el comportamiento del motor diésel, específicamente en términos de potencia al freno, torque y consumo de combustible instantáneo, cuando el motor es sometido a pruebas dinamométricas. Estas predicciones se basan en la lectura de parámetros de funcionamiento clave

como la Presión del Múltiple de Admisión (kPa), la Presión del Riel de Combustible Ordenada (bar), la Aplicación del Pedal del Acelerador (%), la Presión del Riel de Combustible Medida (bar) tanto en el lado derecho como izquierdo, y la Temperatura del Aire del Múltiple de Admisión (°C).

El análisis de la correlación entre los parámetros de entrada y las salidas predichas revela que la Presión del Múltiple de Admisión (kPa) tiene el mayor impacto en las predicciones. Este parámetro, además, es fácilmente medible tanto en condiciones de laboratorio como en campo, lo que lo convierte en una variable fundamental para la evaluación del rendimiento del motor, ya sea en un entorno controlado o en su aplicación en vehículos operativos.

Referencias

- Deheri, C., & Acharya, S. K. (2023). Experimental and statistical evaluation of biohythane fuelled thermal barrier coated engine using machine learning algorithms. *International Journal of Global Warming*, 31(1). <https://doi.org/10.1504/IJGW.2023.133208>
- Géron, A. (2019). Hands-on machine learning with Scikit-Learn, Keras and TensorFlow: concepts, tools, and techniques to build intelligent systems. In *O'Reilly Media*.
- Müller, A. C., & Guido, S. (2017). Introduction to Machine Learning with Python: a guide for data scientist. *O'Reilly Media, Inc.*
- Shateri, A., Yang, Z., & Xie, J. (2024). Utilizing Artificial intelligence to identify an Optimal Machine learning model for predicting fuel consumption in Diesel engines. *Energy and AI*, 16. <https://doi.org/10.1016/j.egyai.2024.100360>
- Sunil, K., Surakasi, R., Patro, S. G. K., Govil, N., Ramis, M. K., Razak, A., Sharma, P., Alsubih, M., Islam, S., Khan, T. M. Y., Almakayeel, N., & Chintakindi, S. (2024). Performance, Combustion, and Emission analysis of diesel engine fuelled with pyrolysis oil blends and n-propyl alcohol-RSM optimization and ML modelling. *Journal of Cleaner Production*, 434. <https://doi.org/10.1016/j.jclepro.2023.140354>

Venkatesh, K., & Murugesan, S.
(2020). Prediction of Engine
Emissions using Linear Regression
Algorithm in Machine Learning.
International Journal of Innovative

*Technology and Exploring
Engineering*, 9(7), 962–968.
[https://doi.org/10.35940/ijitee.
G5707.059720](https://doi.org/10.35940/ijitee.G5707.059720)