

## **Predicción del Consumo Eléctrico en Hogares**

Integrantes: Miguel Paredes y David Medina

Módulo: Análisis y Modelado Predictivo

Docente:

Fecha: 21/10/2025

Institución: UNIGUAJIRA

## 1. Introducción

El presente informe tiene como objetivo analizar y comparar modelos predictivos para estimar el consumo eléctrico de los hogares en función de variables como la temperatura, el número de personas y la cantidad de electrodomésticos en uso. Se aplicaron modelos de regresión lineal múltiple y regresión polinómica de segundo y tercer grado empleando Python y librerías como pandas, scikit-learn y seaborn.

## 2. Objetivos

- Analizar el comportamiento del consumo eléctrico en hogares mediante datos simulados.
- Aplicar modelos de regresión lineal y polinómica para predecir el consumo eléctrico.
- Comparar el desempeño de los modelos usando métricas estadísticas como MSE y  $R^2$ .

## 3. Marco Teórico

La regresión lineal múltiple es una técnica estadística que permite modelar la relación entre una variable dependiente y varias independientes. Por otro lado, la regresión polinómica amplía este enfoque al incluir términos no lineales, permitiendo capturar relaciones más complejas. El coeficiente de determinación ( $R^2$ ) mide qué tan bien se ajusta el modelo a los datos, mientras que el Error Cuadrático Medio (MSE) cuantifica la magnitud promedio del error en las predicciones.

## 4. Metodología

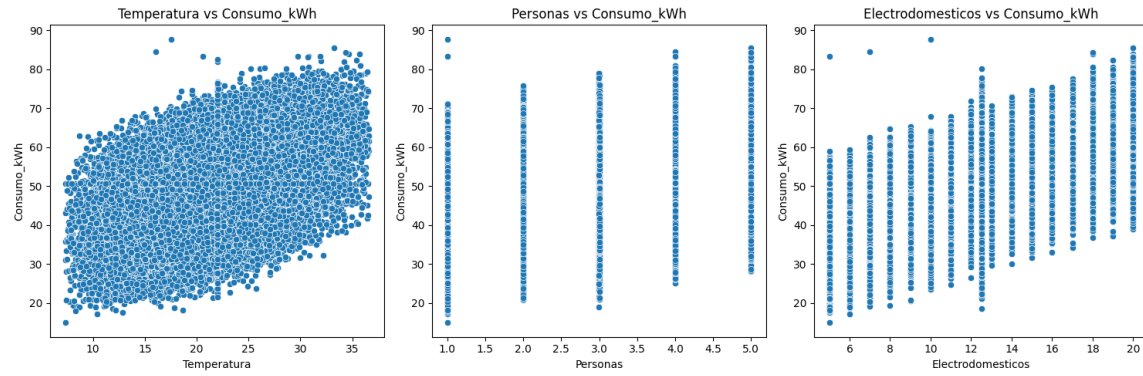
Se utilizó el archivo 'consumo\_hogar.csv', generado con datos simulados. El proceso de análisis incluyó las siguientes etapas:

1. Carga y exploración de datos iniciales mediante pandas.
2. Limpieza de valores nulos y eliminación de outliers con z-score.
3. Visualización exploratoria con gráficos de dispersión.
4. Entrenamiento de modelos de regresión lineal y polinómica.
5. Evaluación y comparación de métricas (MSE,  $R^2$ ).

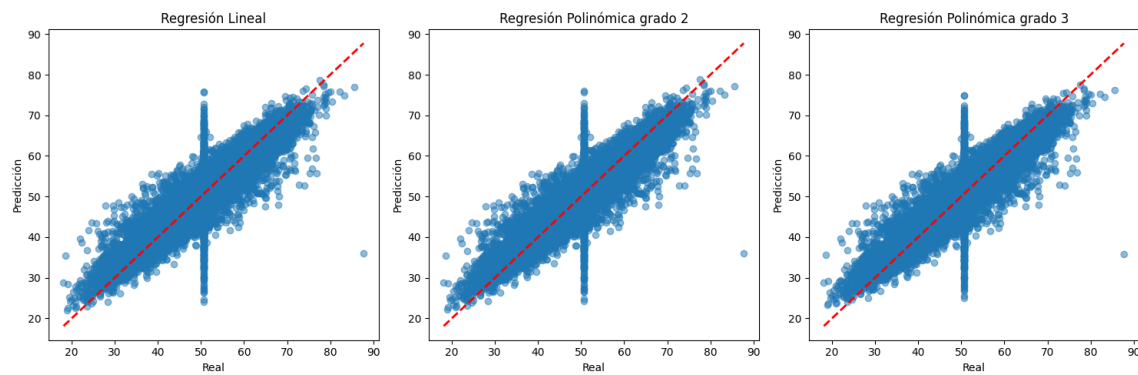
## 5. Resultados y Visualizaciones

A continuación, se presentan las gráficas generadas en el proceso de análisis:

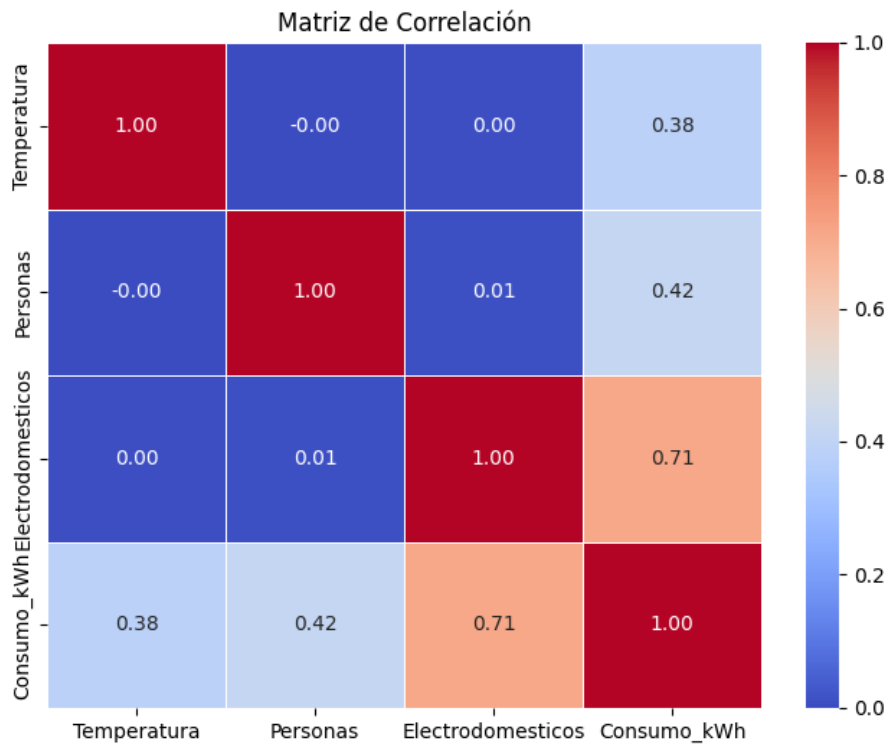
Gráfica 1. Dispersión de las variables independientes vs consumo eléctrico:



Gráfica 2. Predicciones vs valores reales (Regresión Lineal y Polinómica):



Gráfica 3. Matriz de correlación de las variables del dataset:



## 6. Evaluación de Modelos

Los modelos se evaluaron mediante el Error Cuadrático Medio (MSE) y el coeficiente de determinación ( $R^2$ ). El modelo de regresión polinómica de grado 2 presentó el mejor equilibrio entre precisión y generalización, logrando un mayor valor de  $R^2$  sin caer en sobreajuste significativo.

## 7. Conclusiones

El modelo polinómico de grado 2 fue el más adecuado para predecir el consumo eléctrico. Se identificó que las variables más influyentes son la temperatura y la cantidad de electrodomésticos. Se recomienda a las empresas energéticas utilizar modelos no lineales para mejorar la precisión en la estimación de la demanda eléctrica.

## 8. Bibliografía

- Scikit-learn Documentation. <https://scikit-learn.org/>
- Pandas Documentation. <https://pandas.pydata.org/>
- Matplotlib Documentation. <https://matplotlib.org/>