

MODELO DE MACHINE LEARNING CAPAZ DE PREDECIR EL CONSUMO DE PETRÓLEO UTILIZANDO ECUACIONES NORMALES

Adriana Carolina Buelvas Bohórquez T00035794

Este problema nos pide predecir el consumo de petróleo en determinada población, este consumo fue medido en 18 estados de EU, para un total de **48** registros de los cuales utilizaremos el **60%** de estos para entrenar el algoritmo, estos serían un total de **29 datos** y para realizar las pruebas utilizaríamos los datos restantes que son el **40%** que corresponde a **19 datos** en total.

Para calcular el porcentaje de error utilizaremos el MAPE. Se había utilizado el método del gradiente descendente para predecir el modelo, en esta oportunidad se utilizará ecuaciones normales el cual facilita el cálculo de theta porque evita realizar algoritmos basados en múltiples iteraciones. Los conjuntos de los datos no serán normalizados.

De acuerdo a los valores obtenidos en la aplicación del algoritmo y aplicando ecuaciones normales para el set de datos **petrol_data.txt**, se evidencia que los valores de **theta** son:

θ_1	θ_2	θ_3	θ_4	θ_5
2,07	-1,04	-8,31	4,00	1,35

Por consiguiente, la función $h(x)$ que nos permite predecir el consumo de petróleo es:

$$h(x) = \theta_0 + (2,07)x_1 + (-1,04)x_2 + (-8,31)x_3 + (4,00)x_4 + (1,35)x_5$$

El valor para el **MAPE** es igual a **9,0%**

Se puede concluir de acuerdo a las actividades realizadas que en el modelo con datos normalizado existe un MAPE elevado, a comparación del MAPE de los datos no normalizados, esto es debido a que, con valores cercanos a cero, este proporciona una imagen distorsionada del error cuando este es promediado.