

A decorative background featuring a network diagram with nodes and lines. The nodes are represented by circles of varying sizes and colors (blue, grey, and white), connected by thin grey lines. Some nodes are highlighted with a blue outline. The network is distributed across the slide, with a denser concentration on the left side and a more sparse arrangement on the right.

Predicciones del Rendimiento del Combustible de un Auto

Alejandra Silva H.
06-10-2022

Contenidos

- Introducción
- Problema
- Conjunto de Datos
- Visualizaciones
- Modelo
- Evaluación
- Recomendaciones



1.

Introducción

Hoy en día se considera fundamental cuidar el medioambiente debido al cambio climático y hay pocos contaminantes tan severos como el petróleo.

¿Se puede hacer algo para minimizar la huella de carbono? Lo primero que se debe hacer es tomar conciencia de este problema e informarse.

En el escenario actual, en que las ventas de autos han superado los récords de años anteriores, ¿se puede tomar alguna decisión que impacte menos en el medioambiente?


¿Se puede predecir el rendimiento del combustible de un auto?



2.

Problema

¿Se puede predecir el rendimiento del combustible de un auto basado únicamente en sus características propias? ¿O hay otros factores externos que inciden?



3.

Conjunto de Datos

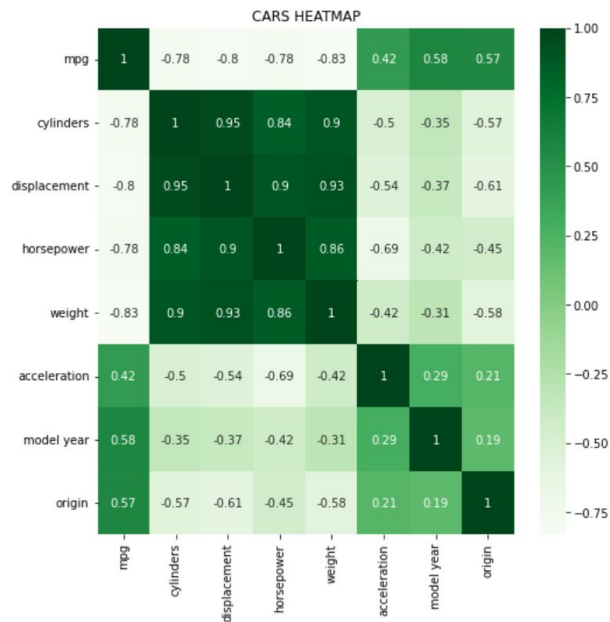
Este es un conjunto de datos bastante pequeño, de un poco menos de 400 registros y 9 columnas.

Entrega información sobre el millaje, caballos de fuerza, año del modelo, peso, velocidad de aceleración de 0 a 60 mph, origen y cantidad de cilindros.

| mpg | cylinders | displacement | horsepower | weight | acceleration | model year | origin | brand |
|------|-----------|--------------|------------|--------|--------------|------------|--------|------------|
| 46.6 | 4 | 86.0 | 65 | 2110 | 17.9 | 80 | 3 | mazda |
| 44.6 | 4 | 91.0 | 67 | 1850 | 13.8 | 80 | 3 | honda |
| 44.3 | 4 | 90.0 | 48 | 2085 | 21.7 | 80 | 2 | volkswagen |

4.

Visualizaciones



Se interpreta del mapa que mpg (nuestra columna target) tiene una relación bastante fuerte con el número de cilindros, cilindrada, caballos de fuerza y peso del auto, por lo que se entiende que dichas columnas inciden directamente en las predicciones que se harán más adelante.

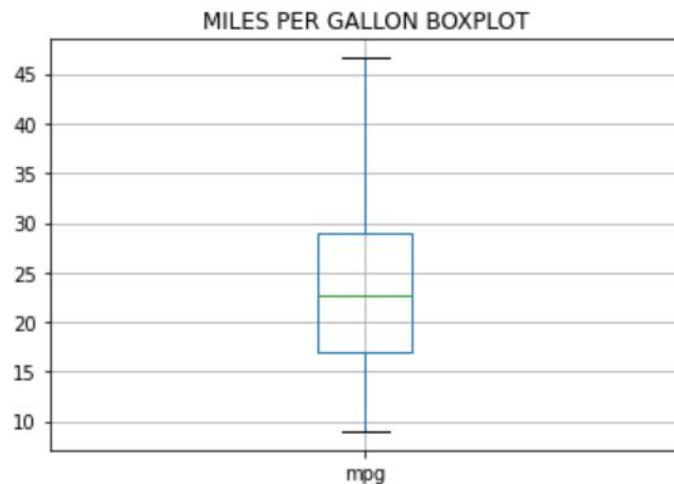
En cambio, el año y la procedencia del auto tienen una relación bastante moderada, más cercanas a la columna aceleración que es la con menos incidencia. Esta última columna mencionada muestra una relativamente baja relación con la columna target.

4.

El boxplot muestra que es una distribución simétrica, dado que la mediana está dentro de la caja, dividiéndola en 2.

En este caso, la media (23.45) es levemente mayor que la mediana (22.75) pero esa pequeña diferencia y la ubicación en la que se corta la caja indica que la distribución es prácticamente simétrica.

Los bigotes de la caja indican los valores menor y mayor encontrados en la columna mpg. Como se aprecia, no se ven valores atípicos o fuera de rango.



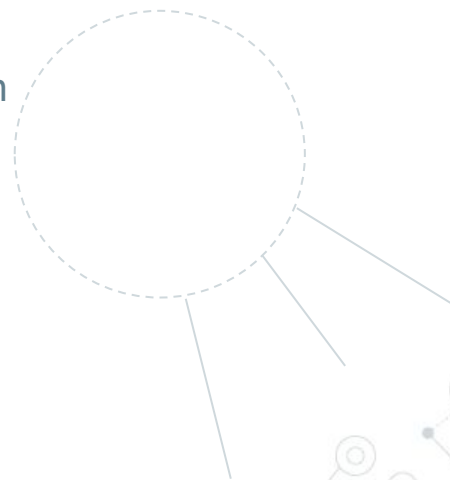


5.

Modelo

Se utilizaron varios modelos, entre ellos:

- Modelos Secuenciales (Redes Neuronales) con y sin regularización
- Árbol de Decisión de Regresión
- Modelo de Regresión Lineal



6.

Evaluación

De todos los modelos, el que tuvo mejores resultados fue el Árbol de Decisión de Regresión.

| INDICADOR | DATO |
|---------------------|---------------------------|
| Dec Tree Test Score | 77% 0.7715364508969077 |
| Dec Tree EAM Test | 2.594635265010326 |
| MSE Test | 14.733589881810326 |
| RMSE Test | 3.8384358639698966 |

Si bien el score es menor al 80%, de todas formas es un excelente punto de partida. El MAE es relativamente pequeño, 2.6 y el RMSE es solo de 3.8, siendo bastante pequeño considerando que las millas por galón van desde 9 a 46.6.



7.

Recomendaciones

Si bien nuestro modelo tuvo un 77% de score, no es un número tan alto como hubiésemos esperado, lo que podría significar que además de todas las características consideradas por el modelo, pueden haber otros factores que inciden en el rendimiento del combustible de un auto.

De todas formas es importante considerar que los factores como la cilindrada, caballos de fuerza y el peso del auto parecen tener una incidencia directa en el rendimiento del combustible.

Usar el aire acondicionado, tener los neumáticos con poco aire y la falta de mantención son otros factores que pueden incidir en un rendimiento más bajo de lo esperado, lo que conlleva gastos extra de combustible y polución en mayor grado.

A decorative background featuring a network diagram with nodes and lines. The nodes are represented by circles of varying sizes and colors (blue, grey, white), connected by thin grey lines. Some nodes are highlighted with a blue outline. The network is distributed across the slide, with a denser concentration on the left side and a more sparse arrangement on the right.

Predicciones del Rendimiento del Combustible de un Auto

Alejandra Silva H.
06-10-2022