

CARACTERÍSTICAS QUE AFECTAN EL CONSUMO DE COMBUSTIBLE EN LOS CARROS

Javier Camilo Rueda Serrano – 2141380

Juan Felipe Silva Garcés – 2141380

Resumen—Este proyecto que se llevó a cabo en el curso de Inteligencia Artificial I, implementa el uso de Machine Learning Regression, donde inicialmente se toman los datos registrados en un dataset que posee diferentes características acerca del desempeño del motor, teniendo en cuenta principalmente el desplazamiento y la cantidad de cilindros, los cuales influyen en el consumo de combustible de un vehículo. Del mismo modo es posible predecir la procedencia de los vehículos gracias al comportamiento que tuvo cada marca en su país de origen, teniendo mejor o peor optimización de potencia, preocupándose por restar peso al vehículo, o por disminuir el consumo. Este dataset comprende vehículos que se fabricaron entre 1970 y 1982, un periodo que marcó un cambio significativo en la industria automotriz de los Estados Unidos. La producción estadounidense sufrió un drástico cambio de los pesados y potentes pero poco optimizados motores de seis y ocho cilindros, a motores más ligeros de cuatro cilindros, que aunque contaban con menos potencia, presentaban mayor eficiencia de combustible. La industria automotriz mundial, incluidos los estadounidenses y sus competidores europeos y japoneses, aumentó la cantidad de millas por galón (MPG) en general al enfocarse en producir automóviles con motores de cuatro cilindros y lograr mayor eficiencia de combustible.

Abstract—This project implements the use of Machine Learning Regression, where it becomes a dataset with different characteristics about the performance of an engine, taking into account mainly the displacement and the number of cylinders, which influence in the fuel consumption of a vehicle. In the same way, it is possible to predict the operation of the vehicles thanks to the behavior that was had in their country of origin, in order to improve the optimization of power, worry about the weight of the vehicle, or to reduce consumption. This dataset includes vehicles that were manufactured between 1970 and 1982, a period that marked a significant change in the United States automotive industry. US production took a drastic change from the heavy and powerful but poorly optimized six and eight cylinder engines, but then introducing the use of lighter four-cylinder engines, although they had less power, were more fuel efficient.

I. INTRODUCCIÓN

En este documento se presentará la descripción del proyecto que se lleva a cabo en el curso de Inteligencia Artificial I, que comprende principalmente el tema de regresión, nos basamos en una problemática que se vive diariamente, como lo es el consumo de combustible de los automóviles, el cual es una

de las principales preocupaciones de las marcas fabricantes de vehículos. Por lo tanto recurrimos a un dataset que comprende diferentes aspectos o características de los motores que influyen en su rendimiento de gasolina, como lo son la cantidad de cilindros, el cilindraje y el peso principalmente. Tomamos estas características y a partir de ello realizamos un análisis gráfico de los datos para poder llevar a cabo una regresión lineal que permitiera predecir el comportamiento a futuro, de acuerdo a los parámetros dados.

Fig. 1
RESUMEN DEL DATASET

	count	mean	std	min	25%	50%	75%	max
ID	398.0	199.500000	115.036951	1.0	100.250	199.5	298.750	398.0
mpg	398.0	23.514573	7.815984	9.0	17.500	23.0	29.000	46.6
cylinders	398.0	5.454774	1.701004	3.0	4.000	4.0	8.000	8.0
displacement	398.0	193.425879	104.269838	68.0	104.250	148.5	262.000	455.0
horsepower	398.0	104.449749	38.541172	46.0	75.000	93.5	125.000	230.0
weight	398.0	2970.424623	846.841774	1613.0	2223.750	2803.5	3608.000	5140.0
acceleration	398.0	15.568090	2.757689	8.0	13.825	15.5	17.175	24.8
model	398.0	76.010050	3.697627	70.0	73.000	76.0	79.000	82.0
origin	398.0	1.572864	0.802055	1.0	1.000	1.0	2.000	3.0

Como podemos apreciar en la tabla (Fig. 1) estos son los diferentes datos que se encuentran en el dataset: el identificador del carro (ID), la distancia que alcanza a recorrer el auto dada en Millas con un Galón de combustible o Millas por Galón (mpg), la cantidad de cilindros que posee el motor (cylinders), el cilindraje del motor dado en pulgadas cúbicas (displacement), la potencia del motor dada en caballos de fuerza (horsepower), el peso del vehículo dado en libras (weight), el tiempo que le toma al vehículo en llegar de 0 a 100km/h (acceleration), el año de producción del vehículo (model), y por último el origen geográfico del vehículo (origin). De esta tabla tomamos las características que más llegan a influir en el consumo del vehículo, las cuales son cylinders, displacement, weight, model y origin.

Usamos principalmente las herramientas brindadas en el curso como clasificadores en los datasets, como lo son RandomForest, Logistic Regression y Decision Tree. Para los cuales nos basamos en el origen y la cantidad de cilindros para encontrar el acierto de la predicción, arrojando los siguientes resultados (Fig. 2 y Fig. 3):

Fig. 2
ORIGEN

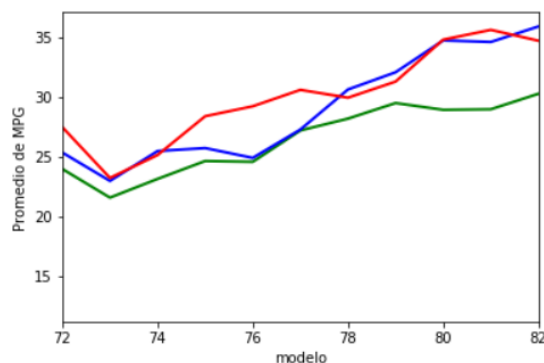
```
Random Forest
acierto en train 0.989949748744
acierto en test 0.768844221106
Logistic Regrssion
acierto en train 0.773869346734
acierto en test 0.768844221106
Decision Tree
acierto en train 0.909547738693
acierto en test 0.748743718593
```

Fig. 3
CILINDROS

```
Random Forest
acierto en train 0.991596638655
acierto en test 0.95625
Logistic Regrssion
acierto en train 0.945378151261
acierto en test 0.9375
Decision Tree
acierto en train 1.0
acierto en test 0.93125
```

Del análisis notamos que el origen de los vehículos tiene gran influencia en el consumo, ya que en el caso de la cultura americana, su tendencia era a los motores de grandes desplazamientos y gran cantidad de cilindros, además de producir automóviles robustos que implicaban mayor peso, por lo tanto, lejos de unirse a la gran tendencia de evolución que si seguía el mercado Europeo (Azul) y Japonés (Rojo), a pesar de que el mercado American (Verde) mejoró sus cifras con los años, seguía arrojando las peores cifras de consumo, como podemos verlo en la gráfica (Fig. 4)

Fig. 4
MODELO VS MPG



La característica presentada que mayor relación tiene con el consumo es el peso, como podemos apreciarlo en la imagen (Fig. 5), donde así mismo podemos ver que va ligado a la cantidad de cilindros, al tener esta última característica, el motor se vuelve más pesado, por lo tanto el consumo aumenta. Del mismo modo se halla una relación existente entre el desplazamiento con el consumo, como es obvio, al tener mayor volumen el motor, se necesita mayor aire para llenar los cilindros y realizar la combustión, así mismo demanda mayor

cantidad de combustible, por lo tanto el consumo aumenta. Al mismo tiempo que al tener mayor desplazamiento, el motor se hace más grande y por tanto más pesado. Y como mencionamos anteriormente, al aumentar el número de cilindros, así mismo aumenta el desplazamiento, por lo tanto, tiene una relación similar con el consumo de combustible como podemos apreciar en la gráfica (Fig. 6).

Fig. 5
PESO VS MPG

(4 CILINDROS – AZUL, 6 CILINDROS – VERDE, 8 CILINDROS – AMARILLO)

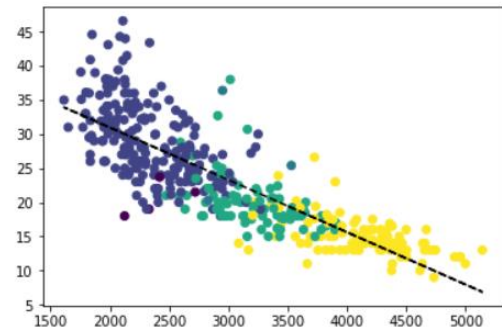
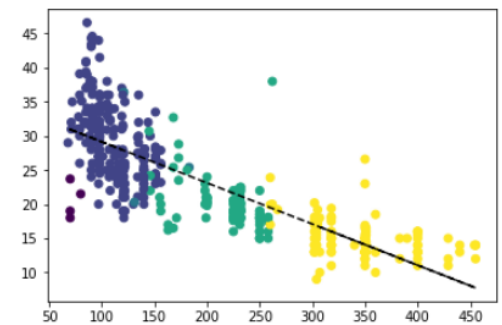


Fig. 6

DESPLAZAMIENTO VS MPG

(4 CILINDROS – AZUL, 6 CILINDROS – VERDE, 8 CILINDROS – AMARILLO)



II. CONCLUSIÓN

El peso de un automóvil es un determinante importante de su eficiencia de combustible, como lo expresan las relaciones que encontramos de acuerdo al consumo de combustible (MPG). Encontramos que a menor cantidad de cilindros, usualmente el desplazamiento es menor, por lo tanto el gasto de combustible es menor al tener menos volumen por llenar, y al tener menos cilindros se reduce la cantidad de veces que se inyecta combustible en cada ciclo. Del mismo modo, al tener un motor más pequeño en desplazamiento y número de cilindros, también logramos menor peso en el vehículo, pues el motor es la parte que más peso agrega a los automóviles. De esta forma, al tener menos peso, el vehículo necesita menor esfuerzo para moverse, por tanto el consumo de combustible se reduce.

Finalmente concluimos que a menor cantidad de cilindros, menor desplazamiento y por lo tanto, menor consumo.

REFERENCIAS

- [1] <https://github.com/RodolfoViana/exploratory-data-analysis-dataset-cars>