

Desafío - Regresión Lineal

En este desafío validaremos nuestros conocimientos aprendidos de regresión lineal. Para lograrlo, necesitarás aplicar los modelos aprendidos en clase, utilizando de apoyo el archivo fish.csv.

Lee todo el documento antes de comenzar el desarrollo **individual**, para asegurarte de tener el máximo de puntaje y enfocar bien los esfuerzos.

Tiempo asociado: 2 horas cronológicas

Descripción

Este conjunto de datos es un registro de las ventas en el mercado de 7 especies comunes de peces diferentes. Con este conjunto de datos, se busca realizar un modelo predictivo utilizando datos adecuados para máquinas y estimar el peso de los peces basándose en la longitud y ancho del pez.

El conjunto de datos incluye información sobre varias dimensiones del pez, que se utilizan para predecir su peso. Cada fila en el conjunto de datos representa una entrada de un pez individual y contiene varias columnas con información relevante. Las dimensiones del pez y otros atributos en el conjunto de datos podrían incluir:

- **Species:** La especie del pez.
- **Weight:** El peso del pez, que es la variable que se intenta predecir.
- **Length1, Length2, Length3:** Diferentes longitudes del pez.
- **Height:** Altura del pez.
- **Width:** Ancho del pez.

Los datos en este archivo permiten realizar análisis estadísticos y modelado predictivo para determinar cómo las dimensiones del pez están relacionadas con su peso. Esto puede ser útil en la industria pesquera y en la investigación científica para comprender mejor las relaciones entre las características físicas de los peces y su peso. Para esto se te solicita:

1. Crea una nueva columna que incluya el volumen de pez, asumiéndolo como un cilindro. Para un cilindro con una base circular, el área de la base (A) es igual a πr^2 , donde "r" es el radio de la base del cilindro, y la altura (h) es la distancia entre las dos bases circulares.
2. Verifica la correlación posible entre las variables, numérica y gráficamente. ¿Aumenta la correlación entre el peso y longitudes de los peces, altura y ancho al añadir el volumen? Explica.

3. Construye un modelo de regresión lineal que relacione el volumen y el peso de los peces. Realiza una breve evaluación del modelo y gráfícalo.
4. ¿Qué sucede si aplicas el modelo anterior por separado a cada especie de peces? Explica.
5. Construye un modelo de regresión que relacione dos o más variables con el peso de los peces. Realiza una breve evaluación del modelo.

Requerimientos

1. Analiza correlaciones entre variables de un dataset, en forma numérica y gráfica.
(2 Puntos)
2. Construye modelos de relación lineal en Python.
(5 Puntos)
3. Evalúa modelos de regresión lineal.
(3 Puntos)



¡Mucho éxito!

Consideraciones y recomendaciones

Considera que el pez es un cilindro por tanto el volumen de este será:

Volumen = Área de la base × Altura

Para un cilindro con una base circular, el área de la base (A) es igual a πr^2 , donde "r" es el radio de la base del cilindro, y la altura (h) es la distancia entre las dos bases circulares.

Por lo tanto, la ecuación general para el volumen (V) de un cilindro es:

```
np.pi*df['Height']*(df['Width']/2)**2
```