

3. Del dataset del punto 1 realice en PYTHON, cinco algoritmos de preprocesamiento; dos de ellos deben ser onehotencoder y otro escalado. Explique por cada uno la razón de su uso.

1. Algoritmo de preprocesamiento: Onehotencoder sobre la columna Sexo

2. Algoritmo de preprocesamiento: Onehotencoder sobre la columna CategoríaObesidad

Primero utilice **LabelEncoder** para transformar etiquetas de texto en números enteros y luego **OneHotEncoder** para convertir esos números en vectores binarios, lo que es útil para mantener la independencia entre las categorías. Esto nos sirve para predecir la clasificación de obesidad con precisión.

3. Algoritmo de preprocesamiento: Imputación sobre las columnas Altura y Peso

Es posible que los datos de altura y peso tengan valores faltantes en tu conjunto de datos. La imputación es el proceso de rellenar esos valores faltantes con valores estimados, como la media o la moda.

4. Algoritmo de preprocesamiento: Normalización sobre las columnas Altura y Peso

5. Algoritmo de preprocesamiento: Escalado sobre las columnas Altura y Peso

En el caso de altura y peso, pueden tener rangos muy diferentes. Por ejemplo, una persona puede tener una altura de 170 centímetros y un peso de 70 kilogramos.

La normalización se trata de ajustar las características para que tengan la misma longitud o magnitud. Por ejemplo la altura se mide generalmente en centímetros y el peso en kilogramos. Estas unidades tienen magnitudes diferentes, es por esa razón que se tuvo que **normalizar** estas características, de tal forma que tengan una misma longitud o magnitud igual a 1.

El **escalado** se trata de ajustar la escala de todas las características para que estén en un rango similar. Por ejemplo en la característica altura y peso, Si la altura va desde 150 cm hasta 180 cm, mientras que el peso va desde 50 kg hasta 100 kg, están en diferentes escalas. El escalado ajusta estas características para que estén en el mismo rango, como de 0 a 1 o -1 a 1.