

Técnica	Qué hace	Ventajas (por qué)	Desventajas (por qué)	Casos de uso recomendados
Label Encoding	Asigna un número entero a cada categoría.	<ul style="list-style-type: none"> - Fácil de aplicar: Solo necesitas una columna con números. - No aumenta columnas: Útil si tienes muchas variables. 	<ul style="list-style-type: none"> - Introduce un orden artificial: El modelo puede pensar que una categoría es mayor o menor que otra, aunque no tenga sentido. - Confunde modelos lineales. 	<ul style="list-style-type: none"> - Cuando la variable tiene un orden real (Ej: "Bajo", "Medio", "Alto") - Modelos basados en árboles, que no se ven afectados por el orden.
One-Hot Encoding	Crea una columna binaria (0/1) por cada categoría.	<ul style="list-style-type: none"> - No introduce orden: Cada categoría es tratada de forma independiente. - Modelos lineales o redes lo entienden bien porque comparan cada columna por separado. 	<ul style="list-style-type: none"> - Aumenta muchas columnas: Si hay muchas categorías, puede multiplicar el número de columnas. - Requiere más memoria y tiempo. 	<ul style="list-style-type: none"> - Cuando la variable no tiene orden (Ej: "Rojo", "Azul", "Verde") - Modelos lineales, SVM o redes neuronales.
Target Encoding	Sustituye cada categoría por el promedio del target asociado.	<ul style="list-style-type: none"> - Capta relación con la variable objetivo: Puede mejorar mucho la predicción si una categoría está muy relacionada con el target. - No aumenta columnas. 	<ul style="list-style-type: none"> - Sobreajuste: Si no se usa validación cruzada o suavizado, el modelo puede aprender demasiado bien los datos de entrenamiento. - No sirve para clasificación sin cuidado. 	<ul style="list-style-type: none"> - Cuando hay muchas categorías y datos suficientes. - Competiciones de machine learning o modelos donde se busca máximo rendimiento.