

Trabajo Práctico: Regresiones Lineal y Logística

Ejercicio 1

Un conjunto de datos registra:

- Velocidad: velocidad promedio del vehículo en km/h
- Consumo: litros de combustible consumidos cada 100 km

Tareas:

1. Ajustar un modelo de regresión lineal que relacione la Velocidad con el Consumo.
2. Representar en un gráfico los datos reales y la línea de regresión.
3. Estimar el consumo de un vehículo que circula a 110 km/h.
4. ¿Qué velocidad se asocia con el menor consumo según el modelo? Justificá si es confiable esa estimación.

Ejercicio 2

Un dataset contiene datos de:

- Edad: en años
- Salario: en pesos

Tareas:

1. Ajustar un modelo de regresión lineal simple para predecir el Salario en función de la Edad.
2. Agregar nuevas edades de prueba (por ejemplo: 25, 35, 45, 55 años) y estimar sus salarios.
3. Mostrar la gráfica con los datos reales, la regresión y las predicciones.

Ejercicio 3

El archivo `estudiantes.csv` contiene:

- `Horas_estudio`: cantidad de horas semanales dedicadas al estudio
- `Nota_final`: nota obtenida en el examen final

Tareas:

1. Crear un modelo de regresión lineal que relacione `Horas_estudio` con `Nota_final`.
2. Evaluar el coeficiente de determinación R^2 .
3. Determinar si el modelo es adecuado para predecir la nota de un estudiante que estudia 12 horas por semana.

Ejercicio 4

Disponés de un dataset con las siguientes variables:

- `Edad`: edad de los clientes
- `Compra`: 1 si realizó una compra, 0 si no

Tareas:

1. Ajustar un modelo de regresión logística para predecir la probabilidad de compra según la edad.
2. Dibujar la curva logística y los datos reales.
3. Calcular la probabilidad de compra para una persona de 30 años.

Ejercicio 5

El archivo `aprobacion.csv` contiene:

- `Horas_estudio`: cantidad de horas estudiadas

- Aprobo: 1 si aprobó, 0 si no

Tareas:

1. Crear un modelo de regresión logística para predecir si un estudiante aprueba.
2. Determinar la probabilidad de aprobación para un estudiante que estudia 8 horas.
3. Calcular la matriz de confusión si se considera como aprobado cuando la probabilidad es mayor a 0.5.

Ejercicio 6

Un dataset contiene:

- Nivel_glucosa: medida de glucosa en sangre
- Diabetico: 1 si la persona tiene diabetes, 0 si no

Tareas:

1. Entrenar un modelo de regresión logística para predecir Diabetico en función del Nivel_glucosa.
2. Graficar la curva logística.
3. Evaluar la precisión del modelo con un conjunto de prueba.