

## ✓ Template para uso de framework (scikit-learn)

En términos generales, debemos seguir los siguientes pasos:

1. Importar módulos
2. Cargar datos
3. Separar datos en subconjuntos
4. Entrenar el modelo
5. Analizar su desempeño
6. Usar el modelo para nuevas estimaciones (datos no vistos)

Los parametros que yo seleccione fueron 20,000 iteraciones con un eta pequeño, de 0.002, para que tenga el espacio y el incremento correcto para ser graficado

```
1 # Importar módulos
2 import pandas as pd
3 # --- linear_model (regresión y clasificación)
4 from sklearn import linear_model
5 # --- model_selection (train_test_split)
6 from sklearn.model_selection import train_test_split
7 # --- metrics (para puntajes)
8 from sklearn import metrics
9 import matplotlib.pyplot as plt
10
11
12 # Cargar datos
13 df = pd.read_csv('/content/Valhalla23.csv')
14 # Separar datos en subconjuntos (usando train_test_split)
15 X = df[['Celsius']] # Variables predictoras
16 y = df['Valks']
17
18 X_train, X_test, y_train, y_test = train_test_split(X, y, te
```

```
19
20 # Entrenar el modelo
21 # --- Crear objeto del modelo
22 model = linear_model.SGDRegressor(eta0=0.002, max_iter=20000
23 # Ajustar el modelo los datos de entrenamiento
24 model.fit(X_train, y_train)
25
26 # Evaluar el modelo en el conjunto de entrenamiento
27 train_score = model.score(X_train, y_train)
28 print(f'Score en entrenamiento: {train_score}')
29
30 # Realizar predicciones en el conjunto de prueba
31 y_pred = model.predict(X_test)
32
33 # Calcular el error cuadrático medio (MSE) en el conjunto de
34 test_error = metrics.mean_squared_error(y_test, y_pred)
35 print(f'Error cuadrático medio en prueba: {test_error}')
36
37 # Graficar los resultados
38 y_pred_all = model.predict(X)
39 plt.figure(figsize=(10, 6))
40 plt.scatter(X, y, color='blue', label='Valores Reales')
41 plt.scatter(X, y_pred_all, color='red', label='Predicciones')
42 plt.plot(X, y_pred_all, color='red', label='Línea de Predicc')
43 plt.xlabel('Celsius')
44 plt.ylabel('Valks')
45 plt.title('Predicciones vs Valores Reales (Todos los Datos)')
46 plt.legend()
47 plt.show()
```



Score en entrenamiento: 0.9759576394445424

Error cuadrático medio en prueba: 172.22316640079526



