



UNIVERSIDAD DE GUAYAQUIL
INGENIERIA INDUSTRIAL
INGENIERÍA EN SISTEMAS DE INFORMACIÓN
TALLER



TÍTULO

REGRESIÓN LINEAL

ALUMNO:

MADELINE MUÑOZ VILLEGAS

MATERIA:

CIENCIA DE DATOS

DOCENTE:

ING. HECTOR DULCEY ASTUDILLO

CURSO:

8VO SEMESTRE NOCTURNO

AÑO LECTIVO:

CICLO I

2023-2024

Desarrollo

La regresión lineal es un modelo matemático que describe la relación entre varias variables. Los modelos de regresión lineal son un procedimiento estadístico que ayuda a predecir el futuro. Se utiliza en los campos científicos y en los negocios, y en las últimas décadas se ha utilizado en el aprendizaje automático. La tarea de la regresión en el aprendizaje automático consiste en predecir un parámetro (Y) a partir de un parámetro conocido X.

Características:

1. **Linealidad:** La regresión lineal asume que la relación entre las variables es lineal, es decir, se puede representar mediante una línea recta o un hiperplano en el espacio.
2. **Mínimos cuadrados:** La técnica de regresión lineal busca minimizar la suma de los cuadrados de las diferencias entre los valores observados y los valores predichos por el modelo.
3. **Coeficientes:** La regresión lineal calcula los coeficientes que ponderan las variables predictoras para predecir la variable dependiente. Estos coeficientes representan el cambio en la variable dependiente por unidad de cambio en la variable independiente.
4. **Supuestos:** La regresión lineal se basa en varios supuestos, como la normalidad de los errores, la homocedasticidad (igual varianza de los errores), la independencia de los errores y la linealidad de la relación.
5. **Independencia:** los residuos deben ser independientes entre sí.
6. **Homocedasticidad:** debe haber homogeneidad en las varianzas de los residuos, es decir, la variabilidad de los residuos debe ser constante.

7. **No multicolinealidad:** las variables explicativas incluidas en el modelo no pueden estar relacionadas entre sí o, al menos, su relación debe ser muy débil.

Ejemplo en Python:

```
import numpy as np
import matplotlib.pyplot as plt
from sklearn.linear_model import LinearRegression

# Generamos datos de ejemplo
np.random.seed(0)
x = np.random.rand(50, 1) * 10
y = 2 * x + 1 + np.random.randn(50, 1) * 2

# Creamos el modelo de regresión lineal
model = LinearRegression()

# Ajustamos el modelo a los datos
model.fit(x, y)

# Coeficientes del modelo
coeficiente = model.coef_[0][0]
intercepto = model.intercept_[0]

# Predicciones
y_pred = model.predict(x)

# Visualización de los resultados
plt.scatter(x, y, label='Datos reales')
plt.plot(x, y_pred, color='red', label='Regresión lineal')
plt.xlabel('Variable Independiente')
plt.ylabel('Variable Dependiente')
plt.legend()
plt.show()

print(f"Coeficiente: {coeficiente}")
print(f"Intercepto: {intercepto}")
```