MODULO 06 - EJERCICIO 04-A

ALEXIS YURI M.

Analiza distintos datasets y determina si la relación entre la variable independiente y la dependiente es lineal o no lineal. Elige el modelo mas adecuado y justifica por qué.

1) Observación de los gráficos.

- Dataset 1 (Horas de estudio vs Nota final): nube de puntos con tendencia creciente casi lineal, sin curvatura evidente.

- Dataset 2 (Variable X vs Variable Y): clara curvatura cóncava (forma de parábola), la pendiente cambia de positiva a negativa (inflexión).

2) Identificación de la relación

- Dataset 1: la relación puede representarse con una línea recta con pendiente positiva.

- Dataset 2: la relación cláramente no es lineal; una recta subestimaría el modelo en los extremos.

3) Clasificación del problema.

- Dataset 1: Se clasifica como una Regresión lineal.

- Dataset 2: Se clasifica como una Regresión no lineal (curvilínea).

4) Propuesta de modelos adecuados.

Dataset 1 (lineal):

- Modelo Propuesto es una Regresión lineal simple.

- Justificación: Tendencia aproximadamente lineal. El modelo es fácil de entender y lo bastante bueno para resolver el problema sin complicarlo de más.

Dataset 2 (no lineal):

- Modelo Propuesto es una Regresión polinómica de grado 2.

- Justificación: El patrón muestra un máximo y luego cambio de pendiente. Un modelo cuadrático lo interpreta correctamente.

5) Plan de validación

1. Dividir datos en entrenamiento y prueba (p. ej. 80/20) o usar validación cruzada.

2. Ajustar modelos:

- Dataset 1: LinearRegression.

- Dataset 2: PolynomialFeatures + LinearRegression (grado 2–3).

3. Evaluar con métricas: R², MAE, RMSE.