

# Modelos de Predicción y Clasificación

Dr. José Juan Castro Alva

Primavera 2025

## Objetivo

Analizar datos mediante la construcción implementación e interpretación de modelos de regresión múltiple, regresión logística, LDA (Linear Discriminant Analysis) y QDA (Quadratic Discriminant Analysis).

## Instrucciones

Descarga una base de datos para la cual puedas implementar los modelos mencionados en el apartado de objetivos. elabora un reporte en formato ipynb con la siguiente estructura.

1. **Portada con datos generales**
2. **Descarga de la Base de Datos** Establece el link de donde obtuviste la base de datos
3. **Análisis exploratorio:** Antes de modelar, realiza un análisis descriptivo y visual de los datos. Usa gráficos de dispersión, histogramas y boxplots para explorar relaciones y distribuciones.
4. **Construcción de los Modelos**
  - **Regresión Múltiple:** Selecciona una variable de interés (por ejemplo, niveles de glucosa) y modela cómo las demás variables la predicen.
  - **Regresión Logística:** Modela la probabilidad de tener la condición médica en función de los factores de riesgo.

- **Análisis Discriminante Lineal (LDA) y Análisis Discriminante Cuadrático (QDA):** Entrena ambos modelos usando las mismas características para predecir la condición médica.

**5. Interpretación de los Modelos**

Explica los coeficientes obtenidos en los modelos de regresión múltiple y logística. Interpreta los resultados de LDA y QDA.

**6. Cálculo e Interpretación de las Métricas de Evaluación y la Matriz de Confusión**

Evalúa los modelos de clasificación (Regresión Logística, LDA y QDA) mediante las métricas: *precisión*, *sensibilidad*, *especificidad* y *F1*. Calcula la *matriz de confusión* para cada modelo. Explica lo que cada métrica revela acerca del rendimiento de cada modelo.

**7. Comparación de Modelos de Clasificación**

Compara el rendimiento de los modelos basándote en las métricas y matriz de confusión. Identifica ventajas y desventajas de cada modelo, considerando aspectos como precisión y facilidad de interpretación.

**8. Conclusiones**

Resume los hallazgos, destacando el modelo de mejor rendimiento. Justifica la elección del mejor modelo en función de su desempeño y aplicabilidad en problemas de clasificación de salud.