**Regresión Logística**

**Presentado Por**

**Luis Rodriguez**

**Jose Olascuaga**

**Kelvis Ariza**

**William Álvarez**

**Víctor Arzuza**

**Corporación Universitaria Latinoamericana**

**(CUL) Proyecto I**

**Ingeniería de sistemas**

**Barranquilla-Atlántico**

**Concepto de Regresión Logística:**

La regresión logística es una técnica estadística utilizada para predecir una variable categórica binaria basada en variables predictoras. A diferencia de la regresión lineal, que se utiliza para variables continuas, la regresión logística estima la probabilidad de que una observación pertenezca a una de las dos categorías posibles.

**Características de la Regresión Logística:**

La variable dependiente es categórica y binaria (por ejemplo, éxito/fallo, sí/no).

Las variables predictoras pueden ser tanto categóricas como continuas.

Estima la probabilidad de pertenecer a una categoría utilizando una función logística.

El resultado se interpreta como las probabilidades de pertenecer a cada categoría.

**Ventajas de la Regresión Logística:**

Es una técnica robusta y ampliamente utilizada para problemas de clasificación binaria.

No requiere supuestos de linealidad entre las variables predictoras y la variable dependiente.

Proporciona información sobre las probabilidades de pertenencia a cada categoría.

Puede manejar variables categóricas y continuas como predictores.

**Desventajas de la Regresión Logística:**

No es adecuada para predecir variables dependientes categóricas con más de dos categorías.

Puede sufrir problemas de multicolinealidad si las variables predictoras están altamente correlacionadas.

Requiere un tamaño de muestra adecuado para obtener resultados confiables.

**Aplicación de la Regresión Logística:**

La regresión logística se aplica en diversos campos, como:

Análisis de riesgo crediticio.

Investigación médica y epidemiología.

Estudios de mercado y análisis de comportamiento del consumidor.

Detección de fraudes.

Predicción de resultados electorales.

Análisis de factores que influyen en el abandono escolar.

**Validación del Funcionamiento de la Regresión Logística:**

Para validar el funcionamiento de un modelo de regresión logística, se utilizan diversas técnicas, entre las que se encuentran:

División de los datos en conjuntos de entrenamiento y prueba para evaluar el rendimiento del modelo en datos no vistos.

Utilización de métricas de evaluación como la precisión, el área bajo la curva ROC o la matriz de confusión.

Validación cruzada, que permite obtener estimaciones más estables del rendimiento del modelo al repetir el proceso de entrenamiento y evaluación en diferentes subconjuntos de datos.

**Referencia bibliográfica:**

Hosmer Jr, D. W., Lemeshow, S., & Sturdivant, R. X. (2013). Applied Logistic Regression (3rd ed.). Wiley.