

# Regresión logística

Práctica III

# Especificaciones

- De manera individual realice lo siguiente:
  - Cargue el dataset *breast-cancer.csv*. En este dataset la columna *diagnosis* es el target y puede tomar los valores M que indica que el cáncer es maligno y B que indica que es benigno. El resto de las columnas son las características
  - Separe el dataset en un conjunto de entrenamiento (90%) y un conjunto de prueba (10%) utilizando *shuffle*
  - Utilizando el conjunto de entrenamiento, cree un conjunto de validación de tres pliegues
  - Utilizando el conjunto de validación realice los siguientes experimentos
    - Entrene un modelo de regresión logística ajustando lo siguientes parámetros
      - Solver. Especifica el algoritmo mediante el cual se ajustan los pesos
      - Penalty. Indica la norma de penalización de los algoritmos de ajuste
    - Debe revisar la documentación para saber cuáles son los valores de estos parámetros y saber cuáles son compatibles entre ellos. [https://scikit-learn.org/stable/modules/generated/sklearn.linear\\_model.LogisticRegression.html](https://scikit-learn.org/stable/modules/generated/sklearn.linear_model.LogisticRegression.html)
    - Reporte la exactitud (accuracy) promedio de los pliegues

# Especificaciones

- Seleccione el modelo con los parámetros que haya obtenido la mejor exactitud promedio y entrénelo con el conjunto de entrenamiento (90% de los datos)
- Utilice el modelo entrenado para clasificar las instancias del conjunto de prueba (10% de los datos)

# Evidencias

- El código fuente
- Un documento donde se reporte lo siguiente:
  - Tabla con los parámetros utilizados y exactitud promedio de los tres pliegues
  - Indicación de los parámetros seleccionados y la exactitud promedio que mejor valor obtuvo
  - Exactitud obtenida por el modelo en el conjunto de pruebas