|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Assignment#01**  **Deadline: 06/07/2021**  Note: Attached papers are just for information.   |  |  | | --- | --- | | **Breast Cancer Coimbra Data Set**  **Abstract**: Clinical features were observed or measured for 64 patients with breast cancer and 52 healthy controls. |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | | **Data Set Characteristics:** | Multivariate | **Number of Instances:** | 116 | **Area:** | Life | | **Attribute Characteristics:** | Integer | **Number of Attributes:** | 10 | **Date Donated** | 2018-03-06 | | **Associated Tasks:** | Classification | **Missing Values?** | N/A | **Number of Web Hits:** | 98971 |   **Source:**  Miguel Patrício(miguelpatricio **'@'** gmail.com), José Pereira (jafcpereira **'@'** gmail.com), Joana Crisóstomo (joanacrisostomo **'@'** hotmail.com), Paulo Matafome (paulomatafome **'@'** gmail.com), Raquel Seiça (rmfseica **'@'** gmail.com), Francisco Caramelo (fcaramelo **'@'** fmed.uc.pt), all from the Faculty of Medicine of the University of Coimbra and also Manuel Gomes (manuelmgomes **'@'** gmail.com) from the University Hospital Centre of Coimbra  **Data Set Information:**  There are 10 predictors, all quantitative, and a binary dependent variable, indicating the presence or absence of breast cancer. The predictors are anthropometric data and parameters which can be gathered in routine blood analysis. Prediction models based on these predictors, if accurate, can potentially be used as a biomarker of breast cancer.  **Features Information:**  Quantitative Attributes: Age (years) BMI (kg/m2) Glucose (mg/dL) Insulin (µU/mL) HOMA Leptin (ng/mL) Adiponectin (µg/mL) Resistin (ng/mL) MCP-1(pg/dL)  **Labels:** 1=Healthy controls 2=Patients |

**TASK 1: (5 Marks)**

Implement the decision Tree classifier to detect breast cancer.

**TASK 2: (5 Marks)**

Implement the logistic regression to detect breast cancer.

Note: Standardize the data before training the model