Description of Datasets

- **BostonHousing.csv**

  Associated Task: Regression
  Number of Instances: 506
  Number of attributes: 14
  Attribute Information:
  1. CRIM - per capita crime rate by town
  2. ZN - proportion of residential land zoned for lots over 25,000 sq.ft.
  3. INDUS - proportion of non-retail business acres per town.
  4. CHAS - Charles River dummy variable (1 if tract bounds river; 0 otherwise)
  5. NOX - nitric oxides concentration (parts per 10 million)
  6. RM - average number of rooms per dwelling
  7. AGE - proportion of owner-occupied units built prior to 1940
  8. DIS - weighted distances to five Boston employment centres
  9. RAD - index of accessibility to radial highways
  10. TAX - full-value property-tax rate per $10,000
  11. PTRATIO - pupil-teacher ratio by town
  12. B - 1000(Bk - 0.63)^2 where Bk is the proportion of blacks by town
  13. LSTAT - % lower status of the population
  14. MEDV - Median value of owner-occupied homes in $1000's

- **HeartDisease.csv**

  Associated Task: Classification
  Number of Instances: 303
  Number of attributes: 14
  Attribute Information:
  1. age: age in years
  2. sex: (1 = male; 0 = female)
  3. cp: chest pain type
  4. trestbps: resting blood pressure (in mm Hg on admission to the hospital)
  5. chol: serum cholestoral in mg/dl
  6. fbs: (fasting blood sugar > 120 mg/dl) (1 = true; 0 = false)
  7. restecg: resting electrocardiographic results
  8. thalach: maximum heart rate achieved
  9. exang: exercise induced angina (1 = yes; 0 = no)
  10. oldpeak: ST depression induced by exercise relative to rest
  11. slope: the slope of the peak exercise ST segment
  12. ca: number of major vessels (0-3) colored by flourosopy
  13. tha: l3 = normal; 6 = fixed defect; 7 = reversable defect
  14. target: 1 or 0

- **RedWineQuality.csv**

  Associated Tasks: Classification, Regression

Characteristics: Multivariate
Number of Instances: 1599
Number of attributes: 12
Attribute Information:
    Input variables
        1. fixed acidity
        2. volatile acidity
        3. citric acid
        4. residual sugar
        5. chlorides
        6. free sulfur dioxide
        7. total sulfur dioxide
        8. density
        9. pH
        10. sulphates
        11. alcohol
    Output variable
        12. quality (score between 0 and 10)

- **LiverPatient.csv**

Associated Task: Classification
Number of Instances: 583
Number of attributes: 11
Attribute Information:
        1. age: Age of the patient in years
        2. gender: Patient Gender: Male or Female
        3. TB: Total Bilirubin
        4. DB: Direct Bilirubin
        5. alkphos: Alkaline Phosphotase
        6. sgpt: Alamine Aminotransferase
        7. sgot: Aspartate Aminotransferase
        8. TP: Total Proteins
        9. ALB: Albumin
        10. A_G: Ratio of Albumin and Globulin
        11. class: Predictor Class: 1 if patient has Liver Disease and 2 if they do not

- **WhiteWineQuality.csv**

Associated Tasks: Classification, Regression
Characteristics: Multivariate
Number of Instances: 4898
Number of attributes: 12
Attribute Information:
    Input variables
        1. fixed acidity
        2. volatile acidity

3. citric acid
4. residual sugar
5. chlorides
6. free sulfur dioxide
7. total sulfur dioxide
8. density
9. pH
10. sulphates
11. alcohol

Output variable

12. quality (score between 0 and 10)