Advanced Data Sciene
Assignment 3

Wine Quality Detection

The Assignment's Description:
Based on the given datasets, train and evaluate four different classifiers to identify the best classifier between them for identifying the wine quality.

1. The targeted Datasets:
    - Wine Quality dataset: https://archive.ics.uci.edu/dataset/186/wine+quality
    - Combine the two datasets, red and white wine datasets, into one dataset and do all the required preprocessing steps based on point of view.
2. Train and evaluate the following classifiers:
    - Support Vector Machine (SVM) Classifier,
    - XGBoost Classifier,
    - Random Forest Classifier (with 50 estimators),
    - Multilayer Perceptron (MLP) Classifier
        - with 22 hidden neurons in one layer.
    - use the following evaluation Metrics:
        - Accuracy
        - F1-Score
        - RAC/AUC

The Assignment's Deliverables:

1. Jupyter Notebook, with sufficient comments, description, and explanations. Use Markdown cells for this task. Keep your result in the notebook file.

2. Word document only (docx) reporting your work. The report must include the following:
    a. **Introduction section** that introduces your work and the general purpose of the work.
    b. **Dataset Preparation** section describes the targeted dataset and lists of the applied data preparation steps, with brief justification for each one.
    c. **Classifiers' Description** section that provides brief essential description for each of the targeted classifiers.
    d. **Result and discussion** section represents the evaluation results for each of the targeted classifiers individually, with your interpretation/comments of the result. Then, comparing the gained result of the targeted classifiers visually as possible.
    e. **Conclusion section** that concludes your work and the gained results briefly.

Submission Deadline: 27/12/2024 23:35