reduction and PCA. This dataset is a dictionary where you can access the data and labels directly with the keys data and target.

1. Download the iris dataset directly from scikit-learn.
2. Create a train and a test set.
3. Build a Logistic Regression model to predict the species of the flower and give the main metrics as accuracy, precision, recall, f1-score
4. Apply the PCA to your data and Build a new Logistic Regression model and fit it with the data after PCA.
5. Give the metrics of this new model, and compare them to the metrics of the first model. Compare them based on the metrics obtained with another model using a simple decision tree. What is your analysis of these results.
6. Apply your previous solution (on Iris dataset) on the wine dataset and compare your results without and with PCA.

III-Project-Step 1: This assignment is to be submitted individually on Moodle at the end of the lab. Duration: 1h30

At this first stage of your project, complete the following tasks by providing your notebook (in .pybn format or converted to HTML):

1. Descriptive analysis of your data.
2. Implementation of the necessary pre-processing.
3. Formalisation of the problem.
4. Selection of a baseline model and implementation of the model.