

Lab 6 : interpretability with LIME and SHAP

Exercise 1

For this exercise, we illustrate LIME on the wine quality dataset that can be downloaded at <https://www.kaggle.com/uciml/red-wine-quality-cortez-et-al-2009>. We first need to import the following libraries :

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
from sklearn.model_selection import train_test_split
from sklearn import preprocessing
from sklearn.ensemble import RandomForestRegressor
import lime
import lime.lime_tabular
```

1. Importation and preprocessing
 - (a) Load the data
 - (b) The target variable is 'quality'. Separate label from features
2. Learning
 - (a) Split the data into train and test data
 - (b) Build the model with the random forest regression algorithm
3. Interpretation with LIME
 - (a) Use the function `LimeTabularExplainer` to define the LIME explainer
 - (b) Interpret now the first record using the explainer using the function `explain_instance`
 - (c) Display the coefficients as a list
4. Interpretation with SHAP
 - (a) Use the function `TreeExplainer` to define the SHAP explainer on the train
 - (b) Display the shap values using `shap_values`. Use `summary_plot` to visualize