Model Interpretation

For Assignment 4, we will train and interpret model on wine-quality dataset. You can access the data from following link. There are two csv files available on the link, but **you only need to work on white-wine dataset.** Treat this dataset as a regression problem where 1 is poor and 10 is excellent quality. Use R-squared metrics for model evaluation.

https://archive.ics.uci.edu/ml/machine-learning-databases/wine-quality/

1. Train a Random Forest Regressor for the dataset. Find the best model based on R-squared value using RandomizedSearchCV. [10 Marks]

2. Use the best model from question 1 for model interpretation and rank the features based on drop feature importance. [15 Marks]

3. Use the best model from question 1 for model interpretation and rank the features based on permutation importance. [15 Marks]

4. Use the best model from question 1 for model interpretation and rank the features based on SHAP algorithm. Install SHAP using pip. [20 Marks]

5. Visualize partial dependence plot for each feature in the dataset using Sklearn. [10 Marks]

6. Visualize ICE plot for each feature using following library. http://austinrochford.github.io/PyCEbox/ [20 Marks]

7. Analyze outputs from each technique and comment that which technique you found most useful and why. [10 Marks]

Please save your notebook with all the images and comments before submitting.