

# CT5132/CT5148 Lab Week 5

James McDermott

1. Visit [https://scikit-learn.org/stable/datasets/toy\\_dataset.html](https://scikit-learn.org/stable/datasets/toy_dataset.html) to see some of the famous “toy” datasets included with Scikit-Learn.
2. Load the Wine Recognition dataset using `load_wine(return_X_y=True)`.
3. Take a look at the dataset. What type and what shape does it have? How many classes are there?
4. Create a train-test split. Set the random seed so that your code is reproducible.
5. Carry out the Import, Instantiate, Train, Evaluate, and Predict workflow using the `LogisticRegression` classifier.
6. Suppose we receive a wine with the following data. What class is it likely from?  
`array([ 12.29, 3.17, 2.21, 18., 88., 2.85, 2.99, 0.45, 2.81, 2.3, 1.42, 2.83, 406.])`
7. Carry out the same procedure using `LogisticRegression` and several `RandomForestClassifiers`, with different parameters. For each, produce a confusion matrix also. But this time, put all the Train and Evaluate steps in a loop, so you don’t have to copy/paste your code, and so that all results are evaluated in exactly the same way.
8. Create a table showing the  $R^2$  value for each classifier on the test set (again, this should be automated, using a loop, not done by hand or copy-pasting).