

Machine Learning

Laboratory on Scikit-learn

Giovanni De Toni

giovanni.detoni@unitn.it

5th December 2022

Preliminaries

Make sure you are using **Python 3** for the following steps

```
> pip install numpy scipy matplotlib sklearn pandas
```

```
> pip install jupyter
```

Then, open the terminal in the folder containing the extracted files and run

```
> jupyter notebook
```

Take-home Exercise

As a self-assessment exercise, you can try to solve a classification task using **Scikit-Learn**. Choose a dataset, train a classifier by performing some model selection (e.g., hyperparameters tuning) and then evaluate the model on the test set.

The following website has many datasets you can use:

<https://archive.ics.uci.edu/ml/datasets.php>

Take-home Exercise

Potential datasets you can use are:

- **Spambase** (<https://archive.ics.uci.edu/ml/datasets/spambase>)
- **OCR Recognition**
(<https://archive.ics.uci.edu/ml/datasets/optical+recognition+of+handwritten+digits>)
- **Abalone** (<https://archive.ics.uci.edu/ml/datasets/abalone>)
- and many others...