Machine learning libraries

Francesco Pugliese, PhD

neural1977@gmail.com

Table of contents

✓ Scikit-Learn

Scikit-Learn

- ✓ Scikit-Learn is python’s **core machine learning package** that has most of the necessary modules to support a basic machine learning project.

- ✓ The library provides a **unified API** for practitioners to ease the use of machine learning algorithms with only writing a few lines to accomplish the predictive or classification task.

- ✓ The package is written heavily in **python**, and it incorporates C++ libraries like LibSVM and LibLinear for support vector machines and generalized linear model implementation.

- ✓ The package depends on **Pandas** (mainly for the dataframe processes), **numpy** (for the ndarray construct) and **scipy** (for sparse matrices).

- ✓ Scikit-learn does one thing and only one thing very well, and that is implementing essential machine learning algorithms.

Where did it come from ?

- ✓ Scikit-learn was initially developed by **David Cournapeau** as a **Google summer of code project** in **2007**.
- ✓ Later **Matthieu Brucher** joined the project and started to use it as apart of his thesis work.
- ✓ In **2010** INRIA got involved and the **first public release** (v0.1 beta) was published in late January 2010.
- ✓ The project now has more than **30 active contributors** and has had paid sponsorship from **INRIA**, **Google**, **Tinyclues** and the Python Software Foundation.

Prerequisites for scikit-learn

- ✓ The library is built upon the SciPy (Scientific Python) that must be installed before you can use scikit-learn.

- ✓ This stack that includes:
  - ✓ **NumPy**
  - ✓ **SciPy**: Fundamental library for scientific computing
  - ✓ **Matplotlib**
  - ✓ **IPython**: Enhanced interactive console
  - ✓ **Sympy**: Symbolic mathematics
  - ✓ **Pandas**

Scikit-Learn flow

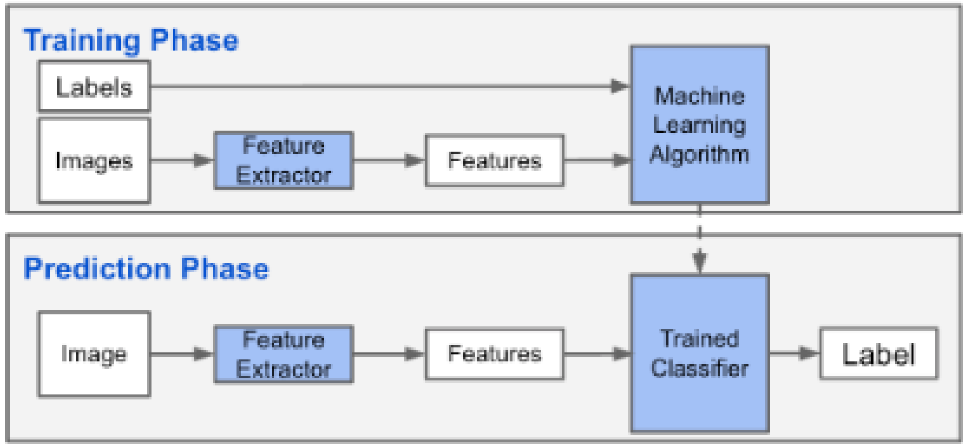
Most of the Scikit-Learn modules follow the same steps.

- ✓ Loading the data
- ✓ Pre-processing the data
- ✓ Train & Test data split
- ✓ Creating your model using supervised & unsupervised learning
- ✓ Fit the model with train set
- ✓ Predicting it with test set and finally
- ✓ Evaluate the model's performance.

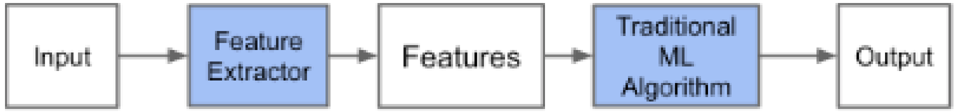
  Note: Scikit-Learn does not provide any **GPU** support

- ✓ More details about scikit-learn methods for performing above steps can be found in this colab notebook

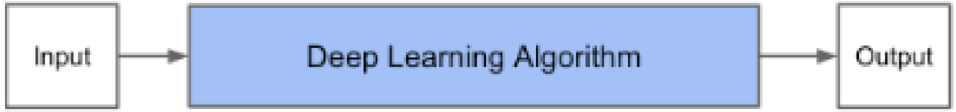
Scikit-Flow



Machine Learning Phases

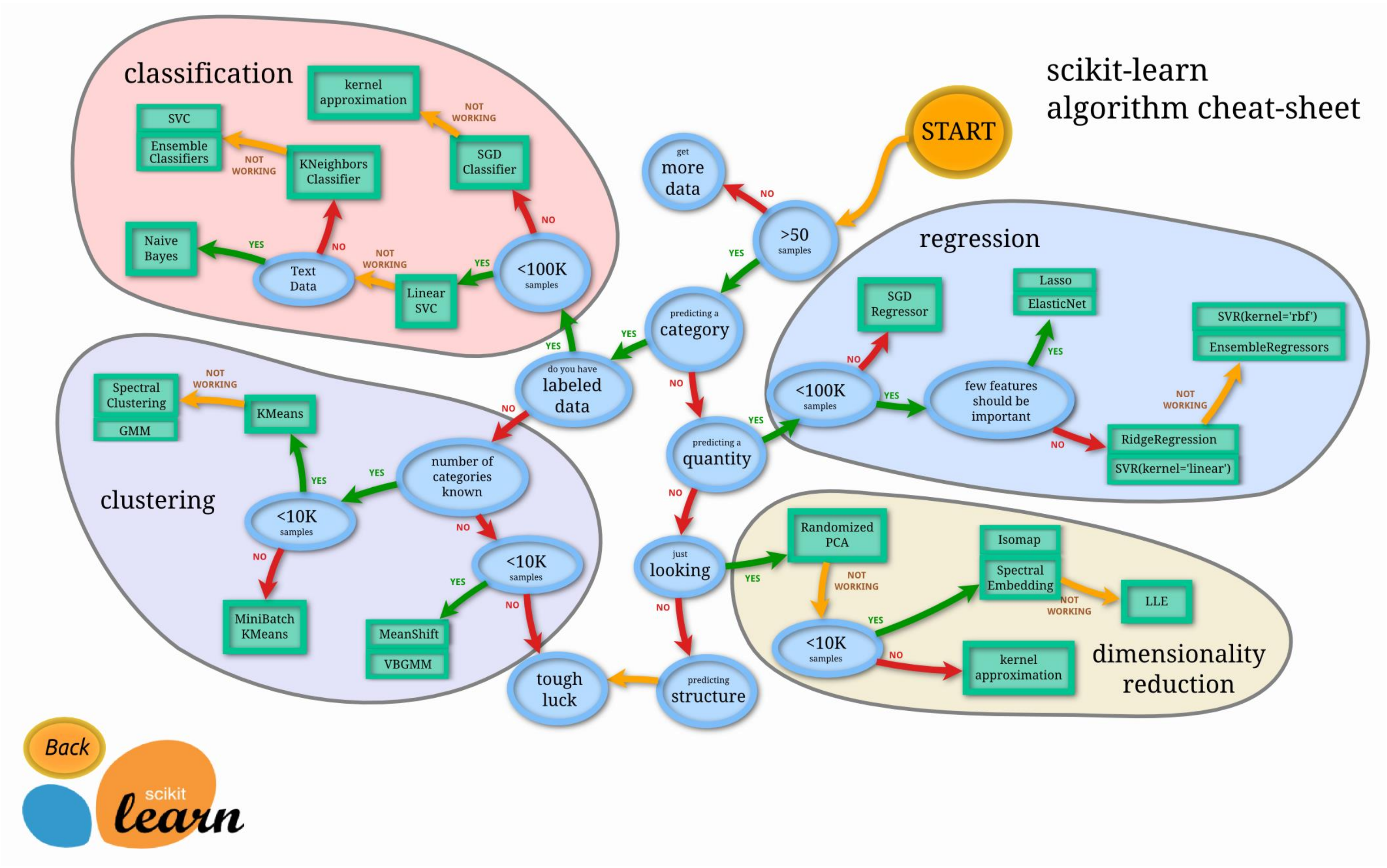


Traditional Machine Learning Flow



Deep Learning Flow

Scikit-learn algorithms cheat sheet



Francesco **Pugliese**