Installing and running scikit-learn

Peter-Tibor Zavaczki

march 7, 2018

Chapter 1

Scikit-learn

1.1 Tool Purpose

Scikit-learn is a machine learning library for Python, which features various classification, regression and clustering algorithms, and is designed to interoperate with the Python numerical and scientific libraries NumPy and SciPy.

1.2 Installing scikit-learn

Prior to using scikit-learn, Python (>= 2.7 or >= 3.3) has to be installed along with the NumPy (>= 1.8.2) and SciPy (>= 0.13.3) libraries.

1.2.1 Installing steps

Some versions of Ubuntu come installed with Python 2.7.12 and Python 3.5.2, so for this installation we will consider that and install scikit-learn for Python 3.5.2. To ease installing packages for Python we will use pip. To install pip, we need the command *sudo apt install python3-pip*. Please note that we have a 3 after python to signal that we will install pip for Python 3.x. After the previous step we install NumPy and SciPy by using the command *sudo pip3 install numpy scipy*. This will download the libraries' latest version and automatically install them. As a final step, we use the command *sudo pip3 install scikit-learn* to install scikit-learn.

Chapter 2

Studided example

2.1 Recognizing hand-written digits

The studied example is **Recognizing hand-written digits** by **Gael Varoquaux**, a handwritten digit classificator by machine learning. It can recognize the 0-9 handwritten digits and convert them to digital characters.

2.2 How to run the example(s)

To run the given example, you need to have matplotlib, installed with *sudo pip3 install matplotlib* and python3-tk, installed with *sudo apt-get install python3-tk*. Then just use the command $python3\ ./plot\_digits\_classification.py$ from the folder of origin to run the example.