

## EECS 510 Social Media Mining Lecture 5: Tutorial Data Mining with Python Software Installation Resources

On Lecture 5 we are going to go through a tutorial on simple Machine Learning methods and data mining tools with interactive tools as iPython. This document provides certain version requirements and useful links to the installation of relevant tools to be expected in class. Although it is not mandated, to better follow the tutorial lecture, students are encouraged to install the relevant packages.

### Overview

The tutorial, which includes simple demonstration on data mining tools such as PCA, K-means Clustering, Linear Regression and Logistic Regression, is going to be given in Python, with the interactive tool IPython Notebook and required libraries scikit-learn, matplotlib and numpy.

Jupyter (newer version of iPython) -

<http://jupyter.readthedocs.org/en/latest/install.html>

Jupyter requires: Python 3.3 or greater, or Python 2.7.

Scikit-learn (latest version is 0.17.0) - <http://scikit-learn.org/stable/install.html>

Scikit-learn requires:

Python ( $\geq 2.6$  or  $\geq 3.3$ ),

NumPy ( $\geq 1.6.1$ ),

SciPy ( $\geq 0.9$ ).

matplotlib - <http://matplotlib.org/users/installing.html>