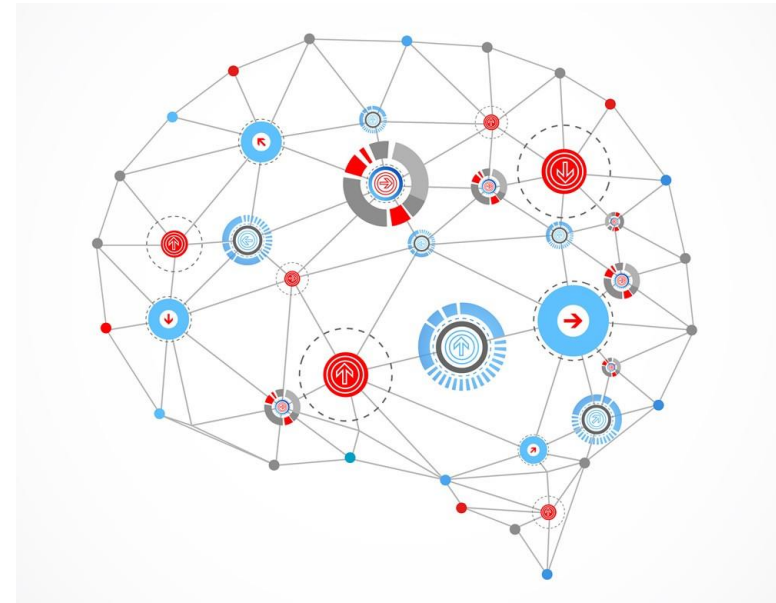# 智慧化企業整合
Intelligent Integration of Enterprise

Setup, Python Numpy Matplotlib Tutorial

助教:陳可馨

Outline

- Python, Anaconda Introduction
- Python
- Numpy
- Pandas
- Matplotlib
- Homework1

• Introduction

✓ Python

❑ Python can :

1. be used on a server to create web applications.

2. be used alongside software to create workflows.

3. connect to database systems. It can also read and modify files.

4. be used to handle big data and perform complex mathematics.

5. be used for rapid prototyping, or for production-ready software development.

❑ Why Python?

1. works on different platforms (Windows, Mac, Linux, Raspberry Pi, etc).

2. has a simple syntax similar to the English language.

3. has syntax that allows developers to write programs with fewer lines than some other programming languages.

4. runs on an interpreter system, meaning that code can be executed as soon as it is written.

Python, Anaconda Introduction

- Introduction & Installation

✓ Anaconda: Python的懶人包

✓ https://www.anaconda.com/distribution/ (Python 3.7 version is recc

The open-source Anaconda Distribution is the easiest way to perform Python/R data science and machine learning on Linux, Windows, and Mac OS X. With over 11 million users worldwide, it is the industry standard for developing, testing, and training on a single machine, enabling *individual data scientists* to:

- Quickly download 1,500+ Python/R data science packages
- Manage libraries, dependencies, and environments with Conda
- Develop and train machine learning and deep learning models with scikit-learn, TensorFlow, and Theano
- Analyze data with scalability and performance with Dask, NumPy, pandas, and Numba
- Visualize results with Matplotlib, Bokeh, Datashader, and Holoviews
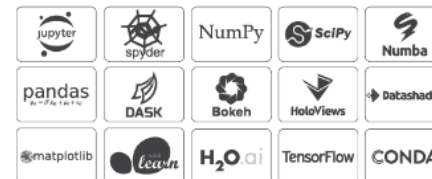
🪟 Windows | 🍎 macOS | 🐧 Linux

Anaconda 2019.03 for Windows Installer

Python 3.7 version
Download
64-Bit Graphical Installer (662 MB)
32-Bit Graphical Installer (546 MB)

Python 2.7 version
Download
64-Bit Graphical Installer (587 MB)
32-Bit Graphical Installer (493 MB)

Python

- # Basic Data Types
  - ✓ Integers and floats
  - ✓ Booleans
  - ✓ Strings

- # Containers
  - ✓ List
  - ✓ Loops(ex: for loop, while loop…)
  - ✓ Dictionaries
  - ✓ Sets
  - ✓ Tuples

- # Functions
- # Classes

x + y	sum of x and y
x - y	difference of x and y
x * y	product of x and y
x / y	quotient of x and y
x // y	floored quotient of x and y
x % y	remainder of x / y
-x	x negated
+x	x unchanged
abs(x)	absolute value or magnitude of x
int(x)	x converted to integer
float(x)	x converted to floating point
complex(re, im)	a complex number with real part re, imaginary part im. im defaults to zero.
c.conjugate()	conjugate of the complex number c
divmod(x, y)	the pair (x // y, x % y)
pow(x, y)	x to the power y
x ** y	x to the power y

Numpy

Numpy is the core library for scientific computing in Python. It provides a high-performance multidimensional array object, and tools for working with these arrays.

Import numpy as np

- Arrays

  - ✓ np.array
  - ✓ .dtype
  - ✓ .reshape
  - ✓ .shape
  - ✓ .T
  - ✓ np.dot
  - ✓ np.arange
  - ✓ np.zeros
  - ✓ np.ones

  - ✓ np.full
  - ✓ np.random.random
  - ✓ np.sum
  - ✓ np. concatenate
  - ✓ np.nan
  - ✓ np.vstack
  - ✓ np.hstack
  - ✓ np.linspace
  - ………

- Array math

  - ✓ np.add
  - ✓ np.subtract
  - ✓ np.multiply
  - ✓ np.divide
  - ✓ np.sqrt()
  - ✓ np.dot
  - ✓ np.sum
  - ✓ np.mean
  - ....

Pandas

import pandas as pd
- Series
- Dataframe

✓ .loc   ✓.merge

✓ .iloc   ....

✓ .ix

✓ .dropna()

✓ .fillna()

✓ .isnull()

✓ .concat

✓ .append

✓ read.csv

Matplotlib

import matplotlib.pyplot as plt

• Plotting

✓ .figure
✓ .plot
✓ .show()
✓ .xlabel
✓ .ylabel
✓ .title
✓ .legend
✓ .subplot

✓.xlim , .ylim
✓.scatter
✓.bar
………..…

Reference

- Python:https://docs.python.org/3/tutorial/
- Numpy:https://docs.scipy.org/doc/
- Pandas: https://pandas.pydata.org/pandas-docs/stable/
- Matplotlib:https://matplotlib.org/3.1.1/tutorials/index.html
- Others:http://www.runoob.com/python3/python3-tutorial.html

Homework

- Please generate matrices $M1$ & $M2$ with random Gaussian distribution of zero mean and standard deviation of 1 & 0.8 respectively.
- $M1 \in \mathbb{R}^{n \times m}$ & $M2 \in \mathbb{R}^{m \times m}$ for $n=123$, $m=345$
- Turn in your work with the format of .ipynb , and please write some brief comments in your ipynb to illustrate your results.
- File name: hw5_Your Chinese Name
- You are asked to:
1. Calculate $M1+M1$, $M1M1T$ and $M1M2$
2. Generate a vector y with m entries and solve $M2x=y$ (Hint: inv(M2))

Note: Please make sure you have saved your result (just press Ctrl+s) before you turn in your homework file.