Practice 01

Seongok Ryu
KAIST Chemistry

Contents

- Why do we use python?
- Python packages for data science and machine learning
- Virtual environment with Anaconda
- Python coding convention
- Linear regression with TensorFlow

Why do we use python?

“Life is too short, You need Python.”


```java
public class HelloWorld {
    public static void main(String[] args) {
        System.out.println("Hello, world!");
    }
}
```

Python 2.x

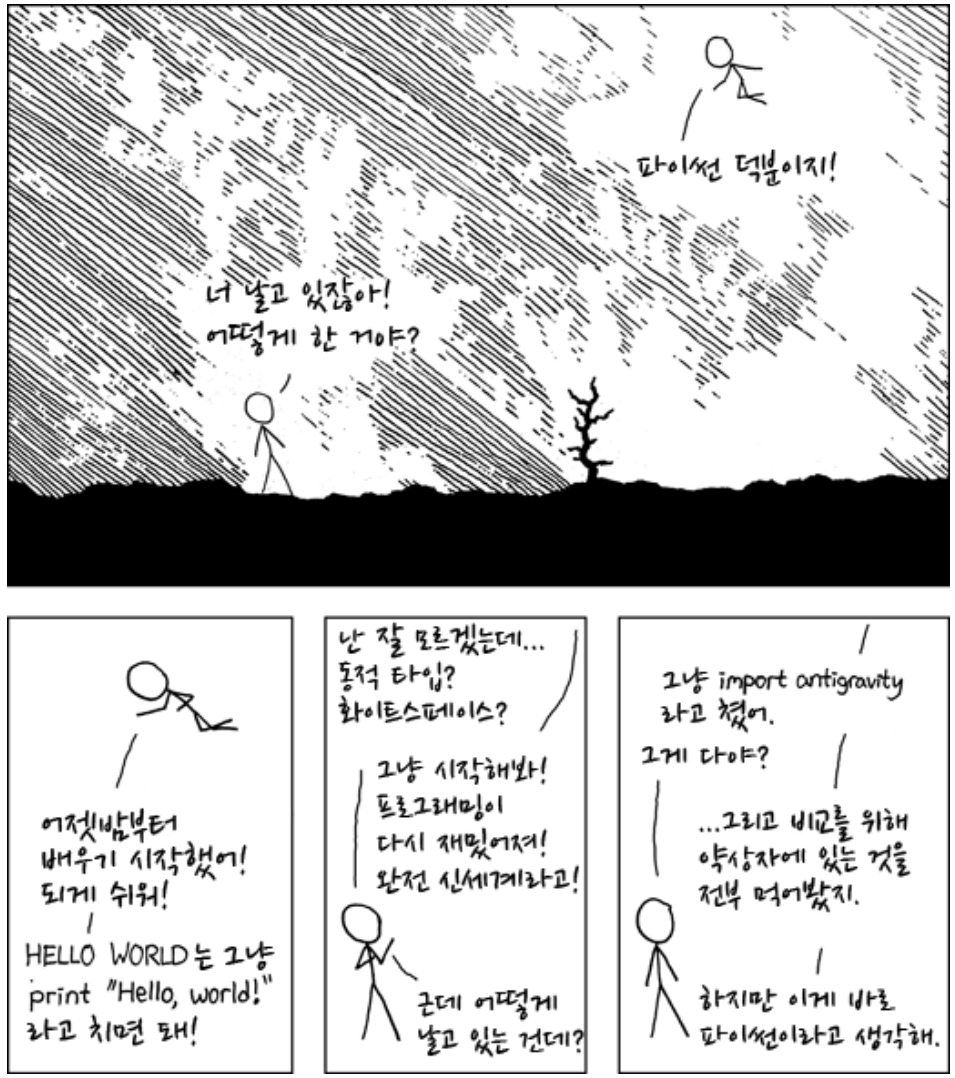
```
1.  print 'hello world!'
```

Python 2.6 ~ 3.x

```
1.  print('hello world!')
```

Why do we use python?

“Life is too short, You need Python.”



Python packages for data science and machine learning

- Scientific works – numpy, scipy, cvxopt
- Statistics & Visualization – pandas, matplotlib, seaborn
- Web programming – Flask, django
- Machine learning – scikit-learn, tensorflow, keras, pytorch

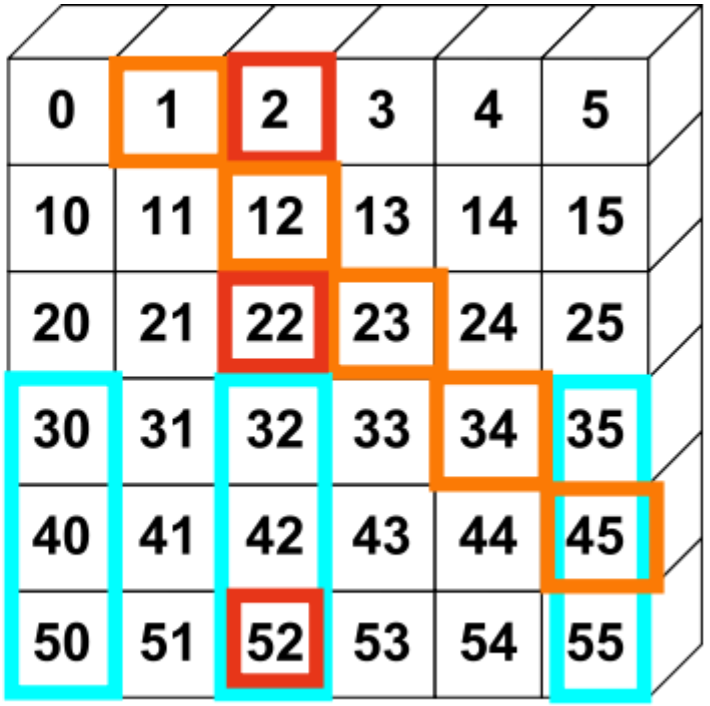
Python packages for data science and machine learning

- **Python libraries**

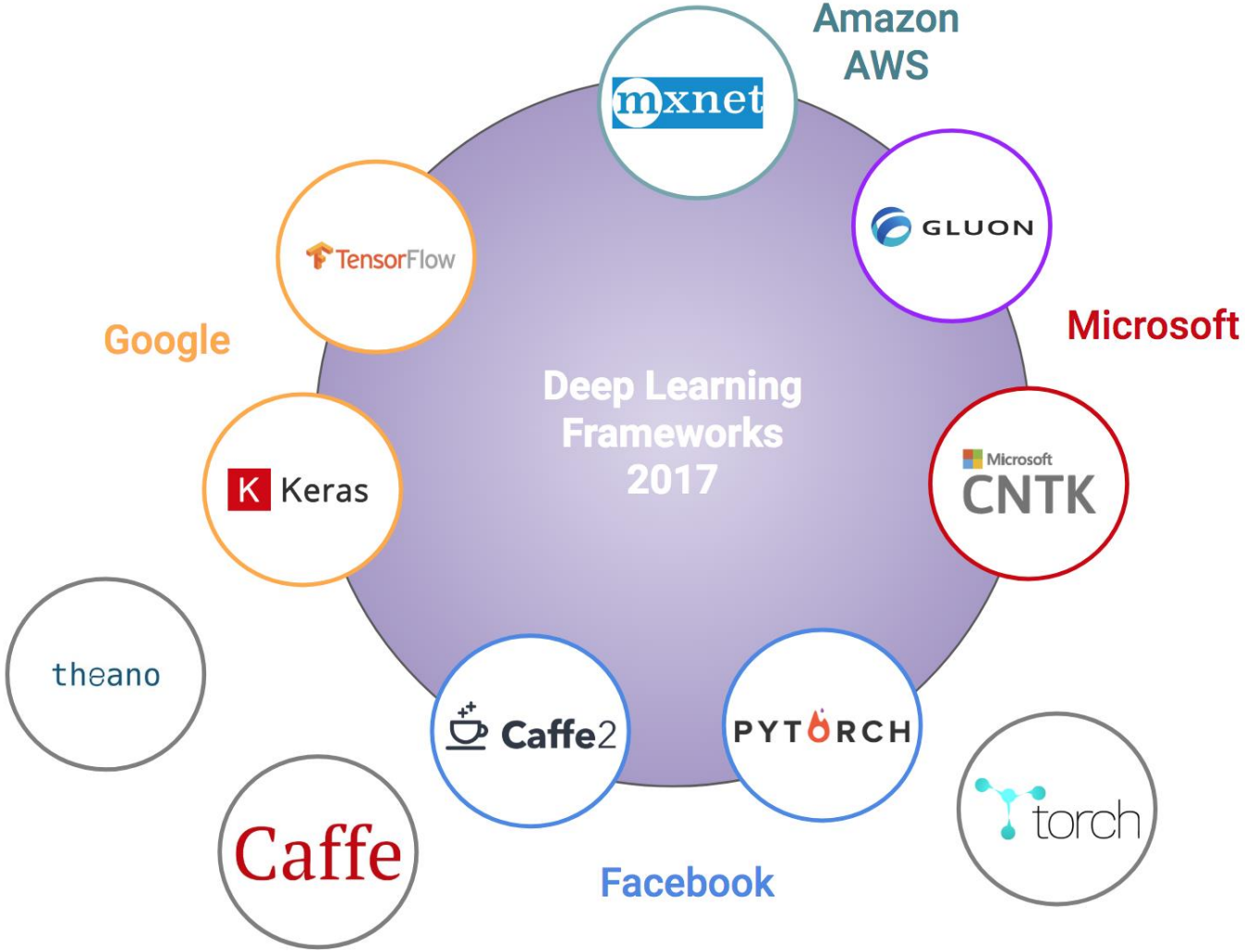

```
>>> a[(0,1,2,3,4),(1,2,3,4,5)]
array([ 1, 12, 23, 34, 45])

>>> a[3:,[0, 2, 5]]
array([[30, 32, 35],
       [40, 42, 45]])
       [50, 52, 55]])

>>> mask = array([1,0,1,0,0,1],
                 dtype=bool)
>>> a[mask,2]
array([2,22,52])
```



Python packages for data science and machine learning



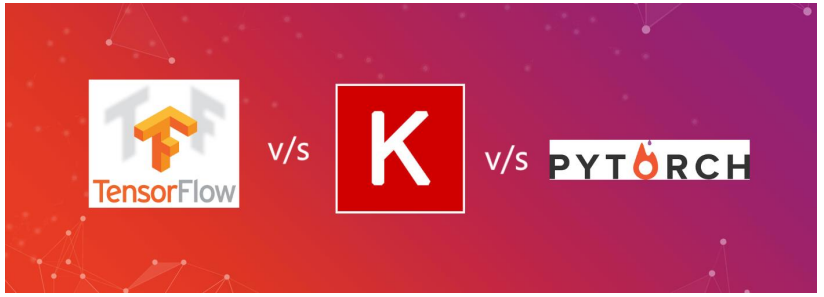
https://towardsdatascience.com/battle-of-the-deep-learning-frameworks-part-i-cff0e3841750

Python packages for data science and machine learning

- **Tensorflow → This practice will mainly use tensorflow for practices.**
  - Google에서 만들고 제공하는 deep learning framework
  - 가장 많은 user 들이 사용 → many practice codes on GitHub
  - Static computational graph

- **Keras**
  - Tensorflow와 theano를 back-end로 하는 framework
  - Very very easy to use
  - Static computational graph

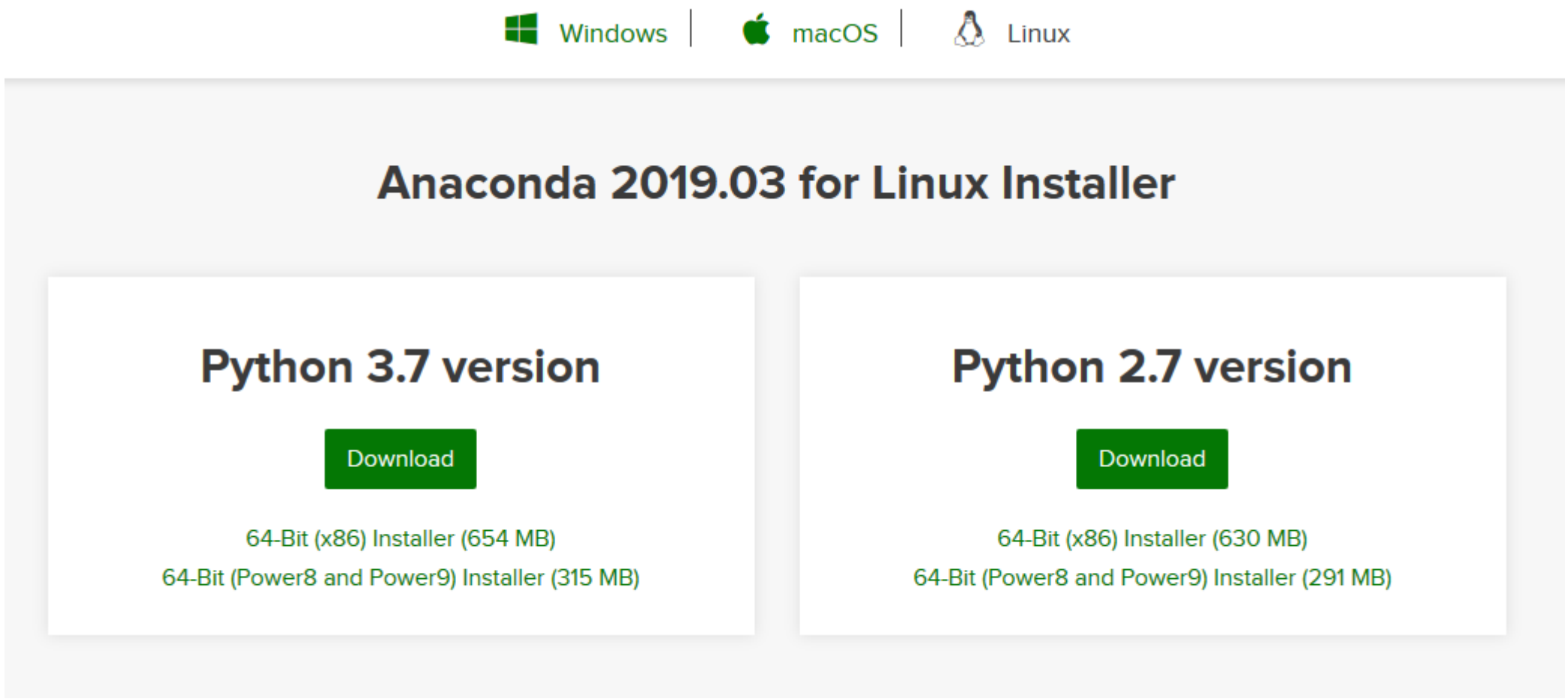
- **PyTorch**
  - Facebook에서 만든 Torch(C++)를 Python으로 wrapping
  - Dynamic computational graph



http://codeinpython.com/tutorials/deep-learning-tensorflow-keras-pytorch/

Virtual environment with Anaconda

1. Install Anaconda https://www.anaconda.com/distribution/



Virtual environment with Anaconda

2. Create your own virtual environment

```
[wykgroup@horus0 ~]$ conda create -n python_seongok
Collecting package metadata (current_repodata.json): done
Solving environment: done

## Package Plan ##

  environment location: /home/wykgroup/appl/anaconda3/envs/python_seongok



Proceed ([y]/n)? y

Preparing transaction: done
Verifying transaction: done
Executing transaction: done
#
# To activate this environment, use
#
#     $ conda activate python_seongok
#
# To deactivate an active environment, use
#
#     $ conda deactivate
```

Virtual environment with Anaconda

3. Activate your virtual environment

```
[wykgroup@horus0 ~]$ source activate python_seongok
(python_seongok) [wykgroup@horus0 ~]$ conda list
# packages in environment at /home/wykgroup/appl/anaconda3/envs/python_seongok:
#
# Name                    Version                   Build  Channel
(python_seongok) [wykgroup@horus0 ~]$
```

* Deactivation of your virtual environment

```
(python_seongok) [wykgroup@horus0 ~]$ source deactivate python_seongok
DeprecationWarning: 'source deactivate' is deprecated. Use 'conda deactivate'.
```

Virtual environment with Anaconda

4. Install python packages in your virtual environment

```
(python_seongok) [wykgroup@horus0 ~]$ conda install numpy
Collecting package metadata (current_repodata.json): done
Solving environment: done

## Package Plan ##

  environment location: /home/wykgroup/appl/anaconda3/envs/python_seongok

  added / updated specs:
    - numpy


The following packages will be downloaded:

    package                    |            build
    ---------------------------|-----------------
    blas-1.0                   |              mkl           6 KB
```

- numpy, scikit-learn
- tensorflow (tensorflow-gpu for GPU utilization)
- rdkit (for chemists), see https://anaconda.org/rdkit/rdkit

Virtual environment with Anaconda

5. Check whether packages are successfully installed

```
(python_seongok) [wykgroup@horus0 ~]$ python
Python 3.7.3 (default, Mar 27 2019, 22:11:17)
[GCC 7.3.0] :: Anaconda, Inc. on linux
Type "help", "copyright", "credits" or "license" for more information.
>>> import numpy as np
>>> import tensorflow as tf
>>> import keras
Traceback (most recent call last):
  File "<stdin>", line 1, in <module>
ModuleNotFoundError: No module named 'keras'
>>>
```

* I didn’t install ‘keras’ in the virtual environment ‘python_seongok’.

Python coding convention

See practice codes in
https://github.com/SeongokRyu/ACE-Team-DLstudy/tree/master/practices/week01

Linear regression with TensorFlow

See practice codes in
https://github.com/SeongokRyu/ACE-Team-DLstudy/tree/master/practices/week01