

Computing for Medicine: Phase 3, Seminar 3 Project

Rabiya Noori

Based on slides by Jennifer Campbell

Seminar 3 Project: Identifying Alzheimer's disease from picture descriptions

- The project is posted handout is posted:
 - https://c4m-uoft.github.io/seminars/Frank_Rudzicz/Seminar3Project.pdf
- Packages used:
 - numpy (included in Anaconda distribution)
 - nltk
 - scipy
 - scikit-learn

Installing Packages

1. *Activate your virtual environment*

Windows: Type **Anaconda** in the search box and choose **Anaconda Prompt** from the list.

Run:

```
> conda activate C4M
```

Mac: Open **terminal** and run:

```
> source activate C4M
```

Installing Packages

2. Install packages

(Windows & Mac) Run the following commands. Wait for each installation to finish before running the next command. Answer 'y' if prompted.

```
> conda install -c anaconda nltk
```

```
> conda install -c anaconda scipy
```

```
> conda install -c anaconda scikit-learn
```

```
> conda install -c anaconda numpy
```

Now you can open jupyter lab as per the instructions on our website and start using these modules.

More on modules used

- nltk (Natural Language Toolkit) – <http://www.nltk.org/api/nltk.html>
- csv (Comma Separated Values files) – <https://docs.python.org/3/library/csv.html>
- math (Mathematical functions) – <https://docs.python.org/3/library/math.html>
- numpy (Numerical Python) – <https://docs.scipy.org/doc/numpy/reference/routines.html>
- sklearn / scikit-learn (Machine Learning in Python) - <https://scikit-learn.org/stable/>

- Click on the links above for documentation for each module/package. You can also use help().

NumPy

- A scientific computing package for Python.
- For this project, you'll use NumPy's N-dimensional array.
- NumPy's 2D array vs Python's nested lists
 - NumPy's array may contain only elements of the same type, whereas Python's lists may contain different types.
 - NumPy's arrays are more efficient and take less space.
 - NumPy supports a variety of array operations.

NumPy 2D array demo

```
>>> import numpy as np
```

```
>>> my_array = np.array([[1, 2, 3], [4, 5, 6]])
```

```
>>> my_array.shape
```

```
(2, 3)
```

```
>>> my_array.size
```

```
6
```

```
>>> my_array.sum()
```

```
21
```

```
>>> my_array.min()
```

```
1
```

```
>>> my_array.max()
```

```
6
```

```
>>> my_array.mean()
```

```
3.5
```

```
>>> my_array.var()
```

```
2.9166666666666665
```



```
>>> my_array.std()
```

```
1.707825127659933
```

Upcoming Seminar

Seminar 4: Chris J. McIntosh

Date: Tuesday, February 12, 2019; 4-6pm

Location: DSC Innovation Lab, Gerstein Library

Topic: Medical Image Analysis

Profile: https://www.researchgate.net/profile/Chris_Mcintosh