

DATA SCIENCE ROADMAP

2021



AI Expert in 2022

Required for any path

Papers With Code

GIT - Version Control

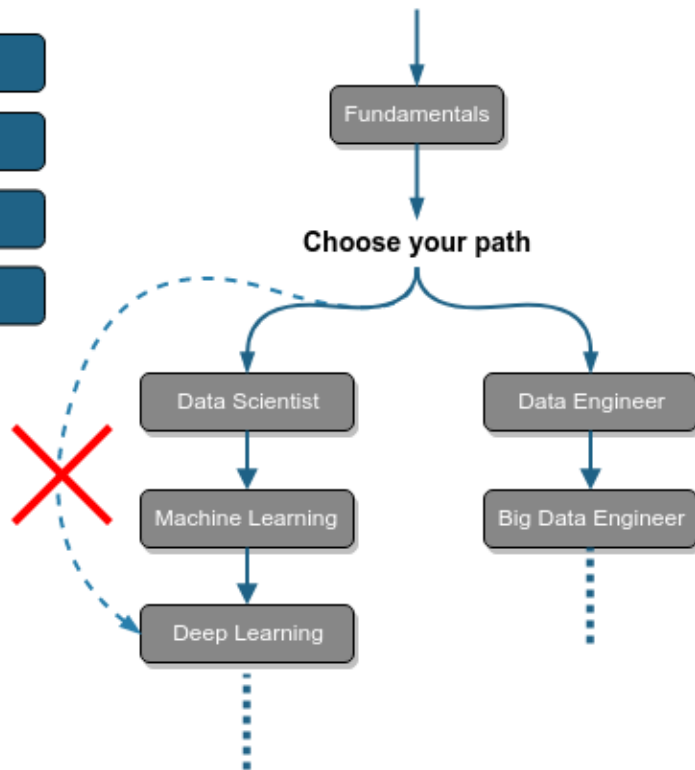
Semantic Versioning

Keep a Changelog

Legend

Personal Recommendation!

Available Options



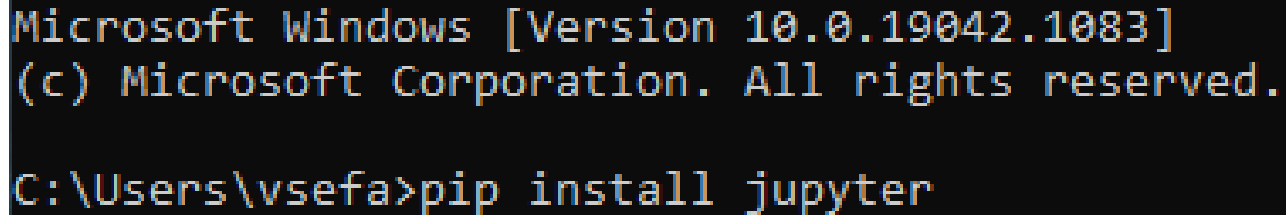
Intro to Jupyter Notebook

01

Instalasi

1. pip

```
pip install jupyter
```

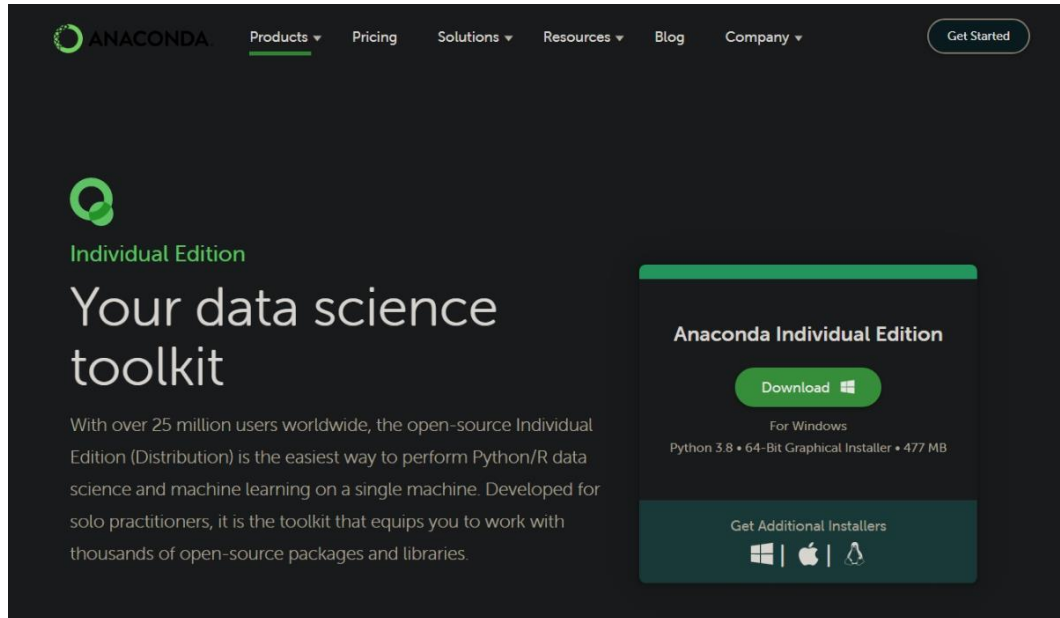


```
Microsoft Windows [Version 10.0.19042.1083]
(c) Microsoft Corporation. All rights reserved.

C:\Users\vsefa>pip install jupyter
```

2. Anaconda


<https://www.anaconda.com/products/individual>



The screenshot shows the Anaconda Individual Edition product page. The navigation bar includes the Anaconda logo, a 'Products' dropdown menu (highlighted with a red underline), 'Pricing', 'Solutions', 'Resources', 'Blog', 'Company', and a 'Get Started' button. The main content area features the Anaconda logo, the text 'Individual Edition', and the headline 'Your data science toolkit'. A paragraph describes the product as an open-source toolkit for Python/R data science and machine learning. On the right, a product card displays 'Anaconda Individual Edition', a 'Download' button with a Windows icon, 'For Windows', and 'Python 3.8 • 64-Bit Graphical Installer • 477 MB'. At the bottom of the card, there is a 'Get Additional Installers' section with icons for Windows, macOS, and Linux.

ANACONDA

Products ▾ Pricing Solutions ▾ Resources ▾ Blog Company ▾ [Get Started](#)

 Individual Edition

Your data science toolkit




With over 25 million users worldwide, the open-source Individual Edition (Distribution) is the easiest way to perform Python/R data science and machine learning on a single machine. Developed for solo practitioners, it is the toolkit that equips you to work with thousands of open-source packages and libraries.

Anaconda Individual Edition

[Download !\[\]\(f94f473f740247fc62afc0e0e5025277_img.jpg\)](#)

For Windows
Python 3.8 • 64-Bit Graphical Installer • 477 MB

Get Additional Installers

 |  | 

2. Anaconda

Apa aja yang termasuk di Anaconda

- ❑ Package dan environment system, conda
- ❑ Library machine learning: TensorFlow, scikit-learn, dll
- ❑ Library data science: pandas, NumPy, dll.
- ❑ Library visualisasi: matplotlib, seaborn, dll.
- ❑ Jupyter Notebook

Package Name	Access	Summary	Updated
xlwings	public	Interact with Excel from Python and Vice versa	2021-07-09
spyder-kernels	public	Jupyter kernels for Spyder's console	2021-07-09
pillow	public	Pillow is the friendly PIL fork by Alex Clark and Contributors	2021-07-09
bokeh	public	Statistical and novel interactive HTML plots for Python	2021-07-09
intervals	public	Python tools for handling intervals (ranges of comparable objects).	2021-07-09
iminuit	public	Interactive Minimization Tools based on MINUIT	2021-07-09
dropbox	public	Official Dropbox API Client	2021-07-09
glib	public	Provides core application building blocks for libraries and applications written in C.	2021-07-09
humanize	public	Python humanize utilities	2021-07-09
fonttools	public	fontTools is a library for manipulating fonts, written in Python.	2021-07-09
google-auth	public	Google authentication library for Python	2021-07-09
flask-jwt-extended	public	A Flask JWT extension	2021-07-09
elasticsearch	public	Python client for Elasticsearch	2021-07-09
boto3	public	Amazon Web Services SDK for Python	2021-07-09
datadog	public	The Datadog Python library	2021-07-09
aws-sam-translator	public	AWS Serverless Application Model (AWS SAM) prescribes rules for expressing Serverless applications on AWS.	2021-07-09
asgiref	public	ASGI in-memory channel layer	2021-07-09
apispec	public	A pluggable API specification generator	2021-07-09
toxinfo	public	The GNU Documentation System.	2021-07-08
orc	public	C++ libraries for Apache ORC	2021-07-08
regex	public	Alternative regular expression module, to replace re	2021-07-08
pyodbc	public	DB API Module for ODBC	2021-07-08
pylint	public	python code static checker	2021-07-08
zippp	public	A pathlib-compatible Zipfile object wrapper	2021-07-08
xlsxwriter	public	A Python module for creating Excel XLSX files	2021-07-08
pytest-asyncio	public	Pytest plugin with advanced	2021-07-08

3. Miniconda

<https://docs.conda.io/en/latest/miniconda.html>



The screenshot shows the documentation page for Miniconda. The left sidebar contains navigation links for Conda, Conda-build, Miniconda, and various installers. The main content area is titled 'Miniconda' and includes a description of the installer and a table of Windows installers. A red circle highlights the 'Name' column of the table.

Docs » Miniconda [Edit on GitHub](#)

Miniconda

Miniconda is a free minimal installer for conda. It is a small, bootstrap version of Anaconda that includes only conda, Python, the packages they depend on, and a small number of other useful packages, including pip, zlib and a few others. Use the `conda install command` to install 720+ additional conda packages from the Anaconda repository.

See if Miniconda is right for you.

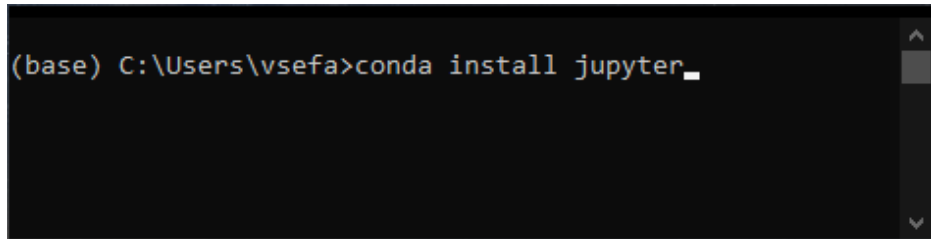
Windows installers

Windows			
Python version	Name	Size	SHA256 hash
Python 3.9	Miniconda3 Windows 64-bit	57.7 MiB	<code>c3a43d6bc4c4fa92454dbfa63eccb859a0456875df602b31ae71b9e0c3fec2b8</code>
	Miniconda3 Windows 32-bit	54.9 MiB	<code>5045fb9dc4405dbba21054262b7d184b0a61a8739c1856038cc08258f233ad646</code>
Python 3.8	Miniconda3 Windows 64-bit	57.0 MiB	<code>4fa22ba0497bab5b6608cb8843545372a99f5331c81200999ae1d083f627c61</code>
	Miniconda3 Windows 32-bit	54.2 MiB	<code>9c2ef76bae97246c95c206733ca30fd1feb0a4b3f90a2a511fea081ce70bc661</code>
Python 2.7	Miniconda2 Windows 64-bit	54.1 MiB	<code>6973025404832944e874bf02bdab4c4594980ceed4707bd51baa8fbd0a4bf332c</code>
	Miniconda2 Windows 32-bit	47.7 MiB	<code>c8049d26f2b0b954057bc04e99ad72d1ff613faab0218e64e641504417b2617b</code>

3. Miniconda

- ❑ Versi minimal Anaconda
- ❑ Hanya termasuk conda, Python, dan beberapa package kecil
- ❑ Install library dengan command: `conda install <nama-library>`

```
(env) conda install jupyter
```



```
(base) C:\Users\vsefa>conda install jupyter_
```

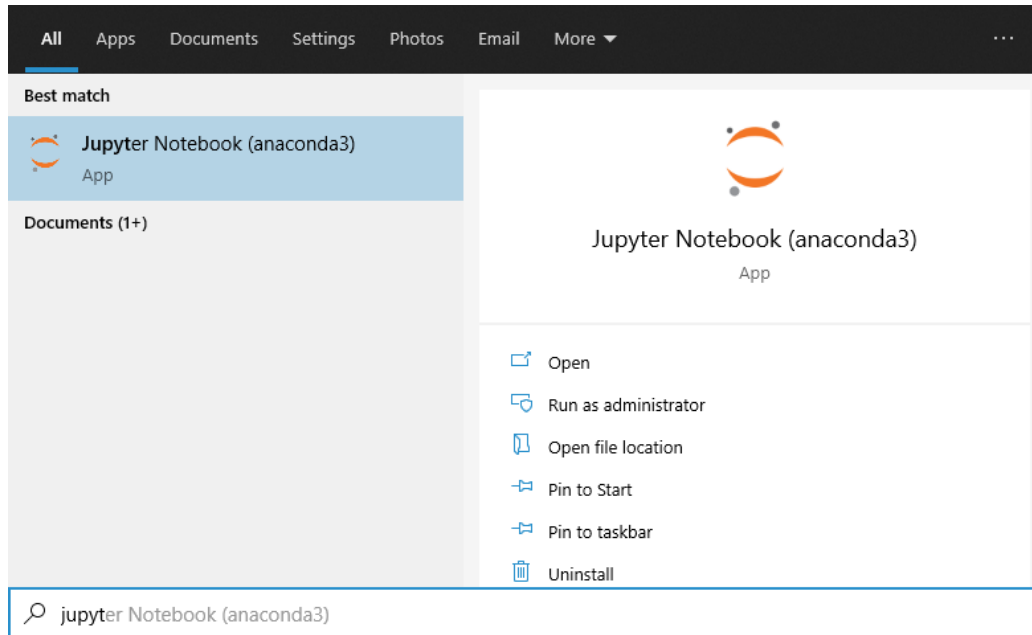
Hands-On

02

Penggunaan

Buka lewat Windows Menu

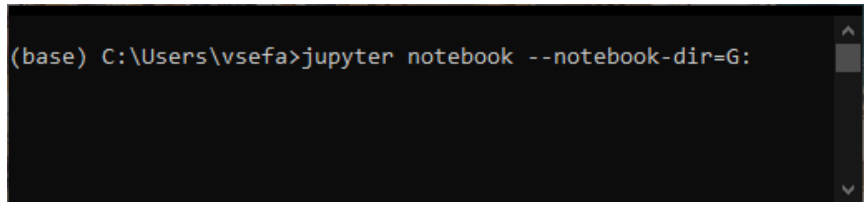
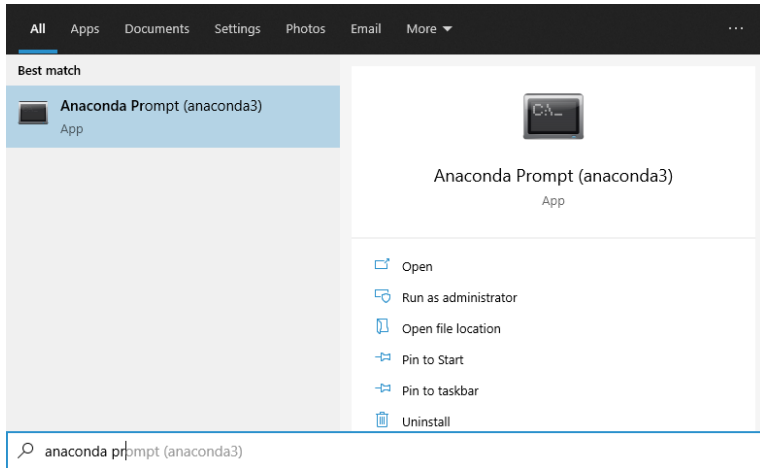
- Terbuka di C:\Users\username



Buka lewat Anaconda Prompt

- Defaultnya terbuka di drive C:\Users\username
- Tambahkan parameter `--notebook-dir` untuk buka di drive lain

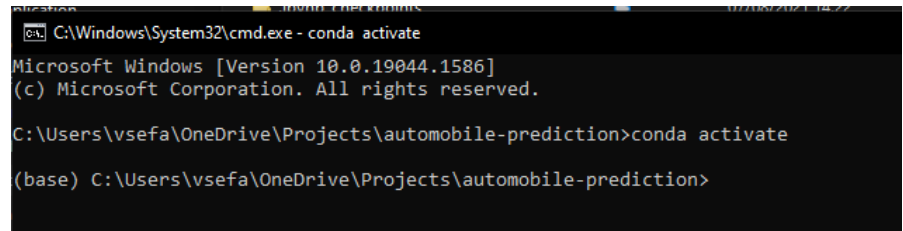
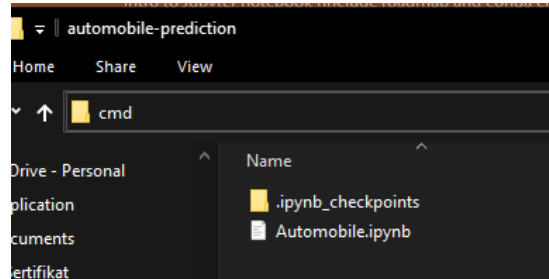
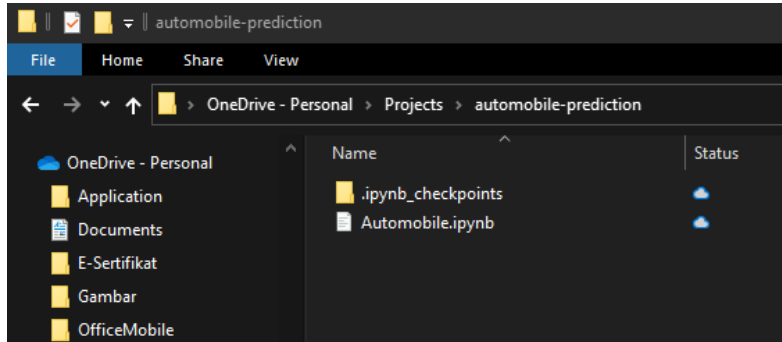
```
(env) jupyter notebook
```



Buka lewat cmd (jika sudah atur path)

- Buka folder project.
- Ketik cmd di bar alamat untuk buka command prompt di lokasi tersebut
- Ketik `conda activate`

`(env) jupyter notebook`



Hands-On

03

Conda Environment

Mengapa Menggunakan Environment

Python 2.7.0 (PC kita)

Tensorflow 2.1.0 • numpy 2.7 • matplotlib 3.1.0 • scipy 1.10.3

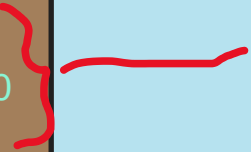
Project Object Detection (2008)

- Python 2.7.0
- Tensorflow 2.1.0
- Numpy 2.7.1
- Matplotlib 3.1.0
- Scipy 1.10.3

Project Text Clustering with Deep Learning

- Python 3.8.0
- Tensorflow 3.0.0
- Numpy 2.9
- Matplotlib 3.1.0
- Scipy 1.10.3

Dependency break



Mengapa Menggunakan Environment

conda

obj-detect-env

python 2.7.0
tensorflow 2.1.0
numpy 2.7
matplotlib 3.1.0
scipy 1.10.3

Project Object
Detection
(2008)

text-cluster-env

python 3.8.0
tensorflow 3.0.0
numpy 2.9
matplotlib 3.1.0
scipy 1.10.3

Project Text
Clustering with
Deep Learning

base

python 3.9.0
tensorflow 3.1.0
numpy 2.7
matplotlib 3.7.0
scipy 4.10.3



Membuat Environment

1. Membuat environment kosong

```
conda create --name myenv
```

2. Membuat environment dengan Python versi spesifik

```
conda create --name myenv python==2.7.0
```

3. Membuat environment dengan Python versi spesifik + packages tambahan

```
conda create -name myenv python==2.7.0 pandas==3.4.5 scipy matplotlib
```

Menginstall Packages di dalam Environment

Aktifkan environment terlebih dahulu

```
conda activate myenv  
(myenv) C:\lokasi\project>
```

Install package (contoh: jupyter notebook)

```
(myenv) C:\lokasi\project> conda install jupyter
```

Hands-On

Berbagi Environment

Anggaplah kita buat sebuah environment berikut

```
conda create -name myenv python==2.7.0 pandas==3.4.5 scipy matplotlib
```

Di dalam environment tersebut, buat file dengan nama *environment.yml* yang isinya package yang dipakai dalam environment kita

```
conda env export --from-history > environment.yml
```

Di komputer orang lain, gunakan *environment.yml* tadi untuk membuat environment yang sama persis dengan milik kita

```
conda env create -f environment.yml
```

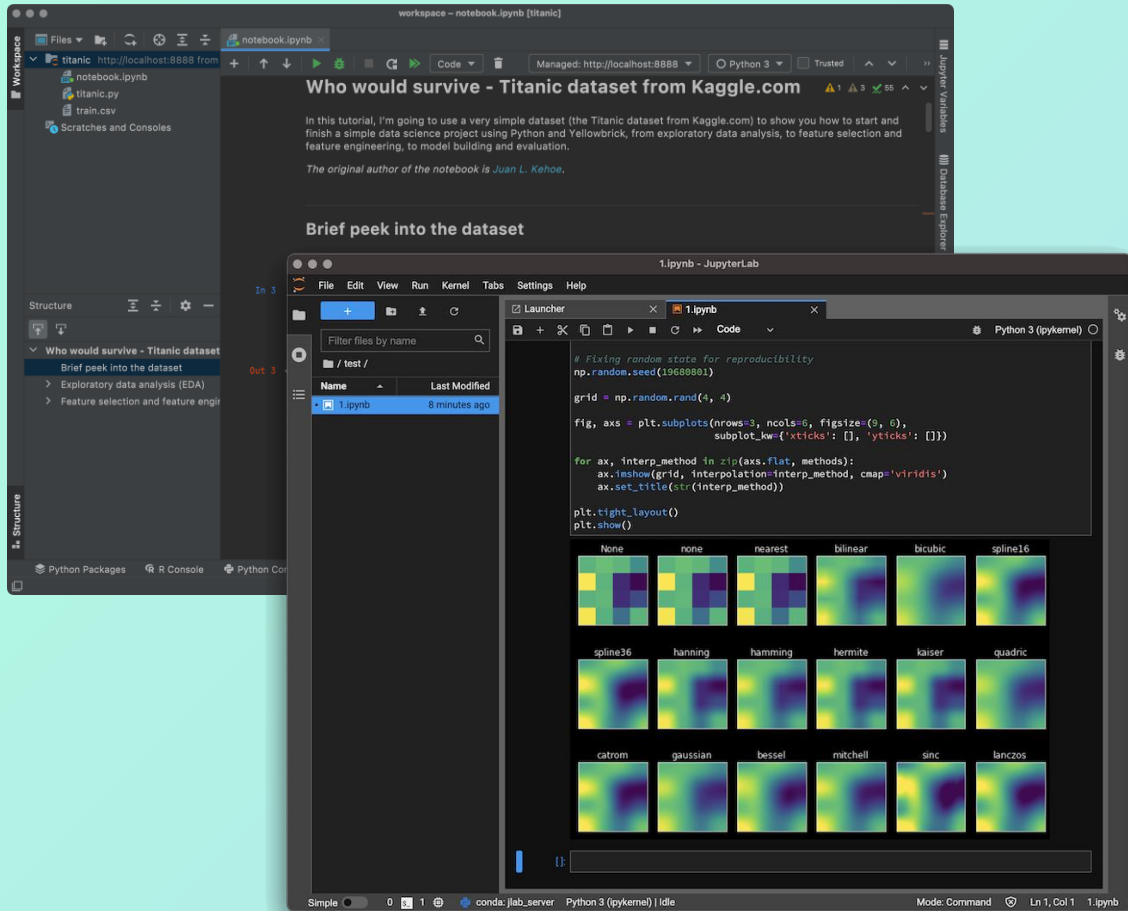
04

Jupyter Notebook Alternative

The image displays two overlapping screenshots of web-based notebook environments. The top screenshot shows the Google Colab interface with a 'Welcome to Colab!' message and a code cell containing `import time` and `import concurrent.futures`. The bottom screenshot shows a Jupyter Notebook interface with an 'Environment' sidebar listing installed packages like `absl-py`, `aiosignal`, and `astunparse`. The main code area contains a script for installing and running a clustering model using `scikit-learn`, `wordcloud`, and `matplotlib`.

Aplikasi Jupyter Notebook di web dan cloud

- Google colab
- Kaggle notebook
- IBM Cloud Data Pak
- JetBrains Datalore



Aplikasi Jupyter Notebook di desktop

- JetBrains dataspell
- JupyterLab Desktop

Referensi

1. <https://www.freecodecamp.org/news/data-science-learning-roadmap/>
2. <https://i.am.ai/roadmap/>
3. <https://docs.conda.io/en/latest/miniconda.html>
4. <https://www.machinelearningplus.com/deployment/conda-create-environment-and-everything-you-need-to-know-to-manage-conda-virtual-environment/>
5. <https://www.codecademy.com/article/setting-up-jupyter-notebook>