Guide to install Python and dependencies

Contents

[1. Installing Python 3 2](#_Toc102663662)

[2. Setting up PIP 2](#_Toc102663663)

[3. Installing Python packages 5](#_Toc102663664)

[4. Setting up TensorFlow 6](#_Toc102663665)

# Installing Python 3

Python version used in the software: 3.6

To download Python 3.6.0, please head to <https://www.python.org/downloads/release/python-360/>. If necessary, security updates for Python 3.6 can be found at <https://www.python.org/downloads/>.

# Setting up PIP

You will need to set up PIP to install Python packages (such as Numpy, Matplotlib etc). To do this, you will need to edit your computer’s environment variable. Follow the steps:

2.1) Type “environment variable” into the Windows search bar and you will see the following option:

Graphical user interface, application, Word

Description automatically generated

2.2) Click on “Edit the system environment variables”, which will lead you to this pop-up window.

Graphical user interface, text, application, email

Description automatically generated

2.3) Click on “Environment Variables…” and you will see another pop-up window as shown below.

Graphical user interface, text, application

Description automatically generated

2.4) We will need to add values to the “Path” variable for **both user and system variables**. Search for the “Path” variable and double click on it. It will lead to the pop-up window as shown below.

Graphical user interface, text, application

Description automatically generated

2.5) You will need to add 3 directories to set up PIP, as shown by the directories marked by a red marker. Note that you will also need to add another 2 directories as indicated by the blue marker, but they will be explained later on.

2.6) To obtain the directories marked in red, open up the Window Explorer, and type “%appdata%” into the directory bar as shown below.

Graphical user interface, application, table, Excel

Description automatically generated

This will lead you to “/AppData/Roaming”. Redirect yourself to the following directory as shown below.

Graphical user interface, application

Description automatically generated

2.7) If you had installed Python 3.6, you would see a folder named “Python36”. The 3 directories that you need are inside this folder – they are “Scripts”, “Lib/site-packages” and the root folder itself. The “Script” folder allows you to use the PIP command in the command prompt to manage your packages. The “Lib/site-packages” tells the system where the Python packages are installed. The root folder itself allows you to run Python directly on the command prompt.

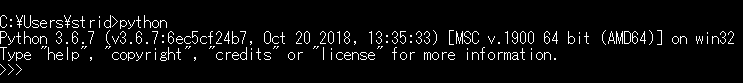
You can open up command prompt and perform the following actions to see if the correct directories have been added.

PIP is working if you can see the following text after typing “pip” and pressing enter:

Text

Description automatically generated

2.8) You can also check if the command “python” works by typing “python” and pressing enter:



If everything is alright, you may now proceed to install the following python packages.

# Installing Python packages

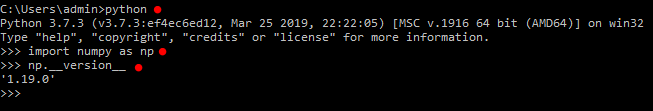
To install Python packages via command prompt, you will be using the “pip install” command. For example, to install the Numpy package, you will key in “pip install numpy” and press enter. However, doing so without specifying the version will cause the system to download and install the latest version of the package. While this is not a problem if you are only installing one or two packages, it can result in conflicts if you are mass installing packages which we will be doing later. This is because some of the packages rely on a certain version of another package (known as the dependency) to work, and using the latest version of the dependency may cause some functions to break.

To install a specific version of a package, for example, key in “pip install numpy==1.18” to install the version 1.18 of the Numpy package. Please note that there should be no spacing when indicating the version.

You will need to install the following packages that were used for this study:

|  |  |  |
| --- | --- | --- |
| **Packages required** | **Version** | **Comments** |
| hdbscan | 0.8.26 | Newer versions may work fine but have not been tested |
| imgaug | 0.4.0 |  |
| matplotlib | 3.1.3 | Newer versions may work fine but have not been tested |
| numpy | 1.18 |  |
| opencv\_contrib\_python | 4.2.0.34 |  |
| scikit-image | 0.18.rc1 | Newer versions may work fine but have not been tested |
| seaborn |  | No specific version required |
| PyQt | 5.14.1 | Please make sure you do not use PyQt6 as the Matplotlib package does not support the newest version of PyQt yet |
| keras | 2.3.1 |  |
| tensorflow\_gpu | 2.1.0 | Please do not install the “tensorflow” package that does not include the term “gpu” in it as that package utilizes the CPU instead. |

To check if the packages have been installed, you can use the command prompt to do so as indicated by the red marker in the picture below (using Numpy as an example):



Alternatively, you may also use the “pip list” command to check both the packages and the version installed. The next section will describe the procedures to set up TensorFlow on your computer.

# Setting up TensorFlow

As indicated previously, we will now need to setup the following directories marked in blue. You will need to install one program before setting up the directories in the environment variable setting.

Graphical user interface, text, application

Description automatically generated

You will first need to install the NVIDIA CUDA toolkit version 10.1, which can be obtained via <https://developer.nvidia.com/cuda-10.1-download-archive-base>.

After installation, go to the following directories and check that you have these 2 files:

|  |  |
| --- | --- |
| **Directory** | **File to check** |
| …/CUDA/v10.1/bin/ | cudart64\_101.dll |
| …/CUDA/v10.1/extras/CUPTI/lib64/ | cupti64\_101.dll |

Once done, you can set up the directories in the environment variable setting. You can also test if you are able to use TensorFlow (GPU version) on your computer by doing the following:

Text

Description automatically generated

The red markers indicate what to type in, while the aqua markers indicate what to look out for. If TensorFlow-GPU was installed successfully and is able to utilize your computer’s GPU, you will also be able to find your GPU model indicated by the first aqua marker.

**You are now ready to run the software for cluster analysis!**