AllenSDK configuration guide

Guide contents

This tutorial three main files:

- This guide, which contains information about installation, configuration of environment and potential troubleshooting.
- `download_data.ipynb`, a simple notebook designed for you to test the environment and download the necessary data to start working.
- `example_notebook.ipynb`, which contains example code on how to work with the dataset, including a simple visual decoding example.

Before the first practical lesson takes place, you are supposed to read this guide, install the AllenSDK, and make sure you can run the `download_data.ipynb`. This will configure a setup and make sure you can download data from the repository.

Later, in the first practical lesson, we will go together through the `example_notebook.ipynb` and comment the code.

If you find any installation/configuration problem, please write us immediately!

Useful links

You can refer to the following links for more information about the experiments, as well as notebooks.

- In the main page of the neuropixels data you can find more tutorials. In particular, our tutorial was done from the Data Access, the Quickstart and part of the Full guide tutorials together, but more information and examples can be found there.
- Information about the natural scenes and the natural movies.
- A cheatsheet with relevant info and common functions.
- The Allen Brain Atlas Data Portal has some nice visualizations and information about the data.
- Allen Brain Interactive Atlas is a web application that allows you to see where the mice brain areas are located. The names used are the same as the ones in the dataset, so it's a nice reference.

Environment setup

In theory, the installation should be simple. However, installation problems may be common, so I **highly recommend** using isolated environments.

Each environment lives in a local folder in your computer and can handle different Python versions, as well as different packages. Thus, if something goes wrong, deleting the environment solves the problem. Conda is an excellent manager that you should download if you haven't done it yet.

During the project we will work with notebooks. Notebooks are interactive documents that can contain texts, show graphs, and run Python code. A popular way to use notebooks is via Jupyter Notebooks. Other code editors such as VSCode can also edit notebooks. Just use your favorite code editor.

Please make sure your Conda installation is OK. Once Conda is ready, we start configuring the environment. Copy the two notebooks (`.ipynb` files) to an empty folder for the project, say, `myproject`.

Create the environment

Open a terminal, create the environment, activate it and install the required packages to work by running

```
conda create -n envallen python=3.11
conda activate envallen
pip install allensdk scikit-learn jupyterlab
```

You might find that `pip` is not installed in the new environment, depending on your configuration. If this is the case, just run `conda install pip` and then install the packages. Also note that `pip` might be called `pip3`. **Do not try to install AllenSDK from `conda`**, as the latest stable version is not available.

If the installation isn't successful, check the Python version of your environment. **The AllenSDK will not work for Python 3.12 or newer**. You can indicate conda the Python version for your environment as we did above. If you find any problem you can try downgrading the Python version.

If that does not work, please send an email to me indicating your commands and the obtained error. We may file an issue on AllenSDK's repo. Sometimes, recent updates of the SDK or its dependencies might break compatibilities and support from the organization is needed.

At the end of each session once you end working with your environment, don't forget calling `conda deactivate`. This will avoid you accidentally installing or updating other packages whose installation might interfere with our setup.

Download the data

Now we will set if the installation above was succesfull. We will do so by requesting the AllenSDK to download some example data that we will use later for the example notebook.

The `download_text.ipynb` contains everything we need to set up the AllenSDK environment, download the dataset, and try to use it.

Initialize the Jupyter notebook by opening a terminal, moving to the "myproject" folder and typing `jupyter lab` (or `jupyter notebook`). Jupyter will open in your web browser. Remember to call `conda activate allensdk` before (if you haven't already) so the default kernel corresponds to your new environment.

It will be important to create a folder for all the data that we are going to download. By default, our notebook is configured to put everything in a folder called `allendata`. **Please create a folder `allendata` inside `myproject`, next to the notebooks.**

Open the `download_text.ipynb` notebook in Jupyter. Read the instructions and execute the code cells one by one. Take into account that the download of data might take some time. In total, it might be up to (20min - 1 hour).

If you get the result of the last cell (the call to `oursession.metadata`) without issues, you'll be able to operate without problems. **I highly recommend you to open a RAM monitor before executing that last cell, especially if you have an 8GB computer. If you get above 80% of your RAM just interrupt the Python kernel and restart it.**

We have realised that **loading the dataset in Python might take 5 to 6 GB of RAM**. Therefore, computers with 8GB of RAM or less might be unable to run it. **We have checked the possibility of extracting the essential data and working with, just without the AllenSDK API.** Basically, I wrote a script that takes the full AllenSDK object and writes it to several small files that you can read. This means that some SDK convenience functions will not be available but it's not a big deal.

Example notebook

If you are able to run the `download_notebook.ipynb` up to the last step, you can use the `example_notebook.ipynb`. We will go together through this notebook in a practical lesson, so I highly recommend you to do the installation and download beforehand, and reaching to us if you find any problem during the procedure.