

Python Debugging FAQs

For some of you, MLRD will be the first time you are running Python locally on your machine. While Unix based systems include a system supported installation of Python, Windows does not, which can present issues when getting set up. In this document, we address some common Python issues that come up in Tick 1.

We aim for MLRD code to be system agnostic, however, it is only tested on Unix based systems (Ubuntu and MacOS), so there may be unforeseen specifics affecting the Windows operating system later on in the course. Our recommendation for the MLRD practicals (and beyond) is to either install a Unix based operating system or to use a virtual machine with one (e.g. [VirtualBox](#), or [WSL](#)). It is the de facto standard when working with machine learning algorithms in real-world applications. If you decide to use Windows, we do our best to provide advice and solutions to problems you might encounter when trying to run MLRD code.

In terms of IDEs, we recommend PyCharm. You can get a fully licenced version for free as a Cambridge University student.

1. Windows: My command line tells me that NLTK is not installed but I installed it.

This is probably because you have two versions of Python, and the one it installed on is not the same as the one you are running. You can check the version you are running using `python3 --version`. Running `pip3 install nltk` usually installs it to this version. If you are working in an IDE like PyCharm, you can access a Python terminal at the bottom; installing NLTK in that terminal should install it correctly in whichever Python environment the IDE is using. In Windows, you might have to write `py` rather than `python3`.

Generally, if you have multiple Python versions installed on your computer, we recommend uninstalling all of them and to reinstalling a single one (3.6–3.10 should be fine). When installing Python, make sure to add it to PATH as a system variable. The easiest way to do that is to tick a small box in the Python installer.

2. The program cannot find the folder `utils`, or cannot find the file or path `data/sentiment_detection/reviews`?

This is happening because you are running the python code in the wrong directory. You need to run the code from the root directory, i.e. from inside `mlrd/` run `python3 exercises/tick1.py`. On Windows, you sometimes have to run it as a module, with `python3 -m exercises.test1`, and on unix you might need to specify to the Python path, `PYTHONPATH='.' python3 exercises/tick1.py`

An IDE might automatically set up the default execution path incorrectly. In PyCharm you can fix this by editing the run configuration (top right, to the left of the green ‘play’ button). Remove `exercises/` from the working directory, so it runs from `mlrd/`.

We do not recommend that you modify the folder structure to fix this problem.

3. Macintosh: NLTK is giving an SSL Certificate Error?

On Macintosh computers, downloading `punct` in NLTK sometimes leads to an SSL Certificate Verification Error. To get around this, run the following command in terminal, changing 3.6 to the Python version you have installed: `/Applications/Python 3.6/Install Certificates.command`. See [this Stack Overflow post](#).