PHYMOL Kick-off conference Hands-on ML tutorial

Dahvyd Wing
University of Luxembourg

September 2023

1 Installation

These instructions create a python virtual environment with PyTorch and SchNetPack[1].

1. Install anaconda and open an anaconda prompt
2. Create a virtual environment:

   `conda create --name ml_tutorial python=3.8`
3. Activate the new environment:

   `conda activate ml_tutorial`
4. Test to make sure that the pip, python, etc. all point to executables in the conda environment:

   If you are using linux or macos, run `which pip`. If you are using windows, run `where pip`. Check that the first/only executable is in the anaconda virtual environment path.
5. Install the following packages

   `pip install schnetpack spyder tensorboard`

You can use a different IDE if you would like, but you have to ensure that they are using the virtual environment.

2 Setup training set

1. Download the `schnet_tutorial` folder from the phymol/2023_workshop_tutorials github repository.
2. Download the md17 ethanol database [2] at `http://quantum-machine.org/gdml/data/npz/md17_ethanol.npz` (this file is 219 MB) and place this in the `schnet_tutorial` folder
3. In the `schnet_tutorial` folder run `python script1_generate_dataset.py`

References

1. Schütt, K. T., Hessmann, S. S. P., Gebauer, N. W. A., Lederer, J. & Gastegger, M. SchNetPack 2.0: A neural network toolbox for atomistic machine learning. *The Journal of Chemical Physics* **158,** 144801. ISSN: 0021-9606 (Apr. 2023).
2. Chmiela, S. *et al.* Machine learning of accurate energy-conserving molecular force fields. *Science Advances* **3,** e1603015 (2017).