CSCI E-104 -
Final Project: ChemicalX: A Deep Learning Library for Drug Pair Scoring
Installation Guide ChemicalX

Daniel Ardila

April 2024

Installation Commands

0.1 Local Installation

0.1.1 Setting up a Python Environment

1. **Install Anaconda or Miniconda:** Begin by installing Anaconda (which includes Python and conda) or Miniconda (a lighter version with only conda and Python) from `https://www.anaconda.com/products/individual` or `https://docs.conda.io/en/latest/miniconda.html`, respectively.

2. **Create a New Conda Environment:** Open the Anaconda Command Prompt or your terminal and create a new environment specifically for this project:

```
# Create a new conda environment
conda create -n myenv python=3.10
```

3. **Activate the Environment:** Activate the newly created environment:

```
# Activate the environment
conda activate myenv
```

After activating your environment, you might want to use this environment in Jupyter Notebook. To do so, follow these steps:

4. **Install IPyKernel:** Install 'ipykernel' which allows the environment to be used as a kernel in Jupyter:

```
# Install ipykernel
conda install ipykernel
```

5. **Add Environment as a Kernel:** Add your newly created environment as a kernel in Jupyter:

```
# Add environment as a kernel
python -m ipykernel install --user --name myenv --display-name "
    Python (myenv)"
```

This will make the environment available as "Python (myenv)" in the kernel options when you open or create Jupyter notebooks. This step is useful if you wish to keep your project's dependencies isolated but still use Jupyter for interactive development.

6. **Install Required Packages:** Within the activated environment, install the necessary libraries:

```
# Install PyTorch, PyTorch Geometric, TorchDrug and ChemicalX
    libraries
pip install torch
pip install torch-scatter torch-cluster -f https://pytorch-
    geometric.com/whl/torch-1.8.0+cu102.html
pip install torchdrug
pip install chemicalx
```

Fixing ImportError Locally

If you encounter an 'AttributeError' when importing 'chemicalx', indicating a missing 'Molecule.dummy_atom' attribute, you will need to modify the 'constants.py' file in the 'chemicalx' package. This error is due to recent changes in 'torchdrug' which 'chemicalx' has not yet accommodated.

Steps to fix the issue:

1. **Locate the File:** Navigate to your Python environment's site-packages directory, usually found at: `<Your Environment Path>/Lib/site-packages/chemicalx`

2. **Edit the File:** Open 'constants.py' and locate the line:

```
TORCHDRUG_NODE_FEATURES = len(atom_default(Molecule.dummy_atom))
```

Replace it with:

```
TORCHDRUG_NODE_FEATURES = len(atom_default(Molecule.dummy_mol.
    GetAtomWithIdx(0)))
```

Save your changes. This modification corrects the reference to the updated 'Molecule' class structure in 'torchdrug'.

0.2 Installation on Google Colab

For running the project on Google Colab, follow these steps to ensure all dependencies are installed:

1. **Start a New Notebook:** Open Google Colab at `https://colab.research.google.com/` and start a new notebook.

2. **Install Required Libraries:** Since Google Colab runs isolated instances, you'll need to install required packages every time you start the notebook. Use the following commands in a Colab cell:

```
# Install necessary libraries in Colab
!pip install torch
!pip install torch-scatter torch-cluster \
    -f https://pytorch-geometric.com/whl/torch-1.8.0+cu102.html
!pip install torchdrug
!pip install chemicalx
```

Fixing ImportError in Google Colab

Due to a compatibility issue between the `chemicalx` and `torchdrug` libraries, the following error may occur when attempting to import `chemicalx`:

```
AttributeError: type object 'Molecule' has no attribute 'dummy_atom'
```

This error is caused by outdated references in `chemicalx` to elements of the `torchdrug` library that have been updated. To fix this error in Google Colab, perform the following steps:

1. Open a new cell in your Google Colab notebook.
2. Execute the following command to modify the `constants.py` file directly using the `sed` stream editor. This command finds and replaces the problematic line:

```
!sed -i 's/dummy_atom/dummy_mol.GetAtomWithIdx(0)/' $(find /
    usr/local/lib/python3.* -type f -name constants.py)
```

3. Restart the runtime in Google Colab by selecting "Runtime" ¿ "Restart runtime" from the menu to ensure that all changes take effect.
4. Re-run the import command to verify that the issue has been resolved:

```
import chemicalx
```

This adjustment corrects the reference to the updated structure of the `Molecule` class in `torchdrug`, resolving the `AttributeError`.