

Essential Bioinformatics - Interactive 3D viewer

CSI 5180 - Machine Learning for Bioinformatics

Marcel Turcotte

2025-01-20

Learning objective

- **Illustrate** the process of identifying and resolving missing library issues in Google Colab.

Example

Missing Library

The following code, when executed in [Google Colab](#), is expected to generate an error due to the absence of the required library in the environment.

```
import py3Dmol
```

Here is the expected message.

```
-----  
ModuleNotFoundError                                Traceback (most recent call last)  
<ipython-input-1-4554d9dd9fa2> in <cell line: 0>()  
----> 1 import py3Dmol  
  
ModuleNotFoundError: No module named 'py3Dmol'
```

This issue can be resolved by adding the following line of code before the first `import` statement. Try it!

```
! pip install py3Dmol
```

In Jupyter Notebooks, the `!` symbol is used to execute shell commands directly within the notebook environment. When utilizing Google Colab, these notebooks operate within a virtual machine running a Unix-based system.

To validate the claim, create an executable cell and run the following commands.

- `uname`
- `uname -a`
- `ls`
- `pwd`
- `ls /`
- `cat /proc/cpuinfo`
- `ls sample_data`
- `cat sample_data/README.md`

We had the option to identify the missing library and ensure its execution automatically. However, I wanted to have this discussion beforehand.

To detect and resolve the issue of a missing library, we can implement a method to enforce its installation. By commenting out the initial import statement and utilizing the subsequent import, the notebook will successfully load in Google Colab.

```
try:  
    import py3Dmol  
except:  
    ! pip install py3Dmol  
    import py3Dmol
```

Retrieving Data

In our initial notebook, we utilized the Ensembl REST API to retrieve relevant data. In the following section, we leverage the capabilities of [py3Dmol](#) to facilitate direct data retrieval from the [RCSB Protein Data Bank \(PDB\)](#).

```
view = py3Dmol.view(query='pdb:1hvr')  
view.setStyle({'cartoon':{'color':'spectrum'}})  
view
```

Unable to display output for mime type(s): application/3dmoljs_load.v0, text/html

```
<py3Dmol.view at 0x114f1beb0>
```

Given that our notebook is executed in a Unix environment on Google Colab, we could alternatively utilize Unix commands, such as `wget`, to download the file and subsequently employ the `unzip` command to decompress it.

When utilizing Google Colab, it is crucial to understand that all files stored during a session will be lost upon its termination. To mitigate this, you can integrate your personal Google Drive with the session, though this approach entails specific security considerations, particularly when employed for academic assignments. Nonetheless, downloading data, as demonstrated below, is generally a straightforward process.

```
! wget http://www.rcsb.org/pdb/files/5RH2.pdb.gz
! gunzip -f 5RH2.pdb.gz
```

```
URL transformed to HTTPS due to an HSTS policy
--2025-03-05 17:59:26-- https://www.rcsb.org/pdb/files/5RH2.pdb.gz
Resolving www.rcsb.org (www.rcsb.org)... 128.6.159.248
Connecting to www.rcsb.org (www.rcsb.org)|128.6.159.248|:443... connected.
HTTP request sent, awaiting response... 301 Moved Permanently
Location: https://files.rcsb.org/download/5RH2.pdb.gz [following]
--2025-03-05 17:59:27-- https://files.rcsb.org/download/5RH2.pdb.gz
Resolving files.rcsb.org (files.rcsb.org)... 18.67.17.66, 18.67.17.27, 18.67.17.28, ...
Connecting to files.rcsb.org (files.rcsb.org)|18.67.17.66|:443... connected.
HTTP request sent, awaiting response... 200 OK
Length: 108944 (106K) [application/octet-stream]
Saving to: '5RH2.pdb.gz'

5RH2.pdb.gz          0%[                ]    0  --.-KB/s          5RH2.pdb.gz

2025-03-05 17:59:27 (3.97 MB/s) - '5RH2.pdb.gz' saved [108944/108944]
```

We are now able to visualize the contents of the file.

```
view = py3Dmol.view()
view.addModel(open('5RH2.pdb', 'r').read(), 'pdb')
view.setBackgroundColor('white')
view.setStyle({'chain': 'A'}, {'cartoon': {'color': 'purple'}})
view.addStyle({'resn': 'UH7'}, {'stick': {'colorscheme': 'yellowCarbon'}})
view.addStyle({'within': {'distance': '5', 'sel': {'resn': 'UH7'}}}, {'stick': {}})
view.zoomTo()
view.show()
```

Unable to display output for mime type(s): application/3dmoljs_load.v0, text/html

References

- Based on [py3DMol.ipynb](#) by [CCBatIIT](#). Visited on 2025-01-20.