

EventDisplay.Py File

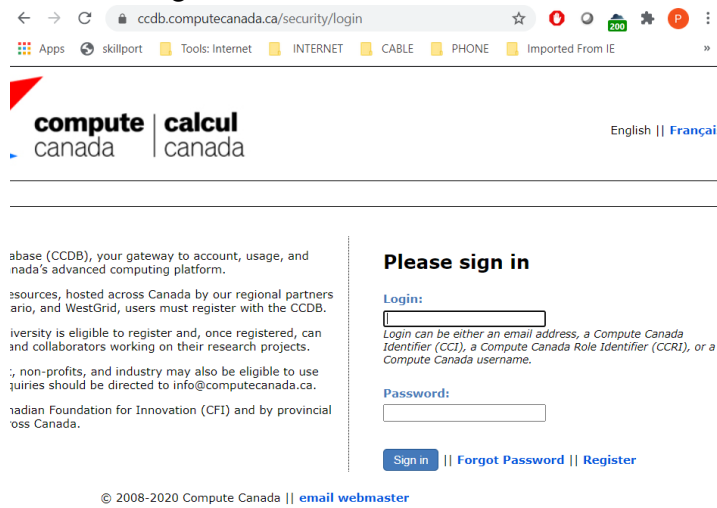
The event display file is a python file which combines events from two npz files for exploratory data analysis. The npz files we use contain the following data arrays ['event_id', 'root_file', 'pid', 'position', 'direction', 'energy', 'digi_hit_pmt', 'digi_hit_charge', 'digi_hit_time', 'digi_hit_trigger', 'true_hit_pmt', 'true_hit_time', 'true_hit_pos', 'true_hit_start_time', 'true_hit_start_pos', 'true_hit_parent', 'track_id', 'track_pid', 'track_start_time', 'track_energy', 'track_start_position', 'track_stop_position', 'track_parent', 'track_flag', 'trigger_time']. We use these arrays to manipulate our data into an h5 file to analyze the events being combined.

Install NPZ File

To run the event display file, we need to have a npz and to do that complete the following steps:

1. Have an active compute Canada cedar account, if yes proceed to step 2.

If you do not have an account, ask your sponsor to provide you a link to register for an account.



The screenshot shows the login page for Compute Canada. The browser address bar displays 'ccdb.compute canada.ca/security/login'. The page header includes the 'compute canada' and 'calcul canada' logos, along with language options for 'English' and 'Français'. On the left side, there is a block of text providing information about the CCDB, its regional partners, and eligibility for registration. The main content area is titled 'Please sign in' and contains a 'Login:' section with a text input field and a note stating that login can be via email, CCI, CCRI, or username. Below this is a 'Password:' section with another text input field. At the bottom of the login area are three buttons: 'Sign in', 'Forgot Password', and 'Register'. The footer of the page shows the copyright '© 2008-2020 Compute Canada' and a link to 'email webmaster'.

2. Open a WSL terminal and enter `ssh <yourusername>@cedar.compute canada.ca`. Then enter your password when given the prompt.

```

pavan@LAPTOP-SLR81BE5:~$ ssh pjosan@cedar.computeCanada.ca
pjosan@cedar.computeCanada.ca's password:
Last login: Thu Nov 12 22:47:12 2020 from d207-023-213-231.wireless.sfu.ca
=====
Welcome to Cedar!

Cedar Power Outage Dec. 1, 06:27
-----
The datacentre that houses the Cedar facility suffered a brief power outage.
All jobs that were running at that time are lost and need to be resubmitted.

There was a problem with GPU drivers that has been fixed. But some GPU jobs may
have failed because of that problem.

Cedar Downtime Dec. 7-9, 2020
-----
The Cedar system will be unavailable Dec. 7 - 9 for system upgrades:
1) Replacement of /scratch metadata storage with SSD based system.
2) Operating system upgrades.
3) Scheduler software (Slurm) upgrade to the latest version. The current version
   19.05 is no longer supported.
4) Lustre filesystem upgrade (/home, /project, /scratch).
5) OmniPath driver upgrade to re-enable GPU-direct support.
6) NVIDIA driver upgrade.
7) Installation of optical taps on Cedar's 100GigE links to prepare for network
   security monitoring.
All jobs that are still running on Dec. 7, 8:00am (Pacific) will be terminated
at that time.

For information see: https://docs.computeCanada.ca/wiki/Cedar
Email support@computeCanada.ca for assistance and/or to report problems.
=====
[pjosan@cedar1 ~]$

```

Figure 1: Signed into cedar successful.

3. Exit the cedar account by entering exit.
4. There are 0-999 npz files. Choose a number between 0-999. Now we will retrieve the npz file using the following command. :

```

scp -r <your username>@cedar.computeCanada.ca:/project/rpp-
blair2k/machine_learning/data/IWCDmPMT_4pi_full_tank/numpy/mu-
-/E0to1000MeV/unif-pos-R371-y521cm/4pi-
dir/IWCDmPMT_4pi_full_tank_mu-_E0to1000MeV_unif-pos-R371-
y521cm_4pi-dir_3000evts_<file number>.npz < local folder path>

```

```
pavan@LAPTOP-SLR81BE5:~$ scp -r pjosan@cedar.compute.canada.ca:/project/rpp-blair2k/machine_learning/data/IWCDmPMT_4pi_full_tank_numpy/mu-/E0to1000MeV/unif-pos-R371-y521cm/4pi-dir/IWCDmPMT_4pi_full_tank_mu-/E0to1000MeV_unif-pos-R371-y521cm_4pi-dir_3000evts_26.npz /mnt/c/Users/pavan/Downloads
pjosan@cedar.compute.canada.ca's password:
IWCDmPMT_4pi_full_tank_mu-/E0to1000MeV_unif-pos-R371-y521cm_4pi-dir_3 100% 107MB 1.6MB/s 01:08
```

Figure 2: Npz file 26 was successfully downloaded to my Downloads folder

- Confirm npz file has been downloaded to folder you specified in the previous command.

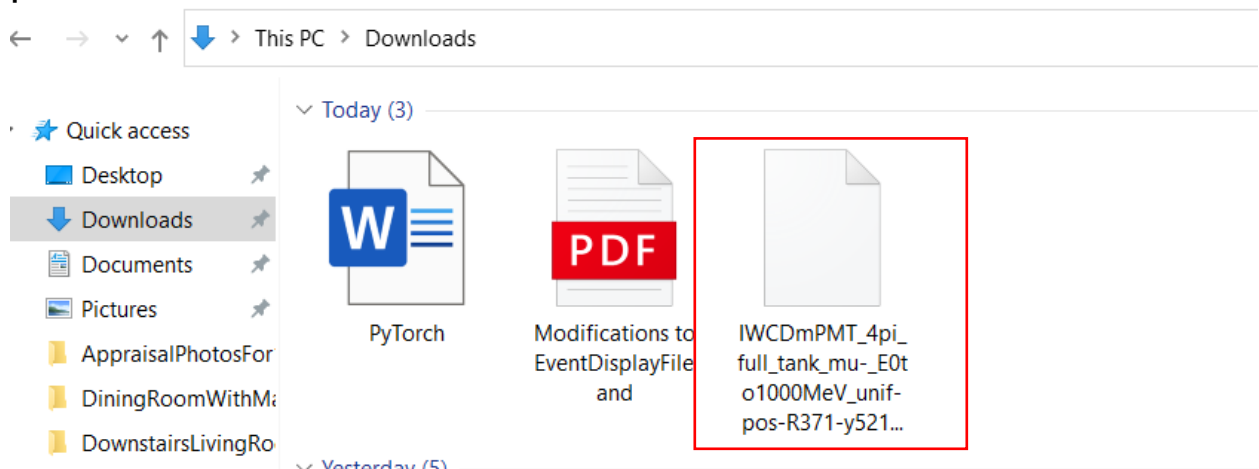


Figure 3: NPZ file is in the directory specified in step 4.

- Rename the file for easy naming convention to event25.npz:
Right click on file, then click on rename.
Change file name to <your choosing>. npz

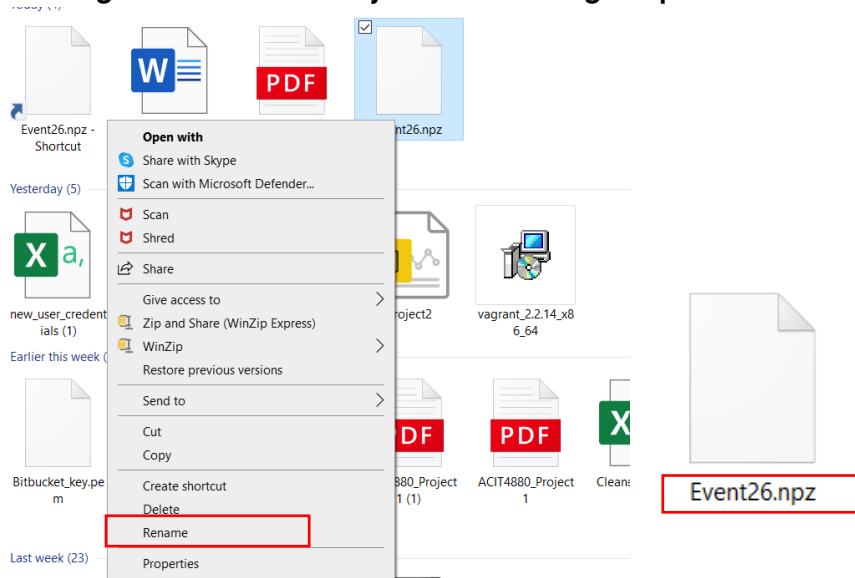


Figure 4: File is renamed to Event25.npz

7. Move NPZ file to project folder

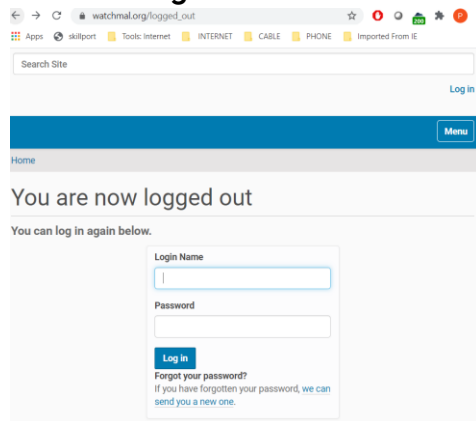
.idea	2020-10-16 8:48 AM	File folder	
venv	2020-10-16 8:50 AM	File folder	
check	2020-10-22 2:14 PM	Python File	1 KB
checknpz	2020-11-05 5:21 PM	Python File	1 KB
cnn_mpmt_dataset	2020-11-26 7:32 PM	Python File	3 KB
dataFile.npz	2020-11-05 3:38 PM	NPZ File	1 KB
Event25.npz	2020-10-16 10:03 AM	NPZ File	110,088 KB
EventDisplay	2020-12-01 6:58 AM	Python File	33 KB
EventDisplay	2020-10-23 8:32 AM	Text Document	0 KB
h5_dataset	2020-11-26 7:26 PM	Python File	3 KB
h5check	2020-11-27 9:10 AM	Python File	0 KB
Histogram	2020-11-26 7:32 PM	Python File	33 KB
mpmt_full_geo.npz	2020-10-16 8:48 AM	NPZ File	96 KB
multimuonfile(2).h5	2020-11-27 9:36 AM	H5 File	1 KB
multimuonfile.h5	2020-10-29 12:04 PM	H5 File	4 KB
npz_to_text	2020-11-26 7:26 PM	Python File	1 KB
PMTlabelSheet3	2020-10-16 8:48 AM	Microsoft Excel Co...	5 KB
smallernpzset	2020-11-05 4:09 PM	Python File	1 KB

Figure 5: NPZ is moved to project folder.

Install Mpmt_full_geo.npz File

1. Have an active WatchMal account, if yes proceed to step 2.

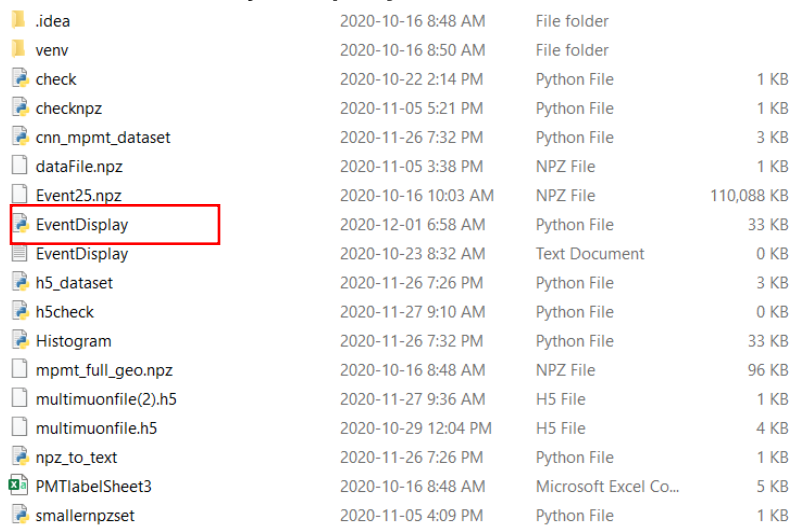
If you do not have an account, ask your sponsor to provide you a link to register for an account.



2. Click the following link to download the mpmt_full_geo.npz file we will require for our EventDisplay.py file:
https://www.watchmal.org/computing/data/detector-geometries/mpmt_full_geo.npz

Install and Run EventDisplay.py File

1. Download the event display file from Jupyter Notebook using the following link: <https://www.watchmal.org/computing/software-resources/jupyter-notebook-tutorials>
2. Add the file to your project folder.



.idea	2020-10-16 8:48 AM	File folder	
venv	2020-10-16 8:50 AM	File folder	
check	2020-10-22 2:14 PM	Python File	1 KB
checknpz	2020-11-05 5:21 PM	Python File	1 KB
cnn_mpmc_dataset	2020-11-26 7:32 PM	Python File	3 KB
dataFile.npz	2020-11-05 3:38 PM	NPZ File	1 KB
Event25.npz	2020-10-16 10:03 AM	NPZ File	110,088 KB
EventDisplay	2020-12-01 6:58 AM	Python File	33 KB
EventDisplay	2020-10-23 8:32 AM	Text Document	0 KB
h5_dataset	2020-11-26 7:26 PM	Python File	3 KB
h5check	2020-11-27 9:10 AM	Python File	0 KB
Histogram	2020-11-26 7:32 PM	Python File	33 KB
mpmc_full_geo.npz	2020-10-16 8:48 AM	NPZ File	96 KB
multimuonfile(2).h5	2020-11-27 9:36 AM	H5 File	1 KB
multimuonfile.h5	2020-10-29 12:04 PM	H5 File	4 KB
npz_to_text	2020-11-26 7:26 PM	Python File	1 KB
PMTlabelSheet3	2020-10-16 8:48 AM	Microsoft Excel Co...	5 KB
smallernpzset	2020-11-05 4:09 PM	Python File	1 KB

Figure 1: EventDisplay.py file is added to a <your project> folder.

3. Open PyCharm and make sure all dependencies are installed for the file. Note: It is assumed you already have python and PyCharm installed and a working interpreter.

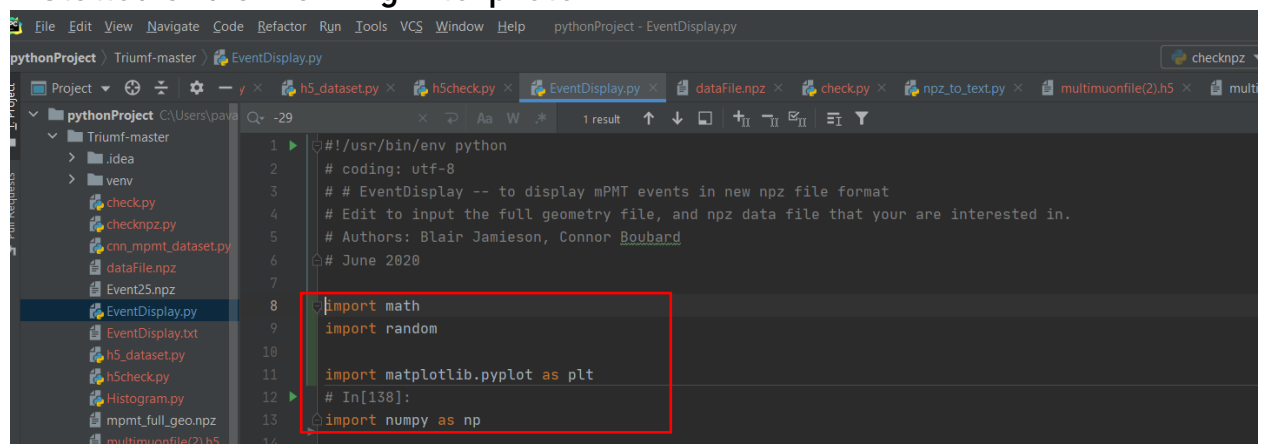
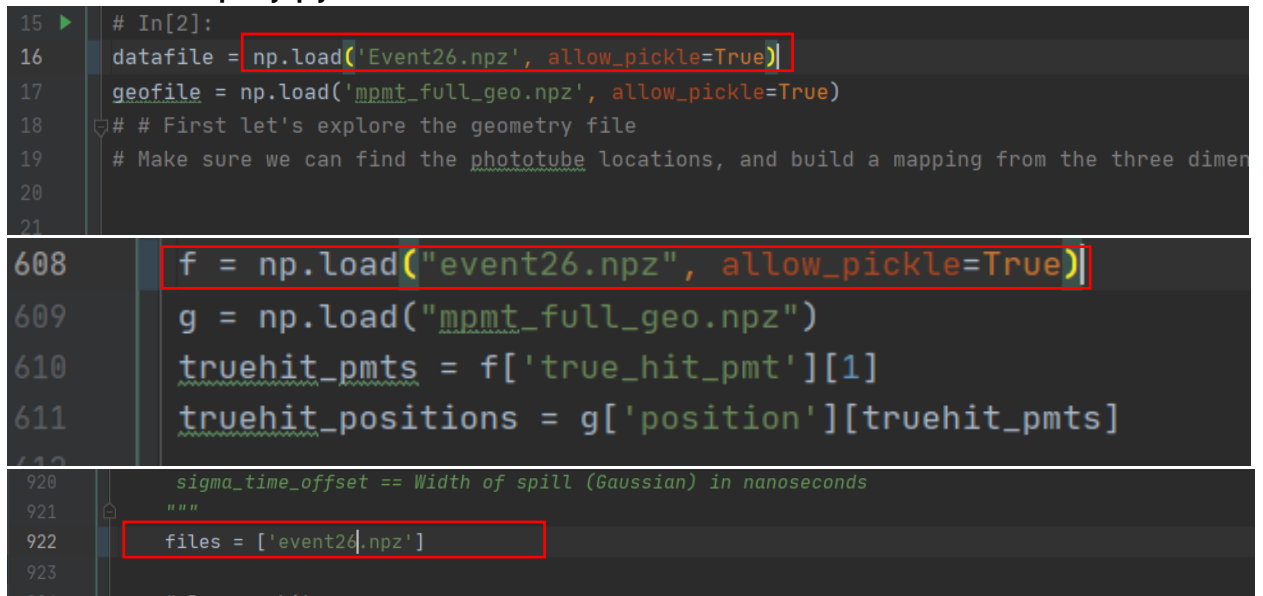


Figure 2: All the dependencies are working because they are added to the python interpreter.

4. Change all occurrence with a different file name to match yours in the EventDisplay.py file



The figure consists of three vertically stacked screenshots of a code editor showing modifications to the EventDisplay.py file. In the first screenshot, line 16 shows `datafile = np.load('Event26.npz', allow_pickle=True)` with a red box around the file name. In the second screenshot, line 608 shows `f = np.load("event26.npz", allow_pickle=True)` with a red box around the file name. In the third screenshot, line 922 shows `files = ['event26.npz']` with a red box around the file name.

```
15 # In[2]:
16 datafile = np.load('Event26.npz', allow_pickle=True)
17 geofile = np.load('mpmt_full_geo.npz', allow_pickle=True)
18 # # First let's explore the geometry file
19 # Make sure we can find the phototube locations, and build a mapping from the three dimer
20
21
608 f = np.load("event26.npz", allow_pickle=True)
609 g = np.load("mpmt_full_geo.npz")
610 truehit_pmts = f['true_hit_pmt'][1]
611 truehit_positions = g['position'][truehit_pmts]
612
920 sigma_time_offset == Width of spill (Gaussian) in nanoseconds
921 """
922 files = ['event26.npz']
923
924 # Remove white spaces
```

Figure 3: Change all occurrences where the npz file will be referenced in the EventDisplay.py file to the name of your file.

5. Run the EventDisplay.py file by clicking on the play icon.

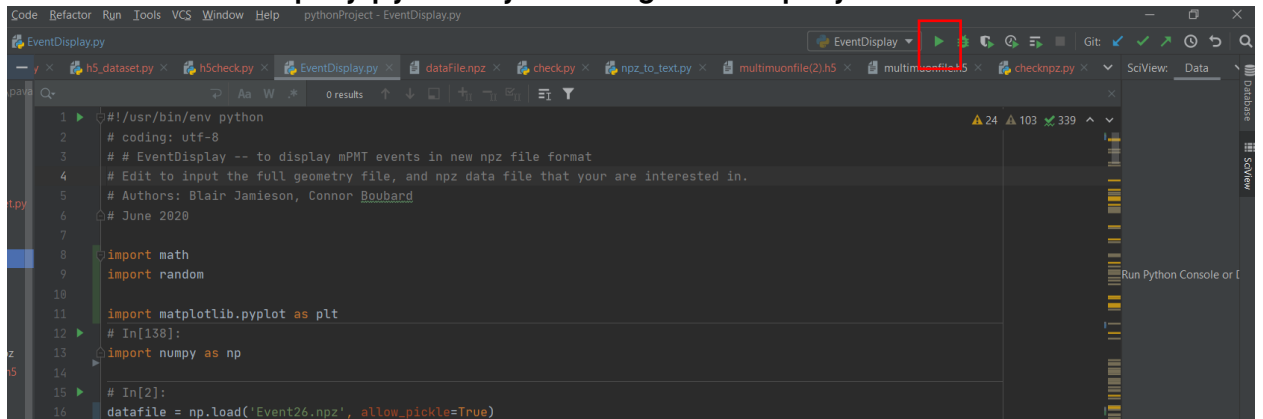


Figure 4: Click the highlighted button to run the file.

6. Running the EventDisplay.py file will generate a h5 file called multimuoanfile(2).h5.

```
950 #sets h5 file to be written
951 h5_file = h5py.File('multimuoanfile(2).h5', 'w')
952 dset_event_data = h5_file.create_dataset("event_data",
953                                         shape=(num_nonzero_events,) + IMAGE_SHAPE,
954                                         dtype=dtype_events)
955 dset_labels = h5_file.create_dataset("labels",
956                                     shape=(num_nonzero_events,),
```

Figure 5: Highlighted file name will be created when the file is run.

Warning: A project level virtual environment is required to install or run the imports.

Note: A container with dependencies installed (singularity) is provided in Compute Canada. Please see the Compute Canada section for instruction on usage.

The static team's task was to fix the bugs within the EventDisplay.py file to composite events and create an H5 file for

Problem: Fix event display script. Debugging took too long as script iterated through 2924 events depending on the npz file which takes about 6 hours.

Approach 1: Shorten the npz file to 10 events and troubleshoot index error faster. Did this manually by hardcoding a subtraction value within the line of code that gets the events. We fixed the index error by setting the for loop to start from zero. Next, we ran into an error with dset_positions. The code that transforms this value from a tensor to a 3D matrix did not work. This was fixed by refactoring the values within the 3D matrix for the

positions being created from the positions array. Then we ran into an error from the line that transform direction using a mathematical equation called `np.arccos`. The issue itself was present because directions returned 1D values and the code was set for two dimensions by previous developer. There were no comments within code to direct of what was happening hence we had to figure this out. The script ran with no problems and an h5 file was produced.

Next our goal was to use the dataloader with the dynamic team to perform an exploratory data analysis on the h5 file. But the current dataloader file does not execute on the h5 file produced therefore this was not feasible given the time left, two weeks.

Steps if we had time

Modify the current dataloader to work with our h5 file and perform an exploratory data analysis from the data collected. Create more histograms for visual presentation of the exploratory data analysis.