Instructions to run our code

The following are instructions to run our code. Note our tutorial is focused on linux machines that run pip virtual environments.

1 Activating the environments

Start by activating an environment type the following into the terminal:

```
virtualenv -p python3 venvE
```

If you use a conda enviornment you can ignore this step; however it is highly recommended to make your own environment. I recommend venvE as we have called it in most our files to run the code.

2 Installing required packages

To install required packages run the following command:

```
pip install -r requirements.txt
```

Otherwise these three packages on top of the CW1-PT environment of COMP0197 will be fine. We had extra packages on requirements however only these three are needed.

1. matplotlib
2. scikit-learn
3. pyav

3 Downloading Datasets

We pre-train our model using iNat we recommend downloading this dataset. Use the following command from torch to first download the dataset. Set root to whichever folder you wish but we recommend root = "./datasets/i-Nat_kingdom/2021_train_mini". We can call the following command

```
iNatdata = datasets.INaturalist(root=root, version='2021_train_mini',
    target_type = 'kingdom')
```

We then transform the above dataset as follows:

```
python image_resizer_imagent.py --in_dir root --out_dir out --size 64
```

Where root = "./datasets/iNat_kingdom/2021_train_mini" and out = "./datasets/iNat64".

Please change the following lines in SimCLR_Scratch_Dev.py starting at line 109:

```
inat_dataset_train = MixedINat128(
'/cs/student/projects3/COMP0087/grp1/simclr_pytorch_new/datasets/iNat64/lanczos',
          transform=transform_augmented,
      )
```

to the following:

```
inat_dataset_train = MixedINat128(
'./datasets/iNat64',
          transform=transform_augmented,
      )
```

4 Saving and Loading models

Throughout our code we have saved our models in custom paths in the UCL shared CS folders however we recommend changing them so that the models get saved and loaded into your folders.

We recommend modifying the *save_path* variable throughout the code to your desired directories.

Similarly when loading the model we call the variable *checkpoint_path* please change this to the *save_path* you saved the model in.

5 Running all our code

After modifying the variables in which our data and models are saved and stored everything can be run with *runner.py*. Type the following into terminal:

```
python runner.py
```