**Introduction to Deep Learning (Fall 2019)**

**Project (150 Pts, Dec 15)**

**Student ID**

**Name**

**Instruction:** We provide two datasets for three classification tasks and a baseline code in Python. Please implement the code to improve the accuracy on the three classification tasks (‘1\_imbalanced’, ‘2\_semisupervised’, ‘3\_noisy’). Compress all ‘the codes you implemented’ & ‘6 test outputs’ & ‘project.docx (this documentation)’ and submit with the filename ‘PROJECT\_STUDENT\_ID.zip’.

**Dataset Instruction**: We provide two datasets – CIFAR, MNIST. The number of classes in CIFAR and MNIST is 100 and 10, respectively. You can check the shape and number of train / valid datasets in the code. You can download the data in the link below.

<https://drive.google.com/drive/folders/1GHmbZpuJ_4AdPx4P9JCAmr8SZqUfYJBV?usp=sharing>

* CIFAR (# class = 100)
* MNIST (# class = 10)

**Task Instruction**. You need to implement the code to find the best accuracy for three classification task. You can implement different models for each clasisifcation task. It would welcome to see some related papers.

* **Imbalanced data classification**: The number of instances of a few classes is different from that of other classes significantly, thus leading to class imbalance.
* **Semi-supervised data classification**: Only a few number of images has labels while others do not have labels.
* **Noisy data classification**: Some labels in the dataset are noisy.

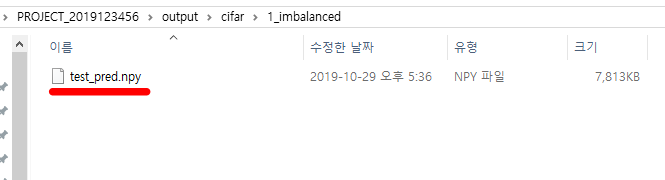
**How to load the data for each task**. These are the options for data and task you want to load in the left-hand side. For example, if you want to load the ‘*cifar*’ for the task ‘*2\_semisupervised*’, then you can refer to an example in the right-hand side.

|  |  |
| --- | --- |
| data = {cifar, mnist}  task = {1\_imbalanced, 2\_semisupervised, 3\_noisy} | data = ‘cifar’  task = ‘2\_semisupervised’ |

**Test outputs**. You must submit 6 test outputs named ‘test\_pred.npy’ under the output directory as below. The type of output file should be ‘numpy array’ and the shape must be (# classes, ). Please match the directory format as below.

(e.g.,) MNIST output example:

(e.g.,) Directory: /output/data(e.g., cifar)/task(e.g., 1\_semisupervised)/test\_pred.npy



1. **[50 pts] Imbalanced data classification (cifar, mnist).**
2. **[30 pts]** Implement the code and explain the key novelty of your model in detail (max 1 page).
3. **[20 pts]** Submittwo test outputs (‘test\_pred.npy’) for CIFAR and MNIST, respectively. Report all the ‘hyperparameters’ you used and the best valid accuracy.

**Answer: Fill the blank in the table.**

|  |  |  |
| --- | --- | --- |
| **Hyperaparmeters** | **Valid Accuracy** | **Test Accuracy** |
|  |  | **Do not write here** |

1. **[50 pts] Semi-supervised data classification.**
   1. **[30 pts]** Implement the code and explain the key novelty of your model in detail (max 1 page).
   2. **[20 pts]** Submittwo test outputs (‘test\_pred.npy’) for CIFAR and MNIST, respectively. Report all the ‘hyperparameters’ you used and the best valid accuracy.

**Answer: Fill the blank in the table.**

|  |  |  |
| --- | --- | --- |
| **Hyperaparmeters** | **Valid Accuracy** | **Test Accuracy** |
|  |  | **Do not write here** |

1. **[50 pts**] **Noisy data classification**.
   1. **[30 pts]** Implement the code and explain the key novelty of your model in detail (max 1 page).
   2. **[20 pts]** Submittwo test outputs (‘test\_pred.npy’) for CIFAR and MNIST, respectively. Report all the ‘hyperparameters’ you used and the best valid accuracy.

**Answer: Fill the blank in the table.**

|  |  |  |
| --- | --- | --- |
| **Hyperaparmeters** | **Valid Accuracy** | **Test Accuracy** |
|  |  | **Do not write here** |