**Assignment #00100 Report**

**Introduction**

In this assignment, I will implement customized Batch Normalization to MLP and CNN models with starter codes provided by Ankur Mali [1]. The dataset used for both models is Fashion MNIST. I will compare the performance of pre-activation and post-activation implementations on both models. Therefore, there will be 4 results in the end. The detailed code can be viewed at my Github: https://github.com/Yuxuan-Liu-Eason/IST-597-01000

**Pre-activation**

**MLP**

After running 5 trials with 10 epochs, the test accuracies are as follows. The optimizer is SGD with lr = 0.1. The batch size is 100.

|  |  |
| --- | --- |
| trial | Test acc |
| 1 | 0.8624 |
| 2 | 0.8730 |
| 3 | 0.8710 |
| 4 | 0.8718 |
| 5 | 0.8681 |

The mean accuracy is 0.8693.

**CNN**

The optimizer is Adams with lr = 0.01. The batch size is 128.

|  |  |
| --- | --- |
| trial | Test acc |
| 1 | 0.9075 |
| 2 | 0.9087 |
| 3 | 0.9068 |
| 4 | 0.9065 |
| 5 | 0.9053 |

The mean accuracy is 0.9069. We can see that CNN has better performance than MLP.

**Post-activation**

**MLP**

The optimizer is SGD with lr = 0.1. The batch size is 100.

|  |  |
| --- | --- |
| trial | Test acc |
| 1 | 0.8691 |
| 2 | 0.8585 |
| 3 | 0.8650 |
| 4 | 0.8680 |
| 5 | 0.8714 |

The mean accuracy is 0.8664, which is slightly lower than pre-activation.

**CNN**

The optimizer is Adams with lr = 0.01. The batch size is 128.

|  |  |
| --- | --- |
| trial | Test acc |
| 1 | 0.8884 |
| 2 | 0.8939 |
| 3 | 0.9076 |
| 4 | 0.8970 |
| 5 | 0.9033 |

The mean accuracy is 0.8983, which is also lower than pre-activation.

**Conclusion**

In this assignment, I implemented different batch normalization functions during training and testing. It is not surprising that CNN has better performance than MLP of the FMNIST dataset. Moreover, I found that pre-activation batch normalization has a slightly better performance.