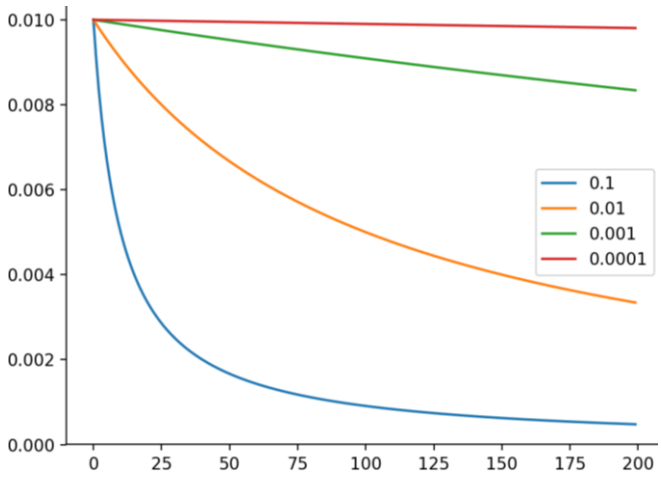
AI42001 – MLFA

1. Download FMNIST dataset from torchvision.
2. Create 3 separate models having the following configuration:
    a. 784-256-10
    b. 784-203-203-10
    c. 784-176-176-176-10
3. Train each model with the following learning rates (use SGD with momentum=0.9) with batch size set to 64:
    a. 0.0001
    b. 0.001
    c. 0.01
    d. 0.1
    e. 1
4. Now fix the learning rate at 0.01 and try the following different batch sizes:
    a. 16
    b. 64
    c. 256
    d. 1024
    e. 2048
5. For each experiment in point 3, plot the **training loss** vs epochs graph. Only one graph should be generated per model containing 5 different plots with corresponding learning rates clearly labelled.



6. For, point 4 report the **validation accuracy** for each model in a tabular form.

Models	Batch size1	Batch size2	Batch size3	Batch size4	Batch size5
Model1	88.35	88.98	87.73	87.03	85.51
Model2	88.43	89.71	89.02	87.12	85.74
Model3	85.51	85.51	85.51	85.51	85.51

7. Finally, increase the swap the model in 2(a) with the model 784-512-10. Use learning rate 0.01 with batch size of 64 and train this model. Report the **validation accuracy** of current model as well as the validation accuracy of the model in 2(a) trained using the same learning rate and batch size. 88.86 and 88.98
8. Now answer the following questions based on your experiments:
    a. How does increasing the learning rate affect the training loss?
    b. How does increasing the batch size affect validation accuracy?
    c. How does increasing depth affect validation accuracy?
    d. How does increasing the number of parameters affect validation accuracy?
9. Bonus round: Can you answer the **why** behind each of the questions asked above?