* We are implementing a simple neural network of 3 layer. The input layer consists of 784 nodes, the “hidden” layer consists of 30 nodes, and the output layer of 10 nodes, i.e the number of classes.
* For the plotting of the graph, three different learning rates were used: 0.001, 0.01, and 0.1.

Lr = 0.001, epoch 1 accuracy = 10.503%, epoch 2 accuracy = 16.939%, time taken = 55.16 seconds.

Lr = 0.01, epoch 1 accuracy = 46.550%, epoch 2 accuracy = 53.449%, time taken = 60.69 seconds.

Lr = 0.1, epoch 1 accuracy = 85.090%, epoch 2 accuracy = 87.423%, time taken = 82.562 seconds.

* On testing, the accuracy produced was 87.98%, roughly equal to the accuracy achieved by lr = 0.1 during training, and the time taken was 8.07 seconds.
* We are using the cross-entropy loss function because it is used for classifier errors.
* The number of epochs is kept small, and the data was accurate, and learning rate of 0.1 was converging to the labels quickly. If we increase the learning rate considerably beyond this, the model will diverge instead of converging.
* Directory structure:
* 