

## Learning Systems & Deep Learning Lab (31245)

### Multi-Layer Perceptron (MLP) Training in Keras with Stochastic Gradient Descent (SGD)

In this lab we will train an MLP to classify items from 10 categories in the Fashion MNIST dataset, which includes 60,000 training images (28x28 pixels) and 10,000 test images:



1. Train the given network three times, each with a different learning rate value: 0.1, 0.01, 0.001. Write down the achieved test set accuracy for each of the learning rates. Which one is best?
2. Use the best learning rate from section (1) and plot a graph of the test set accuracy as a function of the hidden layer size (number of perceptrons) in the range of  $2^k$ , with

$k=4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14$ . Which value of  $2^k$  provided the best results?

3. Use the best learning rate from section (1) and number of perceptrons = 512. Plot a graph of the test set accuracy as a function of the mini-batch size (number of samples per mini-batch) in the range of  $2^k$ ,  $k=3, 4, 5, 6, 7, 8$ . Which value of  $2^k$  provided the best results?
4. Use the best learning rate from section (1), number of perceptrons = 512, and mini-batch-size=32. Plot a graph of the test set accuracy as a function of the number of epochs, in the range of 10, 11, ..., 29, 30.
5. Repeat (1+2) with 2 and 3 hidden layers (use the same number of perceptrons for all layers).