Laboratory 3 – Exercises

1. Estimate the following linear model $y = 1 \cdot X - 1 + \epsilon$, where $X \in \mathbb{R}^{60\times 1}$ consists of generated values from -3 to 3 and ϵ obeys a normal distribution with mean 0 and standard deviation of 0.1. Train the linear regression model as follows:

➢ set the weights and bias to 1;

➢ set the batch size to 10;

➢ **implement** the mean squared error loss;

➢ **implement** the mini-batch gradient descent optimizer, $lr = 0.05$ (do not use the torch.optim package);

➢ set the total number of iterations to 25;

After the model is trained, compare the learned model parameters with the actual parameters.

2. Implement the softmax regression model in order to classify the images from the MNIST dataset (28×28 images of handwritten digits, 10 classes, labels are the digits from 0 to 9, 60000 training images and 10000 testing images) as follows:

➢ set the batch size to 100;

➢ the validation dataset consists of 30000 images and the rest represents the training dataset;

➢ set the weights to 1;

➢ use the cross-entropy loss function;

➢ use mini-batch stochastic gradient descent with a learning rate of 0.15 as the optimization algorithm;

➢ set the number of epochs to 15;

➢ plot the accuracy for the training and validation sets;

➢ use the trained model to classify 20 images.

Evaluate the trained model on the test set.

3. Classify the MNIST dataset (28×28 images, 10 classes, 60000 training images and 10000 testing images) using an MLP as follows:

➢ two hidden layers with 256 units;

➢ add the hyperbolic tangent activation function and a dropout layer after each hidden layer (dropout probability of 40%);

➢ use 45000 training images and 15000 validation images;

➢ use the cross-entropy loss function;

➢ apply Xavier initialization to the weights and set the biases to zero;

➢ set the batch size to 512 and the learning rate to 0.15;
➢ train the model for 10 epochs.
Evaluate the trained model on the test set.