Homework 2: Chapter 11 Training Deep Neural Network

Do Exercise 11.8 of the textbook.

Practice training a deep neural network on the `CIFAR10` image dataset:

- Build a DNN with 20 hidden layers of 100 neurons each (that's too many, but it's the point of this exercise). Use He initialization and the ELU activation function.
- Using Nadam optimization and early stopping, train the network on the `CIFAR10` dataset. You can load it with `keras.datasets.cifar10.load_data()`. The dataset is composed of 60,000 32×32–pixel color images (50,000 for training, 10,000 for testing) with 10 classes, so you'll need a softmax output layer with 10 neurons. Remember to search for the right learning rate each time you change the model's architecture or hyperparameters.
- Now try adding Batch Normalization and compare the learning curves: Is it converging faster than before? Does it produce a better model? How does it affect training speed?
- Try replacing Batch Normalization with `SELU`, and make the necessary adjustements to ensure the network self-normalizes (i.e., standardize the input features, use LeCun normal initialization, make sure the DNN contains only a sequence of dense layers, etc.).
- Try regularizing the model with alpha dropout. Then, without retraining your model, see if you can achieve better accuracy using MC Dropout.
- Retrain your model using 1cycle scheduling and see if it improves training speed and model accuracy.

Deadline: [Nov 15, 2025]