Heinrich Heine Universität Düsseldorf

HEIDERLAB

Exercise 10 on Machine Learning WS 2023/24
Prof. Dr. Dominik Heider
M. Sc. Arsam (Mohammad) Tajabadi
Submission: Until 17.01.2024, 23:59 on Ilias

Note:
Solve tasks 1-3 by hand (i.e., not programmatically!). You may use a calculator, but you have to write down the math.
You can work on the bonus task using any framework in Python. For example, Keras/Tensorflow or PyTorch. The bonus task is optional. However, you will only receive bonus points if your program code runs without errors! Submit only your program code. To ensure reproducibility, you can set seeds in your code.

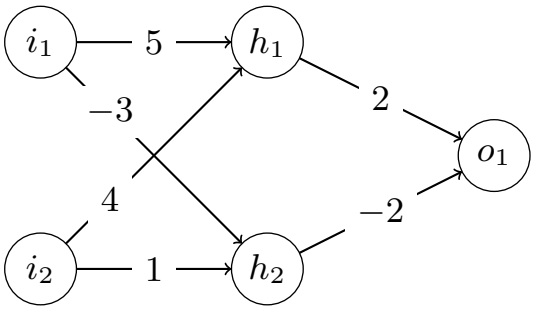
Task 1. McCulloch/Pitts Networks (2 points)

a) Construct a McCulloch/Pitts network that simulates the NOR function. (1 P.)

b) How is a McCulloch/Pitts neuron defined? (1 P.)

Task 2. Backpropagation (6 points)

Given the following neural network:



a) Calculate the output for the following inputs $i_1 = 5$ and $i_2 = 2$. The output function is the identity function. (0.5 P.)

b) Calculate the error (MSE) for the input from part a). The actual output for the input is 100 (0.5 P.)

c) Apply the Backpropagation algorithm once and update all weights of the network. Use a learning rate of 0.0005. Calculate the output and error again for the input from part a). (4 P.)

d) How would your result change if you increased or decreased the learning rate by a factor of 10? (0.5 P.)

e) What would happen if you initialized all weights with 0? (0.5 P.)

Task 3. CNN (2 points)

Given the following matrix:

$$\begin{bmatrix} 5 & 9 & 5 & 3 & 5 \\ 4 & 2 & 4 & 1 & 7 \\ 6 & 4 & 2 & 0 & 3 \\ 8 & 3 & 6 & 9 & 5 \\ 8 & 3 & 6 & 4 & 9 \end{bmatrix}$$

a) Calculate the output for the 3x3 Vertical Line Filter with a Stride of 1.

-1	1	-1
-1	1	-1
-1	1	-1

i. Without padding (0.5 P.)

ii. With Padding (0.5 P.)

b) Pool your results from part a). Use Minimum Pooling (3x3) with a Stride of 2. (1 P.)

Bonus Task (5 Bonus Points)

For this task, use the MNIST dataset. The dataset consists of 60,000 training data and 10,000 test data of handwritten digits (0-9). The grayscale images have dimensions of 28x28.

a) Visualize 2 different examples for each digit.

b) Training:

i. Train a fully connected neural network with one hidden layer and 10 output neurons. Use Softmax as the activation function for the output layer.

ii. Train a Convolutional Neural Network (CNN) with one Conv2d layer and one pooling layer. The output should also consist of 10 neurons with Softmax activation.

Use the Adam optimizer for training, sparse categorical crossentropy as the loss function, and a batch size of 64. Train each model for only 5 epochs. Which of your models performs better?

c) Optimize your models. You can change any parameter except the number of epochs. Experiment with Batch Normalization, Dropout layers, etc. Visualize the architecture of your best model.

d) Why should you normalize the input of neural networks?