TECHNISCHE UNIVERSITÄT DRESDEN

Professur für Informationstheorie und maschinelles Lernen, Fakultät Elektrotechnik und Informationstechnik

Prof. Rafael Schaefer, Dr. Martin Mittelbach, Muah Kim, Dr. Rick Fritschek   30.06.2023

5. Exercise of the Course
Einführung in das maschinelle Lernen

Remark: In addition to the problems below we will discuss Problem 13 of the last exercise.

Problem 14: (Convolution and downsampling)

Let X and K be two 2-dimensional arrays of size $m_x \times n_x$ and $m_k \times n_k$, respectively, where $m_x > m_k$ and $n_x > n_k$. We denote by

$$X * K$$

the convolution of X and K.

a) What is the size of the array $X * K$?

b) Implement "from scratch" a function in PYTHON, which calculates the convolution of two arrays.

c) Explicitly calculate (by hand or with the implemented PYTHON function) the convolution $X * K_j$, when X and K_j are given by

$$X = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \quad K_1 = \begin{pmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{pmatrix}, \quad K_2 = \begin{pmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{pmatrix}$$

d) Implement "from scratch" a function in PYTHON, which downsamples an array by a $p \times p$ max-pooling operation.

e) Apply 2×2 max-pooling to the results of Task c).

f) Load the images `square.png`, `circle.png`, `triangle.png`, and `octagon.png` and calculate the convolutions of the images with the arrays K_1 and K_2 of Task c). Illustrate the convolution results.

Hint: You can use

```
from PIL import Image
image = Image.open("filename")
```

to load the images.

g) Apply $p \times p$ max-pooling to the convolution results of Task f) for $p = 2, 4, 16$ and illustrate the results. What is the array size after downsampling?

Problem 15: (CNNs, convolutional layers)

Let the input of a CNN be given by RGB color images of size 256×256 pixels.

The CNN architecture has the following specifications:

- convolutional layer 1: 5×5 kernel, 6 output channels
- convolutional layer 2: 5×5 kernel, 12 output channels
- convolutional layer outputs are subject to 2×2 max-pooling

a) What is the number of input nodes of a linear layer that is connected to the second max-pooling layer?

b) What is the number of trainable parameters in convolutional layer 1 and convolutional layer 2?

c) What is the total number of trainable parameters of the CNN considered in Problem 13?