Machine learning, exercise sheet 7

DNN

Consider a DNN consisting of the followng layers (input: 5 elements)

1. affine with 3 outputs
2. ReLU
3. affine affine with 10 outputs
4. Softmax

What are output sizes for all layers? outputs? What are the shapes of the matrices

$$W^{(X)}$$

and bias vectors

$$\vec{b}^{(X)}$$

for the affine layers?

CNN 1

Assume that a CNN processes data samples of size (32,32,3). Furthermore assume the following structure for the first 6 layers:

- Conv2D with filter size 3x3, strides 1x1 and 32 filters
- ReLU
- MaxPooling2D with kernel size 2x2
- Conv2D with filter size 4x4, strides 1x1 and 64 filters
- ReLU
- MaxPooling2D with kernel size 2x2
- Softmax

Please give the dimensions of all layers as tuples of (H, W, C)!

CNN 2

Please mark the following statements with true or false, no justification required!

- All filters in a Conv2D-layer must be identical
- The number of filters in a conv2D-layer can be chosen arbitrarily
- The width and height of a conv2D-layer output can be chosen arbitrarily
- A max-pooling layer must have a kernel size of 2x2
- A Conv2D-layer must be followed by ReLU
- A CNN classifier must contain at least one affine layer
- A CNN classifier can contain an arbitrary number of Conv2D-layers
- Conv2D layers are special cases of affine layers

CNN 3

A Conv2D layer X receives input $A^{(X-1)}$ of dimensions $H^{(X-1)} = 28$, $W^{(X-1)} = 28$ and $C^{(X-1)} = 10$.
a) Assuming a filter size $f_X = f_Y = 4$ and strides $\Delta_X = \Delta_Y = 2$: compute the width and height of $A^{(X)}$.
b) How can you compute $C^{(X)}$ from the given facts?
c) Please describe a way to obtain $C^{(X)} = 256$.

CNN 4

A Conv2D layer X receives input $A^{(X-1)}$ of dimensions $H^{(X-1)} = 28$, $W^{(X-1)} = 28$ and $C^{(X-1)} = 10$. How do you need to set filter sizes, strides and nr of filters to obtain $H^{(X)} = W^{(X)} = 24$ and $C^{(X)} = 10$?

Programming practice 2

Have a peek at the TF/Keras slides and implement the ReLU and softmax functions in low-level TF2! They should work with 2D tensors of arbitrary shape. Test them with the two tensors

$$X_1 = \begin{pmatrix} 1 & 2 & -3 \\ 0 & 0 & 1 \end{pmatrix}$$

and

$$X_2 = \begin{pmatrix} 1 \\ -2 \end{pmatrix}$$

.