Architektur Neuronaler Netze für Generative KI

Vertiefungsaufgaben

Hochschule Worms • Fachbereich Informatik
Prof. Dr. Stephan Kurpjuweit

Hochschule Worms
University of Applied Sciences

Aufgabe 1

- In logistic regression given $\mathbf{x}$ and parameters $w \in \mathbb{R}^{n_x}, b \in \mathbb{R}$. Which of the following best expresses what we want $\hat{y}$ to tell us?

$$\boxed{P(y = 1|\mathbf{x})}$$

$$\sigma(W\,\mathbf{x})$$

$$P(y = \hat{y}|\mathbf{x})$$

$$\sigma(W\,\mathbf{x} + b)$$

Aufgabe 2

Consider the Numpy array x:

$$x = np.array([[[1], [2]], [[3], [4]]])$$

What is the shape of x?

$\boxed{(2,2,1)}$

(4,)

(2, 2)

(1, 2, 2)

Aufgabe 3

What does a neuron compute?

A neuron computes a linear function $z = Wx + b$ followed by an activation function

A neuron computes an activation function followed by a linear function $z = Wx + b$

A neuron computes the mean of all features before applying the output to an activation function

A neuron computes a function g that scales the input x linearly (Wx + b)

Aufgabe 4

Consider the following random arrays a and b, and c:

$$a = np.random.randn(2, 3) \, \# \, a.shape = (2, 3)$$

$$b = np.random.randn(2, 1) \, \# \, b.shape = (2, 1)$$

$$c = a + b$$

What will be the shape of c?

The computation cannot happen because the sizes don't match. It's going to be "Error"!

c.shape = (3, 2)

c.shape = (2, 3)

c.shape = (2, 1)

Aufgabe 5

Consider the two following random arrays a and b:

$$a = np.random.randn(4, 3)\ \#\ a.shape = (4, 3)$$

$$b = np.random.randn(1, 3)\ \#\ b.shape = (1, 3)$$

$$c = a * b$$

What will be the shape of c?

- c.shape = (4, 3)
- c.shape = (1, 3)
- The computation cannot happen because the sizes don't match.
- The computation cannot happen because it is not possible to broadcast more than one dimension.

Aufgabe 6

Suppose you have n_x input features per example. Recall that $X = [x^{(1)} x^{(2)} ... x^{(m)}]$. What is the dimension of X?

$$\boxed{(n_x, m)}$$

Aufgabe 7

Recall that $np.dot(a, b)$ performs a matrix multiplication on a and b, whereas $a * b$ performs an element-wise multiplication.

Consider the two following random arrays a and b:

$a = np.random.randn(12288, 150)$

$a.shape = (12288, 150)$

$b = np.random.randn(150, 45)$

$b.shape = (150, 45)$

$c = np.dot(a, b)$

What is the shape of c?

$$\boxed{\text{c.shape = (12288, 45)}}$$

Aufgabe 8

Consider the following code snippet:

$$a.shape = (3, 4)$$

$$b.shape = (4, 1)$$

```
for i in range(3):
 for j in range(4):
 c[i][j] = a[i][j] + b[j]
```

How do you vectorize this?

c = a.T + b

c = a + b

$\boxed{\text{c = a + b.T}}$

c = a.T + b.T

Aufgabe 9

Consider the following arrays:

$$a = np.array([[1, 1], [1, -1]])$$

$$b = np.array([[2], [3]])$$

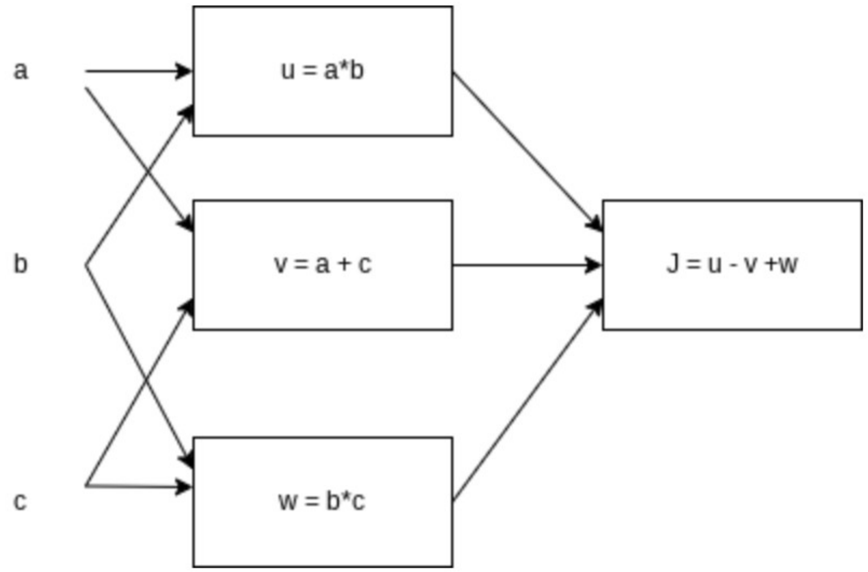
$$c = a + b$$

Which of the following arrays is stored in c?

$$\boxed{\begin{matrix} 3 & 3 \\ 4 & 2 \end{matrix}}$$

Aufgabe 10

Consider the following computational graph.



What is the output of J?

$(c-1)(a+c)$

$(a-1)(b+c)$

$ab+bc+ac$

$\boxed{(a+c)(b-1)}$

Aufgabe 11

Suppose our input batch consists of 8 grayscale images, each of dimension 8x8. We reshape these images into feature column vectors $\mathbf{x}^j$. Remember that $X = \left[\mathbf{x}^{(1)}\mathbf{x}^{(2)}\cdots\mathbf{x}^{(8)}\right]$. What is the dimension of X?

(8, 64)

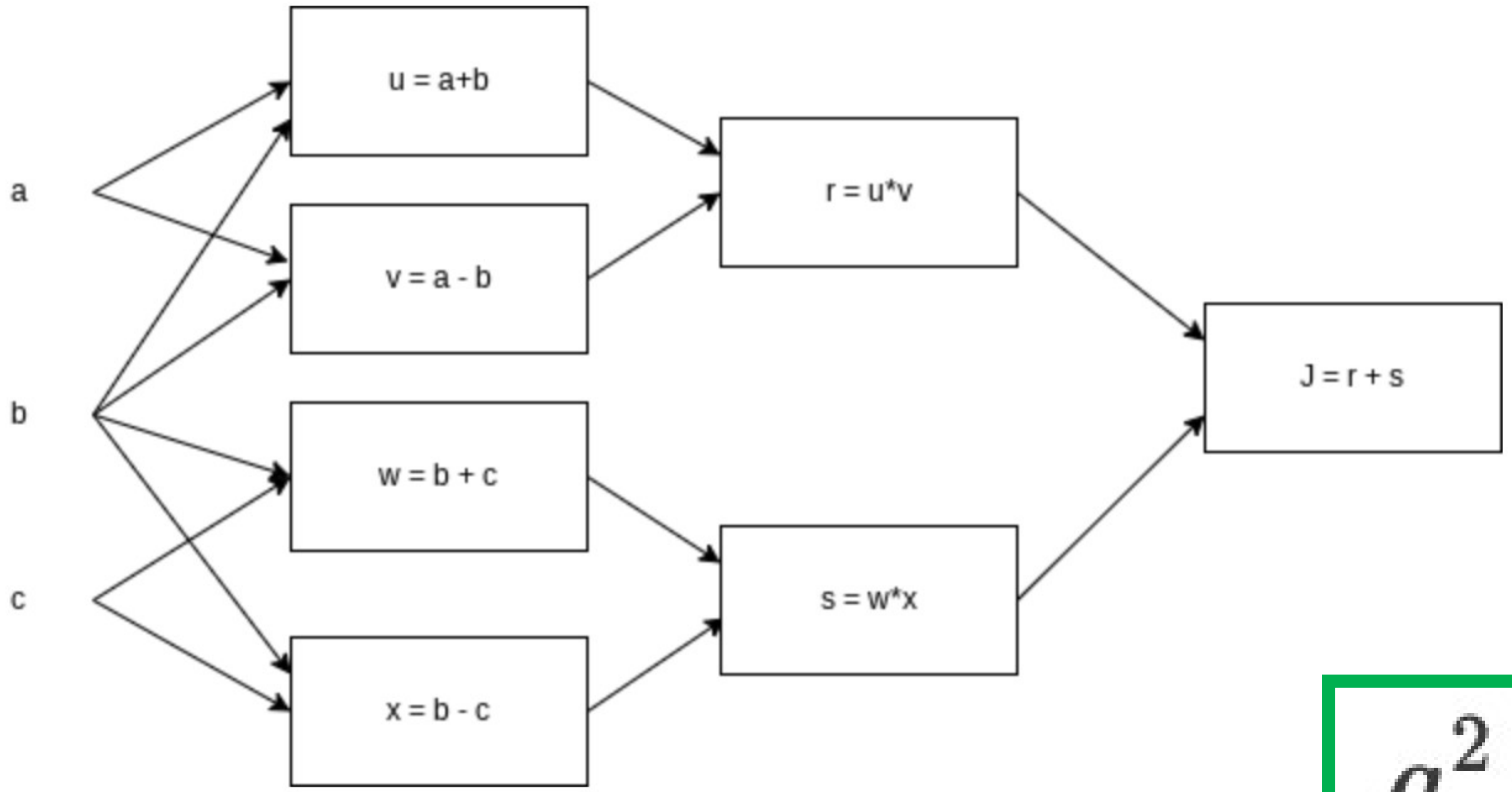
(512, 1)

$\boxed{(64, 8)}$

(8, 8, 8)

Aufgabe 12

ⁱ. Consider the following computational graph.



$$\boxed{a^2 - c^2}$$

What is the output of J?

Aufgabe 13

Consider the following array:

$$a = np.array([[2, 1], [1, 3]])$$

What is the result of $np.dot(a, a)$?

$$\boxed{\begin{pmatrix} 5 & 5 \\ 5 & 10 \end{pmatrix}}$$

Aufgabe 14

Which of the following are true? (Check all that apply.)

$a^{[2]}$ denotes the activation vector of the second layer.

$w_3^{[4]}$ is the row vector of parameters of the fourth layer and third neuron.

$w_3^{[4]}$ is the column vector of parameters of the fourth layer and third neuron.

$a_3^{[2]}$ denotes the activation vector of the second layer for the third example.

$a^{[3](2)}$ denotes the activation vector of the second layer for the third example.

$w_3^{[4]}$ is the column vector of parameters of the third layer and fourth neuron.

Aufgabe 15

Which of these is a correct vectorized implementation of forward propagation for layer l, where $1 \leq l \leq L$?

$$Z^{[l]} = W^{[l]} A^{[l]} + b^{[l]}$$
$$A^{[l+1]} = g^{[l+1]}(Z^{[l]})$$

$$\boxed{\begin{aligned} Z^{[l]} &= W^{[l]} A^{[l-1]} + b^{[l]} \\ A^{[l]} &= g^{[l]}(Z^{[l]}) \end{aligned}}$$

$$Z^{[l]} = W^{[l]} A^{[l]} + b^{[l]}$$
$$A^{[l+1]} = g^{[l]}(Z^{[l]})$$

$$Z^{[l]} = W^{[l-1]} A^{[l]} + b^{[l-1]}$$
$$A^{[l]} = g^{[l]}(Z^{[l]})$$

Aufgabe 16

The use of the ReLU activation function is becoming more rare because the ReLU function has no derivative for $c = 0$. True/False?

$\boxed{\text{False}}$

Aufgabe 17

Consider the following code:

```
#+begin_src python

x = np.random.rand(4, 5)

y = np.sum(x, axis=1)

#+end_src
```

What will be y.shape?

(4, 1)

$\boxed{(4,)}$

(1, 5)

(5,)

Yes. By using axis=1 the sum is computed over each row of the array, thus the resulting array is a column vector with 4 entries. Since the option keepdims was not used the array doesn't keep the second dimension.

Aufgabe 18

Suppose you have built a neural network with one hidden layer and tanh as activation function for the hidden layer. You decide to initialize the weights to small random numbers and the biases to zero. The first hidden layer's neurons will perform different computations from each other even in the first iteration. True/False?

True Yes. Since the weights are most likely different, each neuron will do a different computation.

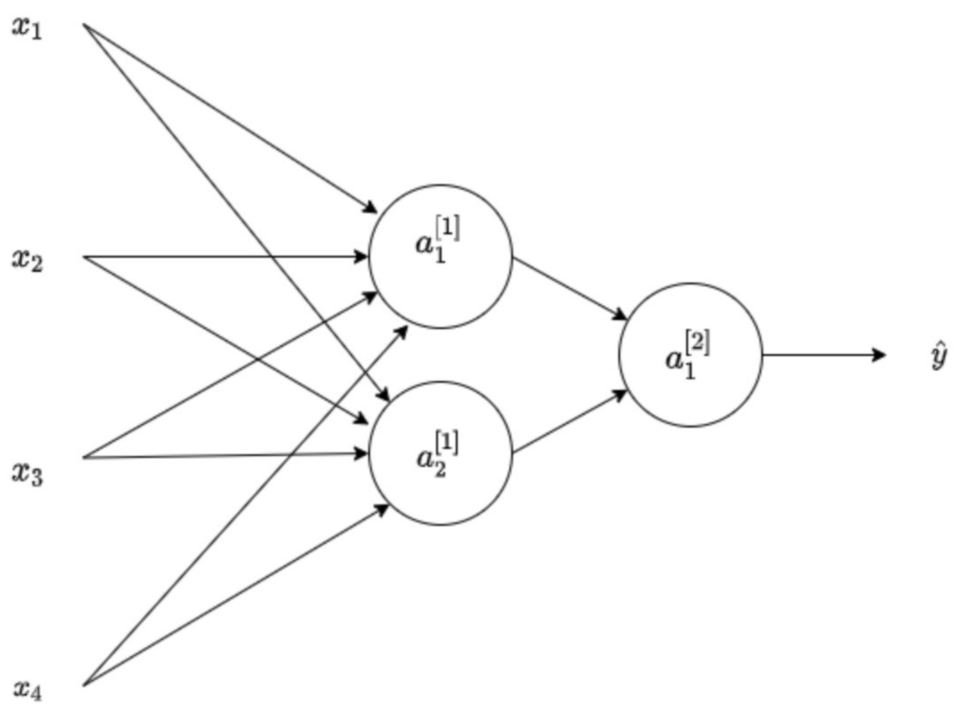
Aufgabe 19

A single output and single layer neural network that uses the sigmoid function as activation is equivalent to the logistic regression. True/False

$\boxed{\text{True}}$

Aufgabe 20

Consider the following 1 hidden layer neural network:

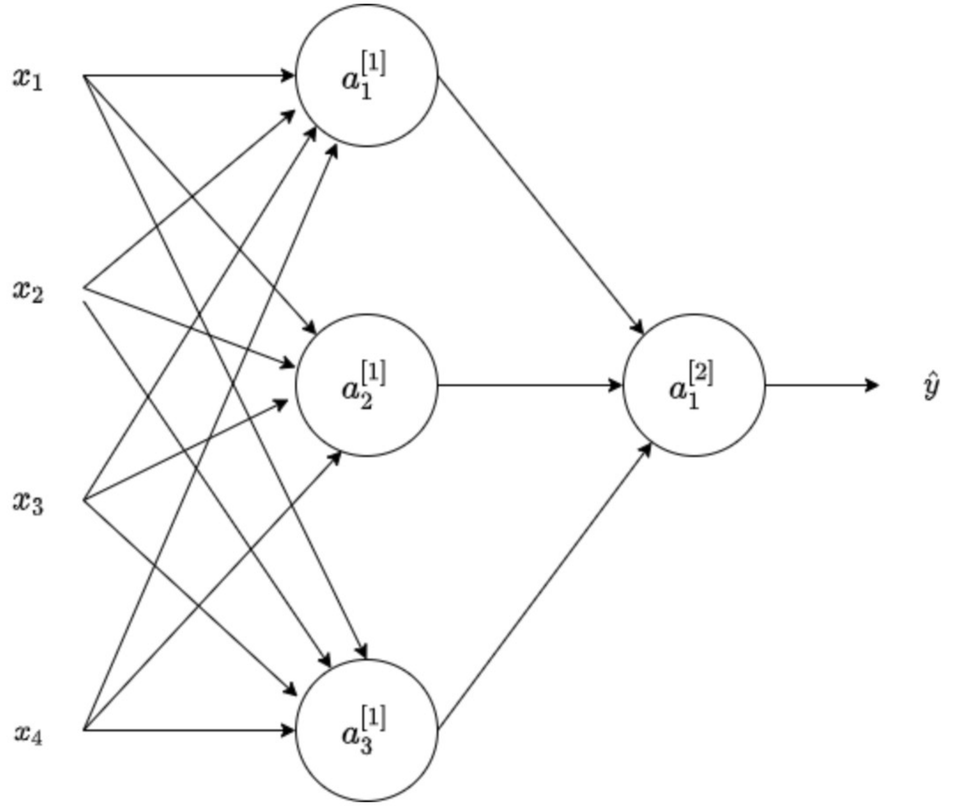


$Z^{[1]}$ and $A^{[1]}$ are (2, m)

What are the dimensions of $Z^{[1]}$ and $A^{[1]}$?

Aufgabe 21

Consider the following 1 hidden layer neural network:



Which of the following statements are True? (Check all that apply).

$W^{[1]}$ will have shape (4, 3).

$b^{[1]}$ will have shape (1, 3)

$b^{[2]}$ will have shape (3, 1)

$\boxed{W^{[1]} \text{ will have shape } (3, 4).}$

$\boxed{b^{[1]} \text{ will have shape } (3, 1).}$

$\boxed{b^{[2]} \text{ will have shape } (1,1)}$

Aufgabe 22

Which of the following are true about the tanh function?

$\boxed{\text{For large values the slope is close to zero.}}$

The derivative at $c = 0$ is not well defined.

$\boxed{\text{The tanh is mathematically a shifted version of the sigmoid function.}}$

For large values the slope is larger.

The slope is zero for negative values.

Aufgabe 23

Among the following, which ones are "hyperparameters"? (Check all that apply.)

weight matrices $W^{[l]}$

size of the hidden layers $n^{[l]}$

number of iterations

activation values $a^{[l]}$

number of layers L in the neural network

learning rate α

bias vectors $b^{[l]}$

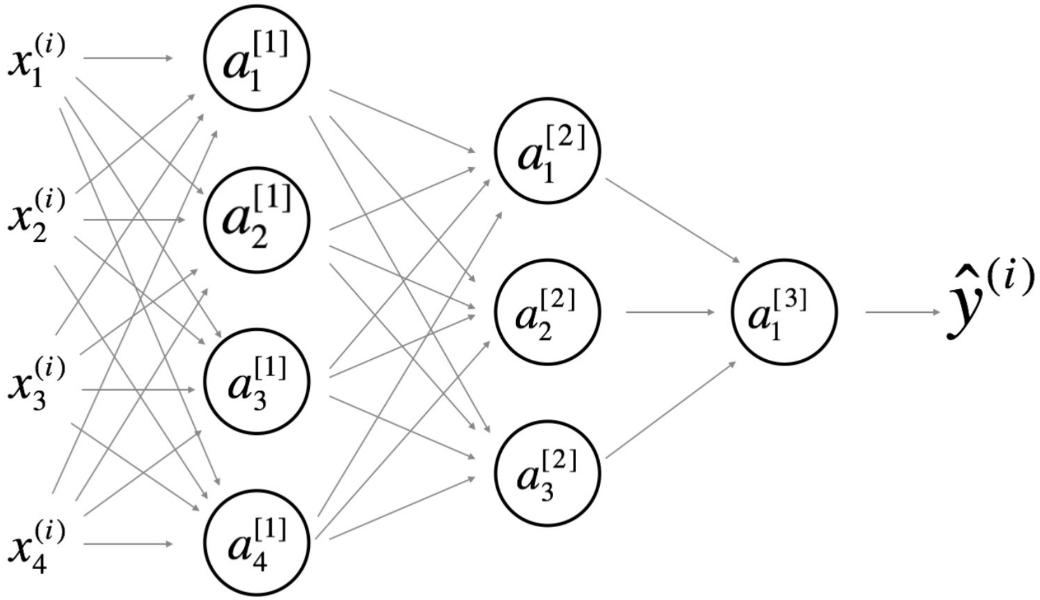
Aufgabe 24

- A shallow neural network with a single hidden layer and 6 hidden units can compute any function that a neural network with 2 hidden layers and 6 hidden units can compute. True/False?

$\boxed{\text{False}}$

Aufgabe 25

Consider the following 2 hidden layer neural network:



Which of the following statements are True? (Check all that apply).

$b^{[2]}$ will have shape (1, 1)

$W^{[3]}$ will have shape (1, 3)

$b^{[2]}$ will have shape (3, 1)

$W^{[2]}$ will have shape (3, 4)

$W^{[3]}$ will have shape (3, 1)

$b^{[3]}$ will have shape (3, 1)

$b^{[1]}$ will have shape (4, 1)

$W^{[1]}$ will have shape (4, 4)

$W^{[1]}$ will have shape (3, 4)

$b^{[3]}$ will have shape (1, 1)

$b^{[1]}$ will have shape (3, 1)

$W^{[2]}$ will have shape (3, 1)

Aufgabe 26

- Whereas the previous question used a specific network, in the general case what is the dimension of W^{[l]}, the weight matrix associated with layer l?

$$\boxed{W^{[l]} \text{ has shape } (n^{[l]}, n^{[l-1]})}$$