

# One Example Task in The Exam

**Instructions:**

- Please write your solutions in the **bold boxes** on the assignment sheets.
- Only pen and paper are allowed.
- **Good Luck!**

Name	
Matriculation number	
Field of study	

**Exercise 1: General Questions [6+4+5+5 points]**

- a) Explain the motivation of batch normalization and how it is performed (calculation of the normalization and specification of learning parameters).

- b) Calculate the output of the following logistic regression model when it is applied to the following test dataset. (Note that in the real exam, we will provide you a pocket calculator to solve the task).

Model:

$$w = \begin{pmatrix} 1 \\ 0 \\ -1 \\ 0.5 \end{pmatrix}, \quad b = -1$$

Dataset:

$$(x_1 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}, y_1 = 1), \quad (x_2 = \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \end{pmatrix}, y_2 = 0), \quad (x_3 = \begin{pmatrix} -1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, y_3 = 1),$$

$$(x_4 = \begin{pmatrix} 1 \\ 2 \\ 0 \\ 1 \end{pmatrix}, y_4 = 0), \quad (x_5 = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix}, y_5 = 0)$$

- c) You need to train a neural network model for a sentence classification task, e.g. dialog act classification. The training and the development data with labels are given.

Which neural network model will you use for this task? Why? Why do you want to train your network with different random initializations? Give reasons for your answer.

- d) Mark whether the following statements are true or false.

Statement	True	False
k-means clustering is typically used for supervised data.		
Convolutional neural networks are used only for image processing.		
Recurrent neural networks have the vanishing gradients problem.		
Parameters are chosen before the training to optimize the results on the training set.		
Training a neural network with a batch size of one is called batch training.		