Universität des Saarlandes
Prof. Dr. Dietrich Klakow
Lehrstuhl für Signalverarbeitung
NNTI Winter Term 2022/2023

Exercise Sheet 4

NNTI Tutorial

Deadline: 07.12.2022 08:00

Exercise 4.1 - Maximum Likelihood Estimate (MLE) (1+1 points)

a) Show how a linear regression procedure can be justified as an MLE procedure, assuming that mean squared error is used as a metric. Recall that

$$MSE = \frac{1}{m}\sum_{i=1}^{m} \left\| \hat{y}^{(i)} - y^{(i)} \right\|^2 .$$

Justify and motivate the assumptions you make along the way. This particular deduction is not covered in the lecture. Consult the book to gain further understanding.

b) Given an i.i.d. sample $X_1, \ldots, X_n$ from a Poisson distribution with parameter λ, find the MLE of the parameter λ. Recall that

$$\Pr(X = x|\lambda) = \frac{\lambda^x e^{-\lambda}}{x!} .$$

Exercise 4.2 - Supervised and Unsupervised Learning (0.5+0.5+1 points)

The following explanations should be kept concise (max. 100 words).

a) Describe some disadvantages of using mean squared error as a metric for a supervised regression task. How would the above answer change if the relevant task was a binary classification task?[1]

b) We have only covered PCA and briefly discussed other methods for unsupervised learning. Research other unsupervised methods and discuss how they differ from PCA and what problems they address.

c) In supervised learning, the *softmax* function can have a temperature parameter T, such that

$$softmax(\mathbf{x})_i = \frac{\exp(x_i/T)}{\sum_{j=1}^{n}\exp(x_j/T)} .$$

Explain what effect a high and a low temperature would have on the function. Do this by plotting the function with distinct T values. Finally, discuss when manipulating the temperature could be useful.

[1]Tasks which aim to assign a category $C \in \{0, 1\}$ to an example x_i.

The topic of temperature is not covered in the lecture and you are encouraged to do some research to reach an answer yourself.

Exercise 4.3 - Gradient Descent (2 points)

a) Let $f : \mathbb{R}^2 \to \mathbb{R}$ be the following real function:

$$f(\mathbf{x}) = \frac{1}{1 + \exp(-(x_1^2 + x_2^2))}$$

Perform gradient descent on f in order to minimize it. Use the following value as x_0

$$x_0 = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \tag{1}$$

Use a learning rate $\epsilon = 0.2$. Do 3 iterations or until the gradient becomes zero, whichever happens first. On each iteration make sure to show the following:

- The value of $\mathbf{x}$
- The value of $\nabla f(\mathbf{x})$
- The value of $f(\mathbf{x})$

Once finished, compare the values $f(\mathbf{x})$ and $f(\mathbf{x}_0)$.

What would happen if a learning rate of ϵ_2 or ϵ_3 were used? Where

$$\epsilon < \epsilon_2 \lll \epsilon_3 \, .$$

Finally, find an $\hat{\mathbf{x}}$ that minimizes f, by using gradient descent. You are free to choose your own learning rate; justify your decision. Is this $\hat{\mathbf{x}}$ guaranteed to be a global minimum given the nature of gradient descent and the function?

Exercise 4.4 - Feedforward Networks (2+2 points)

a) Implement your own feed-forward neural network from scratch. Check the attached `ipynb` file.

b) Implement your own feed-forward neural network using `pytorch`. Check the attached `ipynb` file.

Submission instructions

The following instructions are mandatory. If you are not following them, tutors can decide to not correct your exercise.

- Please submit the assignment as a **team of two to three** students.
- Write the Microsoft Teams user name, student id and the name of each member of your team on your submission.

- Hand in zip file containing a `single` PDF with your solutions and the completed ipython notebook. Do not include any data or cache files (e.g. __pycache__).
- Important: please name the submitted zip folder and files inside using the format: **Name1_id1_Name2_id2**.
- Your assignment solution must be uploaded by only **one** of your team members to the 'Assignments' tab of the tutorial team (in **Microsoft Teams**). Please remember to press the **Hand In** button after uploading your work.
- If you have any trouble with the submission, contact your tutor **before** the deadline.