Machine Assignment 1: Neural Networks

Luca Ambrosio : Machine Learning

Deadline: Friday 19 Nov 2021, 21:00

1 Exercise 1.1

If we assume a batch of input data in form of $\mathbf{X} \in R^{nxk}$, the dimension of the output vector $\mathbf{y}$ is n $\times 1$

2 Exercise 1.2

the vectorized equation for the output $\mathbf{y}$ =$\mathbf{X}$* $\mathbf{W}$ + $\mathbf{1}_n^T$ * b

3 Exercise 1.3

MSE = $\frac{1}{2n} * (\mathbf{y} - \mathbf{t})^T * (\mathbf{y} - \mathbf{t})$

4 Exercise 1.4

$\frac{\delta L}{\delta w_j} = \frac{\delta L}{\delta y_i} * \frac{\delta y_i}{\delta w_j}$

$\frac{\delta L}{\delta y_i} = \frac{(y_i - t_i)}{n}$

$\frac{\delta L}{\delta w_j} = x_i$

$\frac{\delta L}{\delta w_j} = \sum_{i=1}^{n} \frac{(y_i - t_i)}{n} * x_i$

5 Exercise 1.5

The equation of the weight update by gradient descent is $w_i^{new} = w_i^{old} - \eta \frac{\delta L}{\delta w_i}$

6 Exercise 1.6

After one step the weights are : [-0.06;-0,23;0,18]

7 Exercise 1.7

Gradient descent is an algorithm used to find the local minimum of a function with more variables and is used in machine learning to update the weights in order to minimize the error. The back-propagation is an algorithm that allows you to calculate the derivative of the error with respect to the weights using the chain rule methodology.