Neural Networks and Learning Theory

Prof. Dr. Klaus Meer, Ameen Naif

Exercise Sheet 1
Version 12.04.2019

Exercise 1.
Consider the following activation function (*the sigmoid function*):

$$\varphi(t) := \frac{1}{1 + e^{-c \cdot t}} \ ,$$

where c is any positive real constant.

(a) Compute $\varphi(0)$, $\lim_{t \to -\infty} \varphi(t)$ and $\lim_{t \to \infty} \varphi(t)$. Draw the graph of φ.

(b) How does the curve of $\varphi(t)$ change depending on c?

(c) Show that the derivative of $\varphi(t)$ with respect to t is given by: $c \cdot \varphi(t)[1 - \varphi(t)]$.

Exercise 2.
Consider yet another sigmoidal function:

$$\varphi(t) := \frac{t}{\sqrt{1 + t^2}} \ .$$

(a) Compute $\varphi(0)$, $\lim_{t \to -\infty} \varphi(t)$ and $\lim_{t \to \infty} \varphi(t)$. Draw the graph of φ.

(b) Show that the derivative of $\varphi(t)$ is given by: $\dfrac{\varphi^3(t)}{t^3}$.

Exercise 3.
(Hesse normal form) We consider in $\mathbb{R}^n$ subsets of the form

$$H := \{x \in \mathbb{R}^n | w^T \cdot x = d\} \tag{1}$$

with fixed $w \in \mathbb{R}^n, d \in \mathbb{R}$.

(a) Recall from linear algebra , what a hyperplane in $\mathbb{R}^n$ is and how to describe it algebraically as affine space of dimension $n - 1$. Prove then that sets of form (1) are precisely the hyperplanes in $\mathbb{R}^n$.

(b) If $\|w\|_2 = 1$ in equation (1) equals 1, then we obtain the *Hesse normal form* of H. In this case, give a geometric interpretation of the w and d and prove that your interpretation is correct.

Exercise 4.
Show that the **Perceptron** can solve the logical problems AND and OR (implement the logical functions AND and OR), but it cannot solve XOR. Find a $Feedforward-network$ for solving the XOR problem.

Exercise 5.

Consider the **McCulloch-Pitts** neuron with threshold parameter $\theta = 0$, weights $w_1, ..., w_m$ and the bias b, which classifies correctly two finite pointsets C_1 and C_2.

(a) Draw the graph of this network and explain how it work and how can it classifies a point $x \in \mathbb{R}^m$.

(b) Show that the weights $w_1, ..., w_m$ can be changed so that the bias gets the value 1, while all points from $C_1 \cup C_2$ still to be correctly classified. Start with the situation $b = 0$.