Institute for Computer Science VI, Autonomous Intelligent Systems, University of Bonn

Dr. N. Goerke
Friedrich-Hirzebruch-Allee 8, 53115 Bonn, Tel: +49 228 73-4167
E-Mail: goerke (at) ais.uni-bonn.de
http://www.ais.uni-bonn.de/WS2223/4204_L_NN.html

Exercises for module Technical Neural Networks (MA-INF 4204), WS22/23

Assignments Sheet 9, due: Monday 19.12.2022

12.12.2022

Group	Name	53	54	56	57	58	59	$\sum$ Sheet 9

Assignment 53 (3 Points)

Explain that the VC-dimension (Vapnik-Chervonenkis dimension, VC-capacity) for the task of binary classification and the family of linear separating lines in 2-dimensions is $h = 3$.
Use a series of diagrams to support your explanation.
Use your own words, do not take text or pictures from others (e.g. Wikipedia).

Assignment 54 (2 Points)

How large is the VC-Capacity h (VC-dimension) of a table with 1000 lines (entries) for the task of binary classification?
Explain your answer.

Assignment 55 (2 Points)

Define the task of binary classification as formal as possible.

Assignment 56 (3 Points)

Show in an analytical way, that the canonical form of a hyperplane in two dimensions, is defining a linear function. Draw a fully labled sketch to depict the resulting hyperplane; depict all parts of your formulas explicitly within the drawing.
$\mathbf{H} = \Big\{ \vec{\mathbf{Z}} \quad | \quad \langle \vec{\mathbf{w}}, \vec{\mathbf{Z}} \rangle + b = 0 \Big\}.$

Assignment 57 (2 Points)

A hyperplane $\mathbf{H}$ is given in its canonical form by:
$\mathbf{H} = \Big\{ \vec{\mathbf{Z}} \quad | \quad \langle \vec{\mathbf{w}}, \vec{\mathbf{Z}} \rangle + b = 0 \Big\}.$

How does a second hyperplane $\mathbf{G}$ differs from $\mathbf{H}$?
$\mathbf{G} = \Big\{ \vec{\mathbf{Z}} \quad | \quad \langle \vec{\mathbf{w}}, \vec{\mathbf{Z}} \rangle - b = 0 \Big\}.$

Assignment 58 (3 Points)

Create a *decision function* $f(\mathbf{z}) = sign\left((\mathbf{w} \cdot \mathbf{z}) + b\right)$ that is using a hyperplane defined by the scalar product $(\mathbf{w} \cdot \mathbf{z})$ to separate the point $x_s = (0.0, 1.0, 1.0)$ from the other points x_1 to x_7. Give the values of $\mathbf{w}$ and b and depict the hyperplane.
$x_1 = (0.0, 0.0, 0.0), x_2 = (0.0, 0.0, 1.0), x_3 = (0.0, 1.0, 0.0), x_4 = (1.0, 0.0, 0, 0)$
$x_5 = (1.0, 0.0, 1.0), x_6 = (1.0, 1.0, 0.0), x_7 = (1.0, 1.0, 1, 0)$

Assignment 59 (2 Points)

Imagine ypu have two linear separable sets of points from the N-dimensional input space. Each set is covered by a convex hull in N dimensions. Briefly describe a method to determine and calculate the smallest distance between these two convex hulls.

Programming assignment PA-F (5 Points, Due: Monday 19.12.2022)

Implement in Python a simple 2-1 MLP with 2 inputs x_1, x_2, one output neuron with hyperbolic tangent, and the three weights (w_0, w_1, w_2) (as usual the BIAS weight is w_0).

The objective for the network is to seperate the linear separable data points (`PA-F-training.txt`) from the 2 dim input space into two classes $\mathcal{A}$ and $\mathcal{B}$:
Set $\mathcal{A} = \{(0,3), (1,3), (1,2), (2,2), (2,4), (5,8)\}$
Set $\mathcal{B} = \{(0,-3), (1,-3), (1,-2), (2,-4), (3,1), (3,0), (3,-2), (4,-1), (5,1)\}$

Determine (and tell us) a set of weights (w_0, w_1, w_2) that is doing the job. The proposal to determine the weights is to use gradient descent via the the $\delta-$rule and the sum of the squarred differences as error function. Depict the data points together with the resulting separating decision function (a line in 2-dim).

Depict (as png or jpeg or pdf) the three (global) error surfaces that you obtain, when you fix one of the weights to the best value you have found, and vary the other two weights between -10.0 and $+10.0$.

Extra(no extra points)
Try other loss-functions when displaying the surfaces.
Try other loss-functions for learning the weights.
Try what happens, if one of the data points is *on the other side* of the separating line, if the data set is no longer linear separable.