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 3, due: Monday 7.11.2022

31.10.2022

Group	Name	14	15	16	17	18	19	$\sum$ Sheet 3

Assignment 14 (2 Points)

A simple Perzeptron with two inputs x_1, x_2, BIAS, and transferfunction $y = f(z) = sgn(z)$, separates the two dimensional input space into two parts with help of a line g.

Calculate for this Perzeptron the weights w_1, w_2, w_b in such a way, that the line g separates the given $P = 8$ patterns (${}^px_1$, ${}^px_2$; ${}^p\hat{y}$) into two classes.
Depict the given patterns and the resulting line g.

${}^1X = (0, 0; +1)$, ${}^2X = (2, 3; -1)$,
${}^3X = (-1 - \varepsilon, 3 - \varepsilon; +1)$, ${}^4X = (-1 + \varepsilon, 3 + \varepsilon; -1)$,
${}^5X = (-1, 2; +1)$, ${}^6X = (3, 1; -1), ;$
${}^7X = (5, -2 + \varepsilon; -1)$, ${}^8X = (5 - \varepsilon, -2; +1)$
Remark: $0 < \varepsilon \ll 1$.

Assignment 15 (2 Points)

Discuss advantages and disadvantes in taking the *ramp function*, *rectified linear unit: relu* function as transfer function for an MLP compared to taking the *logistic function* or the *hyperbolic tangent*. Typical aspects could be: complexity to implement the function and the derivative, saturation effects for large positive and large negative values, determining the gradient, small or even zero gradients.

Assignment 16 (3 Points)

Apply the learning rule Backpropagation of Error to calculate the weight change ΔW_{nh} for a weight w_{nh} at a hidden layer neuron h in a 3 layer N-H-1MLP. The transfer function in hidden and output layer is the hyperbolic tangent (tanh). The learning rate is $\eta = 0.1$, the teacher value is $\hat{y}_{m=1} = 0.2$, the output value is $y_{m=1} = 0.8$, the weight at the output neuron coming from hidden neuron h is $w_{hm} = 2.5$, the output of the hidden neuron h is $o_h = 0.3$, the input to the hidden neuron is $x_n = 15$.
Write down all intermediate steps of the calculation.

Assignment 17 (2 Points)

Compare the two Backpropagation of error variants:
single step learning and *cumulative learning*.
Describe both and discuss advantages and disadvantages of the two methods.

Assignment 18 (4 Points)

A simple 2-1 neural network is given with 2 inputs x_1, x_2, one output neuron with hyperbolic tangent, and the three weights (w_0, w_1, w_2). The objective for the network is to implement the Boolean function XOR, the input values are 0.0 or 1.0, the desired output values $\hat{y}$ are -1 or $+1$.

Depict the (global) error surface $G(w_0, w_1, w_2)$, for the BIAS-weight fixed to $w_0 = 2.0$, and the weights $-10.0 \leq w_1, w_2 \leq +10.0$.

Extra task (no extra points):
Change the setup to desired values 0 and 1, and the transfer function to be *relu*.

Assignment 19 (4 Points)

Calculate, how to change the **input vector** ${}^p\mathbf{X}$ of a 3-layer, N-H-M MLP to decrease the single error ${}^pE$ using gradient descent.
All weights $w_{i,j}$ of the network are to remain constant, the transfer functions in hidden layer and output layer are the hyperbolic tangent.

The way how to change the input values can be determined in analogy to the derivation of backpropagation of error.
Derive the formulas for a 3-layer N-H-M MLP to calculate the partial derivative of the ${}^pE({}^px_n)$ with respect to components ${}^px_n$ of the N-dimensional input vector ${}^p\mathbf{X}$.

$$\frac{\partial\, {}^pE({}^px_n)}{\partial\, {}^px_n} = ?$$