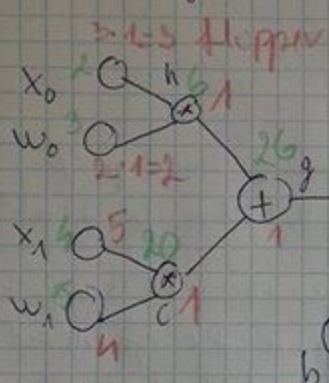


ZH-Sam:

- Backprop minimizes $\rightarrow$ neural net
- (variable minimizes)
- Rosenblatt XOR, OR, AND networks
- 1-2 iterations
- Elman's networks: pl. net as a feedforward?

Suzuki's network:



hidden global's gradients
 $\frac{\partial L}{\partial L} = 1$ grad flow
 Loss: min loss grad out
 grad in

$$\frac{\partial L}{\partial a} = \frac{\partial L}{\partial b} \cdot \frac{\partial b}{\partial a} \quad \frac{\partial L}{\partial b}$$

$$a' = \frac{\partial b}{\partial a} \quad \text{chain rule}$$

$$t = g + b \quad \frac{\partial t}{\partial t} = 1 \quad b' = \frac{\partial t}{\partial b} = \frac{\partial t}{\partial t} \cdot \frac{\partial t}{\partial b} = 1$$

$$\frac{\partial t}{\partial g} = \frac{\partial t}{\partial t} \cdot \frac{\partial t}{\partial g} = 1 \quad g = h + i$$

$$h = w_0 \cdot x_0$$

$$\frac{\partial t}{\partial h} = \frac{\partial t}{\partial t} \cdot \frac{\partial t}{\partial g} \cdot \frac{\partial g}{\partial h} = 1$$

ReLU 1×10
 $0 \rightarrow 1$

$\frac{1}{2} \frac{1}{1+e^{-x}}$

Lokalisiieren: $\delta = \begin{cases} e^{(l_i^{(m)})} \cdot (d_i - y_i^{(m)}) & \text{output layer} \\ e^{(l_i^{(m)})} \cdot \left(\sum_j w_{ij} \cdot \delta_j^{(m+1)} \right) & \text{hidden layer} \end{cases}$

$$w_{ij}^{(m)}(k) = w_{ij}(k) + \eta \cdot \delta_i^{(m)} \cdot y_j^{(m-1)}$$

24. November propagations



A für vigen mündig 1 von.

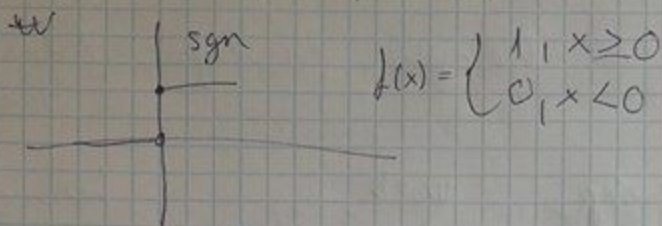
$$\frac{\partial f}{\partial x} = \frac{\partial f}{\partial h} \cdot \frac{\partial h}{\partial w} \cdot \frac{\partial w}{\partial x} = 0.5$$

$$\frac{\partial f}{\partial w} = \frac{\partial f}{\partial h} \cdot \frac{\partial h}{\partial w} \cdot \frac{\partial w}{\partial h} = 0.5$$

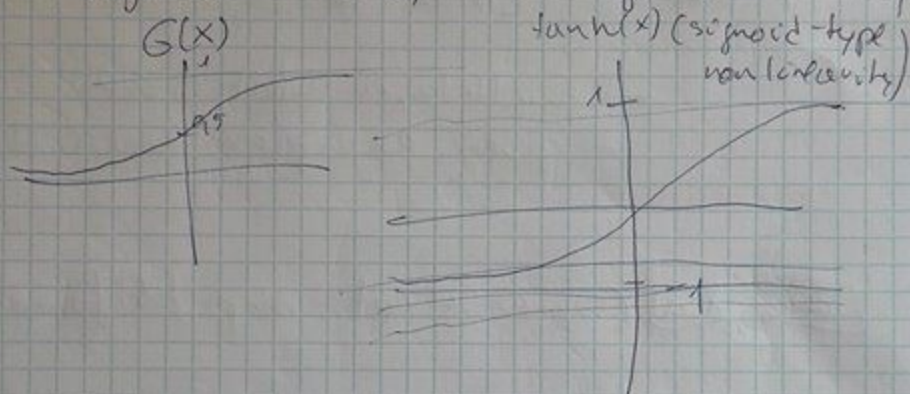
result

3.
 • fogalom tá: a leenelt dolgot

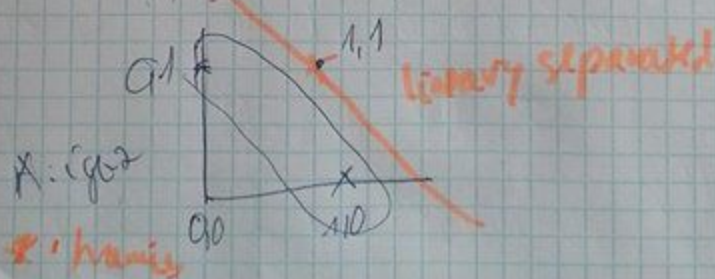
- Artificial neuron: node, ports, model
- activation func, bias



sigmoid linearity vs sigmoid nonlinearity



- element set separation
(binaries for linear separability)
'binary' is well



h.]

- Rosen-Bach algorithm

↳ tablas volt

How it works and we should use it.

- Perceptron convergent (never till infinity)

Initial step

$w_0 = b$
↓ x_1 x_2
↓ ↓ ↓

XOR $x = [1 \ 1 \ -1]$

Loss function =
= error function

- multilayer what is it?

- given examples how is it working?

Suma dia (suma school: deep feed forward network)

Gravitationsfeld wird als

Logarithmus

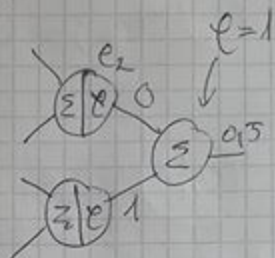
$$\left[\begin{aligned} e(u) &= \frac{1}{1 + e^{-u}} & e_a(u) &= \frac{1}{1 + e^{-au}} \\ f'(u) &= e(u)(1 - e(u)) & e'_a(u) &= a \cdot e(u)(1 - e(u)) \end{aligned} \right]$$

1.) FORWARD PROPAGATION

2.) ERROR CALCULATION

3.) LOCAL ERROR CALCULATION

4.) UPDATE WEIGHTS



① $y_3 = 0.5 \quad y_2 = 1 \quad y_1 = 0$ (müsste 0.5 sein)

② $E_i = d_i - o_i \Rightarrow E_1 = 1 - 0.5 = 1/2$

lokales Loss an 1. Neuron

③ $e_1^{(2)} = e_3 = 1 \cdot 1/2 = 1/2 \leftarrow \text{Error backpropagieren}$

$e_1^{(1)} = e_2 = \text{ReLU}(0) = 0 \leftarrow e_2$

$e_2^{(1)} = e_1 = \text{ReLU}'(1) \cdot 0.5 \cdot 0.5 = 0.25$

lok. Min $\uparrow$ $\uparrow$ $\uparrow$
Sitz $\uparrow$ $\uparrow$ $\uparrow$
vorne propagiert hinten

④ $w_{ij}^{(n)}(-1) = w_{ij}^{(n)} + \eta \cdot e_i^{(n)} \cdot y_i^{(n-1)}$

$w_{12}^{(2)} = 0.5 + 1 + e_3 \cdot 1 = 1 \leftarrow \text{zwei جايتايت}$

$w_{11}^{(2)} = -2 + 1 \cdot e_1^{(2)} \cdot 0 = -2$

ZH: Kapazitäten groß genug: davor hat es nicht funktioniert
• 4.5.21 10.000 Epochen