Technical Neural Networks
Assignments Sheet 3
Group: Svetlana Seliunina, Aleksei Zhuravlev

Assignment 15

- Complexity to implement a function. ReLU function is much easier to implement because it does not require calculating exponents or cos, sin functions.
- Gradient. Logistic function and hyperbolic function derivatives can be expressed using the function value and are continuous. ReLU is non-differentiable at z=0 and the gradient vanishes at z<0.
- Saturation effects. ReLU function is unbound for large z so there is no saturation effect. Logistic and hyperbolic functions are both bounded and there is a saturation effect for large positive and negative values.

Assignment 17

Single step learning - weight update is performed for every pattern.
Cumulative learning - weight changes for each pattern are accumulated and the weight update step is performed at the end of the epoch.
Single step learning approach will quickly reach an adequate solution because weight updates are performed more frequently. Cumulative learning is slower but calculates the actual network error so it may reach the optimal solution easier.

Assignment 19

$$\frac{\partial E(x_n)}{\partial x_n} = \frac{\partial E(x_n)}{\partial net} \cdot \frac{\partial net}{\partial x_n} = \frac{\partial E(x_n)}{\partial out} \cdot \frac{\partial out}{\partial net} \cdot \frac{\partial net}{\partial x_n} = -\delta \frac{\partial net}{\partial x_n}$$

δ depends on a chosen error function.

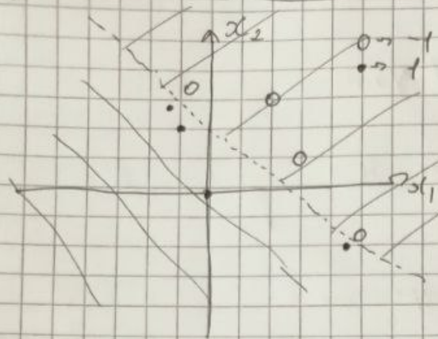
For a hidden neuron:

$$\frac{\partial net_h}{\partial x_n} = \frac{\partial}{\partial x_n}(x_n w_{nh}) = w_{nh}$$

For an output neuron:

$$\frac{\partial net_m}{\partial x_n} = \frac{\partial}{\partial x_n}(x_n w_{nh} + \tilde{out}_h w_{hm}) = w_{nh} + \frac{\partial out_n}{\partial x_n} w_{hm} = w_{nh} + f'(x_n) w_{hm}$$

Assignments 14,16,18



to separate the points, the line should connect $(-1;3)$ and $(5;-2)$

$sgn(\omega_0 \cdot 1 + x_1\omega_1 + x_2\omega_2) =$

$$\begin{cases} sgn(\omega_0 + 0\cdot\omega_1 + 0\cdot\omega_2) = 1 \\ \cancel{sgn(\omega_0 + (-1-\varepsilon)\omega_1 + (3-\varepsilon)\omega_2) = 1} \\ \cancel{sgn(\omega_0 + (-1+\varepsilon)\omega_1 + (3+\varepsilon)\omega_2) = -1} \end{cases} \Rightarrow \begin{cases} \cancel{\omega_0 > 0} \\ \cancel{\omega_0} \end{cases}$$

$\cancel{6y = -5x + 13}$

$\cancel{y = -\frac{5}{6}x + \frac{13}{6}}$

$\cancel{+\frac{5}{6}\omega_1 + \omega_2 = \frac{13}{6}\omega_0 = 0}$

line that connects $(-1;3)$ and $(5;-2)$: $-\frac{5}{6}x_1 - 1\cdot x_2 + 2\frac{1}{6} = 0$

so, if we set $(\omega_0, \omega_1, \omega_2) = \left(+2\frac{1}{6}, -\frac{5}{6}, -1\right)$, the points will be properly separated by $sgn(\omega_0 + \omega_1 x_1 + \omega_2 x_2)$

example: $sgn\left(2\frac{1}{6} + 0\cdot\left(\frac{5}{6}\right) - 1\cdot 0\right) = 1$

Assignment 16

tanh $\quad \eta = 0{,}1 \quad \hat{y}_{m=1} = 0{,}2 \quad y_{m=1} = 0{,}8 \quad \omega_{hm} = 2{,}5 \quad \theta_h = 0{,}3$

$x_n = 15$

$\Delta\omega_{ij} = \eta\, \delta_j\, out_j$

$\delta_m = (\hat{y}_m - y_m)\, f'(net_m)$

$\delta_h = \left(\sum_{k=1}^{K} \delta_k \omega_{hk}\right) f'(net_h)$

for output neuron:

$\Delta\omega_o = 0{,}1 \cdot \underbrace{(0{,}2 - 0{,}8)}_{-0{,}6} \cdot \tanh'\underbrace{(2{,}5 \cdot 0{,}3)}_{0{,}6} \cdot 0{,}2 = -0{,}007$

$\delta_o = -0{,}36$

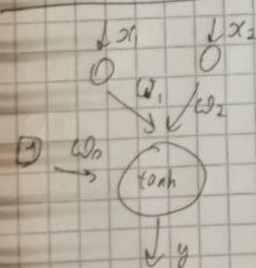
~~$\omega_{hm} = 2{,}49$~~

for hidden neuron:

~~$\Delta\omega_h = 0{,}1 \cdot (15) \cdot \tanh'(15) \cdot$~~

$\Delta\omega_h = 0{,}1 \cdot \left(2{,}5 \cdot (-0{,}36)\right) \cdot \tanh'(\overset{15 \cdot \omega_{nh}}{\cancel{15}}) \cdot 0{,}3 = \underline{0{,}027 \cdot \tanh'(15 \cdot \omega_h)}$

Assignment 18



XOR

X_1	X_2	y
0	0	-1
0	1	1
1	0	1
1	1	-1

$y = \tanh(\omega_0 + x_1\omega_1 + x_2\omega_2)$

${}^{p}E = \frac{1}{2}\left({}^{p}\hat{y} - {}^{p}y\right)^2$

$E = \sum_{p=1}^{P} {}^{p}E$

$x_1=0,\ x_2=0$

$y=-1 \Rightarrow {}^1E = \frac{1}{2}[-1-\tanh(\omega_0)]^2$

$x_1=0,\ x_2=1$

${}^2E = \frac{1}{2}[1-\tanh(\omega_0+\omega_2)]^2$

$x_1=1,\ x_2=0$

${}^3E = \frac{1}{2}[1-\tanh(\omega_0+\omega_1)]^2$

$x_1=1\ \ x_2=1$

${}^4E = \frac{1}{2}[-1-\tanh(\omega_0+\omega_1+\omega_2)]^2$

$$\sigma = {}^1E+{}^2E+{}^3E+{}^4E = \frac{1}{2}\{[-1-\tanh\omega_0]^2+[1-\tanh(\omega_0+\omega_1)]^2+$$
$$+[1-\tanh(\omega_0+\omega_2)]^2+[-1-\tanh(\omega_0+\omega_1+\omega_2)]^2\}$$

plot z = ((-1 - tanh(2 + x + y))^2 + (1 - tanh(2 + x))^2 + (1 - tanh(2 + y))^2 + (-1 - tanh(2))^2)/2 from x=-10 t

NATURAL LANGUAGE MATH INPUT EXTENDED KEYBOARD EXAMPLES UPLOAD RANDOM

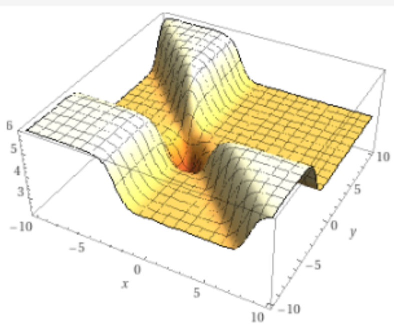
Input interpretation

plot	$z = \frac{1}{2}\left((-1-\tanh(2+x+y))^2+(1-\tanh(2+x))^2+(1-\tanh(2+y))^2+(-1-\tanh(2))^2\right)$	$x = -10$ to 10 $y = -10$ to 10

$\tanh(x)$ is the hyperbolic tangent function

3D plot

Show contour lines



Contour plot