Exc. 2

a) If we remove the Relu activation functions we get the Projection Pursuit Regression Model (PPR).

b) Dataset: $\left\{ \left([1,\ 0.1]^T,\ 2,3\right);\ \left([-0.5,\ 1]^T,\ 1.1\right) \right\}$

weights: $\begin{bmatrix} w_1 & u_1 \\ w_2 & u_2 \end{bmatrix}$, biases: $\begin{bmatrix} w_0 \\ u_0 \end{bmatrix}$. $V = [v_1,\ v_2]$

$V_0 = [v_0]$

Point 1:

compute 1st Layer

$$\text{layer1} = \begin{bmatrix} -2 & 2 \\ 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 0,1 \end{bmatrix} + \begin{bmatrix} 5 \\ -5 \end{bmatrix} =$$

$$= \begin{bmatrix} -1.8 \\ 1.1 \end{bmatrix} + \begin{bmatrix} 5 \\ -5 \end{bmatrix} = \begin{bmatrix} 3,2 \\ -3.9 \end{bmatrix}$$

Activate neuron:

$$\text{act1} = \text{relu}(\text{layer1}) = \begin{bmatrix} 3.2 \\ 0 \end{bmatrix}$$

Output

$$\hat{y} = [1\ \ 1] \begin{bmatrix} 3.2 \\ 0 \end{bmatrix} + [1] = 3.2 + 1 = 4.2$$

Update v_1

$$\frac{\partial E}{\partial v_1} = \frac{\partial E}{\partial \hat{y}} \cdot \frac{\partial \hat{y}}{\partial v_1} = (4,2-2.3)^2 \times 3.2 = 11.552$$

$$v_1 = v_1 - \mu(11.552) = 1 - 0.3(11.552) = 1 - 3.47 = \underline{\underline{-2.47}}$$

Update v_2

$$\frac{\partial E}{\partial v_2} = \frac{\partial E}{\partial \hat{y}} \cdot \frac{\partial \hat{y}}{\partial v_2} = (4,2-2.3)^2 \times 0 = 0$$

$$v_2 = v_2 - \mu(0) = v_2 = 1$$

Update w_1

$$\frac{\partial E}{\partial w_1} = \frac{\partial E}{\partial \hat{y}} \cdot \frac{\partial \hat{y}}{\partial z_1} \cdot \frac{\partial z_1}{\partial a_1} \cdot \frac{\partial a_1}{\partial w_1} =$$

$$= (4.2-2.3)^2 \times v_1 \times 1 \times 1 =$$

$$= 3.61 \times (-2.47) \times 1 \times 1 = -8.92$$

$$w_1 = w_1 - 0,3(-8.92) = -2 - (-2,676) = \underline{\underline{0,676}}$$

Update μ_1

$$\frac{\partial E}{\partial \mu_1} = \frac{\partial E}{\partial \hat{y}} \cdot \frac{\partial \hat{y}}{\partial z_1} \cdot \frac{\partial z_1}{\partial a_1} \cdot \frac{\partial a_1}{\partial \mu_1} =$$

$$= (4.2-2.3)^2 \times v_2 \times 1 \times 0,1 = 0,361$$

$$\mu_1 = \mu_1 - 0.3\,\underset{0.361}{(\;\;)} = 1.892$$

Point 2:

compute 1st layer:

$$\text{layer\_1} = \begin{bmatrix} -2 & 2 \\ 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} -0.5 \\ 1 \end{bmatrix} + \begin{bmatrix} 5 \\ -5 \end{bmatrix} =$$

$$= \begin{bmatrix} 3 \\ 0.5 \end{bmatrix} + \begin{bmatrix} 5 \\ -5 \end{bmatrix} = \begin{bmatrix} 8 \\ -4.5 \end{bmatrix}$$

Activate neuron:

$$\text{act} = \text{relu}(\text{layer\_1}) = \begin{bmatrix} 8 \\ 0 \end{bmatrix}$$

Output:

$$\hat{y} = \begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} 8 \\ 0 \end{bmatrix} + [1] = [8] + [1] = [9]$$

Update v_1

$$\frac{dE}{dv_1} = \frac{dE}{d\hat{y}} \cdot \frac{d\hat{y}}{dv_1} = (9 - 1.1)^2 \times 8 = 499.28$$

$$v_1 = v_1 - \mu\,(499.28) = 1 - 0.3(499.28) = -148.784$$

Update V_2

$$\frac{\partial E}{\partial v_2} = \frac{\partial E}{\partial \hat{y}} \cdot \frac{\partial \hat{y}}{\partial v_2} = (9-1.1)^2 \times 0 = 0$$

$$v_2 = v_2 - 0{,}3(0) = v_2 = 1$$

Update w_1

$$\frac{\partial E}{\partial w_1} = \frac{\partial E}{\partial \hat{y}} \cdot \frac{\partial \hat{y}}{\partial z_1} \cdot \frac{\partial z_1}{\partial a_1} \cdot \frac{\partial a_1}{\partial w_1} =$$

$$= (9-1.1)^2 \times v_1 \times 1 \times (-0.5) = [illegible] = 4642.805$$

$$w_1 = w_1 - 0.3\,(\overset{4642.805}{[illegible]}) = -2 - \overset{1392.84}{[illegible]} = \overset{-1394.84}{[illegible]}$$

Update u_1

$$\frac{\partial E}{\partial u_1} = \frac{\partial E}{\partial \hat{y}} \cdot \frac{\partial \hat{y}}{\partial z_1} \cdot \frac{\partial z_1}{\partial a_1} \cdot \frac{\partial a_1}{\partial u_1} =$$

$$= (9-1.1)^2 \times V_2 \times 1 \times 1 = 62.41 \times 1 = 62.41$$

$$u_1 = u_1 - 0{,}3(62.41) = 2 - 18.723 = -16.723$$