Aprendizagem 2023/24

Homework III – Group 004

(ist1102695, ist1102703)

I. Pen-and-paper

1)

1

a)

$$\phi\begin{pmatrix}0.7\\-0.3\end{pmatrix} = e^{\frac{-\left\|\begin{pmatrix}0.7\\-0.3\end{pmatrix}-\begin{pmatrix}0\\0\end{pmatrix}\right\|^2}{2}} = e^{-0.29}$$

$$= e^{\frac{-\left(\sqrt{(0-0.7)^2-(0+0.3)^2}\right)^2}{2}}$$

$$= e^{-\frac{0.58}{2}} = e^{-0.29}$$

$$\phi\begin{pmatrix}0.7\\-0.3\end{pmatrix} = e^{\frac{-\left\|\begin{pmatrix}0.7\\-0.3\end{pmatrix}-\begin{pmatrix}1\\-1\end{pmatrix}\right\|}{2}} = e^{-0.29}$$

$$\phi\begin{pmatrix}0.7\\-0.3\end{pmatrix} = e^{\frac{-\left\|\begin{pmatrix}0.7\\-0.3\end{pmatrix}-\begin{pmatrix}-1\\1\end{pmatrix}\right\|}{2}} = e^{-2.29}$$

$$\phi\begin{pmatrix}0.4\\0.5\end{pmatrix} = e^{\frac{-\left\|\begin{pmatrix}0.4\\0.5\end{pmatrix}-\begin{pmatrix}0\\0\end{pmatrix}\right\|}{2}} = e^{-0.205}$$

$$\phi\begin{pmatrix}0.4\\0.5\end{pmatrix} = e^{\frac{-\left\|\begin{pmatrix}0.4\\0.5\end{pmatrix}-\begin{pmatrix}1\\-1\end{pmatrix}\right\|}{2}} = e^{-1.305}$$

$$\phi\begin{pmatrix}0.4\\0.5\end{pmatrix} = e^{\frac{-\left\|\begin{pmatrix}0.4\\0.5\end{pmatrix}-\begin{pmatrix}-1\\1\end{pmatrix}\right\|}{2}} = e^{-1.105}$$

$$\phi\begin{pmatrix}-0.2\\0.8\end{pmatrix} = e^{\frac{-\left\|\begin{pmatrix}-0.2\\0.8\end{pmatrix}-\begin{pmatrix}0\\0\end{pmatrix}\right\|}{2}} = e^{-0.34}$$

$$\phi\begin{pmatrix}-0.2\\0.8\end{pmatrix} = e^{\frac{-\left\|\begin{pmatrix}-0.2\\0.8\end{pmatrix}-\begin{pmatrix}1\\-1\end{pmatrix}\right\|}{2}} = e^{-2.34}$$

$$\phi\begin{pmatrix}-0.2\\0.8\end{pmatrix} = e^{\frac{-\left\|\begin{pmatrix}-0.2\\0.8\end{pmatrix}-\begin{pmatrix}-1\\1\end{pmatrix}\right\|}{2}} = e^{-0.34}$$

$$\phi\begin{pmatrix}-0.4\\0.3\end{pmatrix} = e^{\frac{-\left\|\begin{pmatrix}-0.4\\0.3\end{pmatrix}-\begin{pmatrix}0\\0\end{pmatrix}\right\|}{2}} = e^{-0.125}$$

$$\phi\begin{pmatrix}-0.4\\0.3\end{pmatrix} = e^{\frac{-\left\|\begin{pmatrix}-0.4\\0.3\end{pmatrix}-\begin{pmatrix}1\\-1\end{pmatrix}\right\|}{2}} = e^{-1.825}$$

$$\phi\begin{pmatrix}-0.4\\0.3\end{pmatrix} = e^{\frac{-\left\|\begin{pmatrix}-0.4\\0.3\end{pmatrix}-\begin{pmatrix}-1\\1\end{pmatrix}\right\|}{2}} = e^{-0.425}$$

$$\Phi = \begin{pmatrix} 1 & e^{-0.29} & e^{-0.29} & e^{-0.29} \\ 1 & e^{-0.205} & e^{-1.305} & e^{-1.105} \\ 1 & e^{-0.34} & e^{-2.34} & e^{-0.34} \\ 1 & e^{-0.125} & e^{-1.825} & e^{-0.425} \end{pmatrix}$$

$$W = \left(\Phi^T \Phi + 0.1I\right)^{-1} \Phi z$$

b)

$$W = \left(\Phi^T \Phi + 0.1 I\right)^{-1} \Phi z$$

$$= \left(\begin{pmatrix} 1 & 1 & 1 & 1 \\ e^{-0.29} & e^{-0.205} & e^{-0.34} & e^{-0.125} \\ e^{-0.29} & e^{-1.305} & e^{-2.34} & e^{-1.825} \\ e^{-0.29} & e^{-1.105} & e^{-0.34} & e^{-0.425} \end{pmatrix} \begin{pmatrix} 1 & e^{-0.29} & e^{-0.29} & e^{-0.29} \\ 1 & e^{-0.205} & e^{-1.305} & e^{-1.105} \\ 1 & e^{-0.34} & e^{-2.34} & e^{-0.34} \\ 1 & e^{-0.125} & e^{-1.825} & e^{-0.425} \end{pmatrix} + \begin{pmatrix} 0.1 & 0 & 0 & 0 \\ 0 & 0.1 & 0 & 0 \\ 0 & 0 & 0.1 & 0 \\ 0 & 0 & 0 & 0.1 \end{pmatrix} \right)^{-1}$$

$$\left(\begin{pmatrix} 1 & 1 & 1 & 1 \\ e^{-0.29} & e^{-0.205} & e^{0.34} & e^{-0.125} \\ e^{-0.29} & e^{-1.305} & e^{-2.34} & e^{1.825} \\ e^{-0.29} & e^{-1.105} & e^{-0.34} & e^{-0.425} \end{pmatrix} \begin{pmatrix} 0.8 \\ 0.6 \\ 0.3 \\ 0.3 \end{pmatrix} \right)$$

$$= \begin{pmatrix} 0.339144 \\ 0.199453 \\ 0.100962 \\ -0.296003 \end{pmatrix}$$

$$\therefore\ z = 0.339144 + 0.199453 x_1 + 0.100962 x_2 - 0.296003 x_3$$

2)

x_1

$$z^{[1]} = W^{[1]} x_1 + b^{[1]} = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 2 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix} \times \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 5 \\ 6 \\ 5 \end{pmatrix} \qquad x^{[1]} = f(z^{[1]}) = \begin{pmatrix} \tanh(0{,}5\times5-2) \\ \tanh(0{,}5\times6-2) \\ \tanh(0{,}5\times5-2) \end{pmatrix} = \begin{pmatrix} 0{,}46212 \\ 0{,}76159 \\ 0{,}46212 \end{pmatrix}$$

$$z^{[2]} = W^{[2]} x^{[1]} + b^{[2]} = \begin{pmatrix} 1 & 4 & 1 \\ 1 & 1 & 1 \end{pmatrix} \times \begin{pmatrix} 0{,}46212 \\ 0{,}76159 \\ 0{,}46212 \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 4{,}9706 \\ 2{,}68583 \end{pmatrix} \qquad x^{[2]} = f(z^{[2]}) = \begin{pmatrix} 0{,}45048 \\ -0{,}57642 \end{pmatrix}$$

$$z^{[3]} = W^{[3]} x^{[2]} + b^{[3]} = \begin{pmatrix} 1 & 1 \\ 3 & 1 \\ 1 & 1 \end{pmatrix} \times \begin{pmatrix} 0{,}45048 \\ -0{,}57642 \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0{,}87406 \\ 1{,}77502 \\ 0{,}87406 \end{pmatrix} \qquad x^{[3]} = f(z^{[3]}) = \begin{pmatrix} -0{,}91590 \\ -0{,}80494 \\ -0{,}91590 \end{pmatrix}$$

$$\delta^{[3]} = (x^{[3]} - t) \circ \phi'(z^{[3]}) = \left(\begin{pmatrix} -0{,}91590 \\ -0{,}80494 \\ -0{,}91590 \end{pmatrix} - \begin{pmatrix} -1 \\ 1 \\ -1 \end{pmatrix} \right) \circ \phi'(z^{[3]}) = \begin{pmatrix} 0{,}0841 \\ -1{,}8049 \\ 0{,}0841 \end{pmatrix} \circ \begin{pmatrix} 0{,}0806 \\ 0{,}1760 \\ 0{,}0806 \end{pmatrix} = \begin{pmatrix} 0{,}00678 \\ -0{,}31225 \\ 0{,}00678 \end{pmatrix}$$

$$\phi' = 0{,}5(1-\tanh^2(0{,}5x-2))$$

$$\delta^{[2]} = W^{[3]T} \delta^{[3]} \circ \phi'(z^{[2]}) = \begin{pmatrix} 1 & 3 & 1 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 0{,}00678 \\ -0{,}31225 \\ 0{,}00678 \end{pmatrix} \circ \begin{pmatrix} 0{,}399 \\ 0{,}334 \end{pmatrix} = \begin{pmatrix} -0{,}36835 \\ -0{,}09976 \end{pmatrix}$$

$$\delta^{[1]} = W^{[2]T} \delta^{[2]} \circ \phi'(z^{[1]}) = \begin{pmatrix} 1 & 1 \\ 4 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} -0{,}36835 \\ -0{,}09976 \end{pmatrix} \circ \begin{pmatrix} 0{,}393 \\ 0{,}21 \\ 0{,}393 \end{pmatrix} = \begin{pmatrix} -0{,}18396 \\ -0{,}33036 \\ -0{,}18396 \end{pmatrix}$$

$$\frac{\partial E}{\partial W^{[1]}} = \delta^{[1]} x_1^T = \begin{pmatrix} -0{,}18396 \\ -0{,}33036 \\ -0{,}18396 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 & 1 \end{pmatrix} = \begin{pmatrix} -0{,}18396 & -0{,}18396 & -0{,}18396 & -0{,}18396 \\ -0{,}33036 & -0{,}33036 & -0{,}33036 & -0{,}33036 \\ -0{,}18396 & -0{,}18396 & -0{,}18396 & -0{,}18396 \end{pmatrix}$$

$$\frac{\partial E}{\partial W^{[2]}} = \delta^{[2]} x^{[1]T} = \begin{pmatrix} -0{,}36835 \\ -0{,}09976 \end{pmatrix} \begin{pmatrix} 0{,}46212 \\ 0{,}76159 \\ 0{,}46212 \end{pmatrix}^T = \begin{pmatrix} -0{,}17022 & -0{,}280532 & -0{,}17022 \\ -0{,}04610 & -0{,}075976 & -0{,}04610 \end{pmatrix}$$

$$\frac{\partial E}{\partial W^{[3]}} = \delta^{[3]} x^{[2]T} = \begin{pmatrix} 0{,}00678 \\ -0{,}31225 \\ 0{,}00678 \end{pmatrix} \begin{pmatrix} 0{,}45048 \\ -0{,}57642 \end{pmatrix}^T = \begin{pmatrix} 0{,}003054 & -0{,}003908 \\ -0{,}14066 & 0{,}17999 \\ 0{,}003054 & -0{,}003908 \end{pmatrix}$$

x_2

$$z^{[1]} = W^{[1]} x_2 + b^{[1]} = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 2 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix} \times \begin{pmatrix} 1 \\ 0 \\ 0 \\ -1 \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \qquad x^{[1]} = f(z^{[1]}) = \begin{pmatrix} -0,90515 \\ -0,90515 \\ -0,90515 \end{pmatrix}$$

$$z^{[2]} = W^{[2]} x^{[1]} + b^{[2]} = \begin{pmatrix} 1 & 4 & 1 \\ 1 & 1 & 1 \end{pmatrix} \times \begin{pmatrix} -0,90515 \\ -0,90515 \\ -0,90515 \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} -4,4309 \\ -1,71545 \end{pmatrix} \qquad x^{[2]} = f(z^{[2]}) = \begin{pmatrix} -0,99956 \\ -0,99343 \end{pmatrix}$$

$$z^{[3]} = W^{[3]} x^{[2]} + b^{[3]} = \begin{pmatrix} 1 & 1 \\ 3 & 1 \\ 1 & 1 \end{pmatrix} \times \begin{pmatrix} -0,99956 \\ -0,99343 \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} -0,99299 \\ -2,99211 \\ -0,99299 \end{pmatrix} \qquad x^{[3]} = f(z^{[3]}) = \begin{pmatrix} -0,98652 \\ -0,99816 \\ -0,98652 \end{pmatrix}$$

$$\delta^{[3]} = (x^{[3]} - t) \circ \phi'(z^{[3]}) = \left(\begin{pmatrix} -0,98652 \\ -0,99816 \\ -0,98652 \end{pmatrix} - \begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix} \right) \circ \begin{pmatrix} 0,013388 \\ 0,00183 \\ 0,013388 \end{pmatrix} = \begin{pmatrix} -0,026596 \\ 0,000003 \\ 0,00018 \end{pmatrix}$$

$$\delta^{[2]} = W^{[3]T} \delta^{[3]} \circ \phi'(z^{[2]}) = \begin{pmatrix} 1 & 3 & 1 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} -0,026596 \\ 0,000003 \\ 0,00018 \end{pmatrix} \circ \begin{pmatrix} 0,00044 \\ 0,00655 \end{pmatrix} = \begin{pmatrix} -0,000012 \\ -0,000173 \end{pmatrix}$$

$$\delta^{[1]} = W^{[2]T} \delta^{[2]} \circ \phi'(z^{[1]}) = \begin{pmatrix} 1 & 1 \\ 4 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} -0,000012 \\ -0,000173 \end{pmatrix} \circ \begin{pmatrix} 0,09035 \\ 0,09035 \\ 0,09035 \end{pmatrix} = \begin{pmatrix} -0,0000167 \\ -0,0000200 \\ -0,0000167 \end{pmatrix}$$

$$\frac{\partial E}{\partial W^{[1]}} = \delta^{[1]} x_1^T = \begin{pmatrix} -0,0000167 \\ -0,0000200 \\ -0,0000167 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & -1 \end{pmatrix} = \begin{pmatrix} -0\ 0000167 & 0 & 0 & 0\ 0000167 \\ -0,0000200 & 0 & 0 & 0,0000200 \\ -0,0000167 & 0 & 0 & 0,0000167 \end{pmatrix}$$

$$\frac{\partial E}{\partial W^{[2]}} = \delta^{[2]} x^{[1]T} = \begin{pmatrix} -0,000012 \\ -0,000173 \end{pmatrix} \begin{pmatrix} -0,90515 \\ -0,90515 \\ -0,90515 \end{pmatrix}^T = \begin{pmatrix} 0,0000109 & 0,0000109 & 0,0000109 \\ 0,000157 & 0,000157 & 0,000157 \end{pmatrix}$$

$$\frac{\partial E}{\partial W^{[3]}} = \delta^{[3]} x^{[2]T} = \begin{pmatrix} -0,026596 \\ 0,000003 \\ 0,00018 \end{pmatrix} \begin{pmatrix} -0,99956 \\ -0,99343 \end{pmatrix}^T = \begin{pmatrix} 0,02658 & 0,02642 \\ -0,000003 & -0,00000298 \\ -0,0001799 & -0,0001788 \end{pmatrix}$$

Homework III – Group 004

(ist1102695, ist1102703)

Cálculo dos pesos:

$$\frac{\partial E}{\partial W^{[1]}} = \frac{\partial E}{\partial w^{[1]}_{x_1}} + \frac{\partial E}{\partial w^{[1]}_{x_2}} = \begin{pmatrix} -0{,}18396 & -0{,}18396 & -0{,}18396 & -0{,}18396 \\ -0{,}33036 & -0{,}33036 & -0{,}33036 & -0{,}33036 \\ -0{,}18396 & -0{,}18396 & -0{,}18396 & -0{,}18396 \end{pmatrix} + \begin{pmatrix} -0{,}0000167 & 0 & 0 & 0\ 0000167 \\ -0{,}0000200 & 0 & 0 & 0{,}0000200 \\ -0{,}0000167 & 0 & 0 & 0{,}0000167 \end{pmatrix} =$$

$$= \begin{pmatrix} -0{,}18398 & -0{,}18396 & -0{,}18396 & -0{,}18394 \\ -0{,}33038 & -0{,}33036 & -0{,}33036 & -0{,}33034 \\ -0{,}18398 & -0{,}18396 & -0{,}18396 & -0{,}18394 \end{pmatrix}$$

$$W^{[1]}_{new} = W^{[1]}_{old} - \eta \frac{\partial E}{\partial W^{[1]}} = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 2 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix} - 0{,}1 \times \frac{\partial E}{\partial W^{[1]}} = \begin{pmatrix} 1{,}018398 & 1{,}018396 & 1{,}018396 & 1{,}018394 \\ 1{,}033038 & 1{,}033036 & 2{,}033036 & 1{,}033034 \\ 1{,}018398 & 1{,}018396 & 1{,}018396 & 1{,}018394 \end{pmatrix}$$

$$\frac{\partial E}{\partial b^{[1]}} = \delta^{[1]}_{x_1} + \delta^{[1]}_{x_2} = \begin{pmatrix} -0{,}18396 \\ -0{,}33036 \\ -0{,}18396 \end{pmatrix} + \begin{pmatrix} -0{,}0000167 \\ -0{,}0000200 \\ -0{,}0000167 \end{pmatrix} = \begin{pmatrix} -0{,}183977 \\ -0{,}33038 \\ -0{,}183977 \end{pmatrix}$$

$$b^{[1]}_{new} = b^{[1]}_{old} - \eta \frac{\partial E}{\partial b^{[1]}} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} - 0{,}1 \times \begin{pmatrix} -0{,}183977 \\ -0{,}33038 \\ -0{,}183977 \end{pmatrix} = \begin{pmatrix} 1{,}0183977 \\ 1{,}033038 \\ 1{,}0183977 \end{pmatrix}$$

$$\frac{\partial E}{\partial W^{[2]}} = \frac{\partial E}{\partial w^{[2]}_{x_1}} + \frac{\partial E}{\partial w^{[2]}_{x_2}} = \begin{pmatrix} -0{,}17022 & -0{,}280532 & -0{,}17022 \\ -0{,}04610 & -0{,}075976 & -0{,}04610 \end{pmatrix} + \begin{pmatrix} 0{,}0000109 & 0{,}0000109 & 0{,}0000109 \\ 0{,}000157 & 0{,}000157 & 0{,}000157 \end{pmatrix} =$$

$$= \begin{pmatrix} -0{,}170209 & -0{,}28052 & -0{,}170209 \\ -0{,}0459 & -0{,}07582 & -0{,}0459 \end{pmatrix}$$

$$W_{new}^{[2]} = W_{old}^{[2]} - \eta \frac{\partial E}{\partial W^{[2]}} = \begin{pmatrix} 1 & 4 & 1 \\ 1 & 1 & 1 \end{pmatrix} - 0{,}1 \times \begin{pmatrix} -0{,}170209 & -0{,}28052 & -0{,}170209 \\ -0{,}0459 & -0{,}07582 & -0{,}0459 \end{pmatrix} =$$

$$= \begin{pmatrix} 1{,}01702 & 4{,}028052 & 1{,}0170209 \\ 1{,}00459 & 1{,}007582 & 1{,}00459 \end{pmatrix}$$

$$\frac{\partial E}{\partial b^{[2]}} = \delta_{x_1}^{[2]} + \delta_{x_2}^{[2]} = \begin{pmatrix} -0{,}36835 \\ -0{,}09976 \end{pmatrix} + \begin{pmatrix} -0{,}000012 \\ -0{,}000173 \end{pmatrix} = \begin{pmatrix} -0{,}368362 \\ -0{,}099933 \end{pmatrix}$$

$$b_{new}^{[2]} = b_{old}^{[2]} - \eta \frac{\partial E}{\partial b^{[2]}} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} - 0{,}1 \times \begin{pmatrix} -0{,}368362 \\ -0{,}099933 \end{pmatrix} = \begin{pmatrix} 1{,}0368362 \\ 1{,}0099933 \end{pmatrix}$$

$$\frac{\partial E}{\partial W^{[3]}} = \frac{\partial E}{\partial W^{[3]}_{x_1}} + \frac{\partial E}{\partial W^{[3]}_{x_2}} = \begin{pmatrix} 0{,}003054 & -0{,}003908 \\ -0{,}14066 & 0{,}17999 \\ 0{,}003054 & -0{,}003908 \end{pmatrix} + \begin{pmatrix} 0{,}02658 & 0{,}02642 \\ -0{,}000003 & -0{,}00000298 \\ -0{,}0001799 & -0{,}0001788 \end{pmatrix} =$$

$$= \begin{pmatrix} 0{,}029634 & 0{,}022512 \\ -0{,}140663 & 0{,}179987 \\ 0{,}002874 & -0{,}004087 \end{pmatrix}$$

$$W_{new}^{[3]} = W_{old}^{[3]} - \eta \frac{\partial E}{\partial W^{[3]}} = \begin{pmatrix} 1 & 1 \\ 3 & 1 \\ 1 & 1 \end{pmatrix} - 0{,}1 \times \begin{pmatrix} 0{,}029634 & 0{,}022512 \\ -0{,}140663 & 0{,}179987 \\ 0{,}002874 & -0{,}004087 \end{pmatrix} =$$

$$= \begin{pmatrix} 0{,}9970366 & 0{,}9977488 \\ 3{,}0140663 & 0{,}9820013 \\ 0{,}9997126 & 1{,}0004087 \end{pmatrix}$$

$$\frac{\partial E}{\partial b^{[3]}} = \delta_{x_1}^{[3]} + \delta_{x_2}^{[3]} = \begin{pmatrix} 0{,}00678 \\ -0{,}31225 \\ 0{,}00678 \end{pmatrix} + \begin{pmatrix} -0{,}026596 \\ 0{,}000003 \\ 0{,}00018 \end{pmatrix} = \begin{pmatrix} -0{,}019816 \\ -0{,}312247 \\ 0{,}00696 \end{pmatrix}$$

$$b_{new}^{[3]} = b_{old}^{[3]} - \eta \frac{\partial E}{\partial b^{[3]}} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} - 0{,}1 \times \begin{pmatrix} -0{,}019816 \\ -0{,}312247 \\ 0{,}00696 \end{pmatrix} = \begin{pmatrix} 1{,}001981 \\ 1{,}0312247 \\ 0{,}999304 \end{pmatrix}$$

TÉCNICO LISBOA

Aprendizagem 2023/24

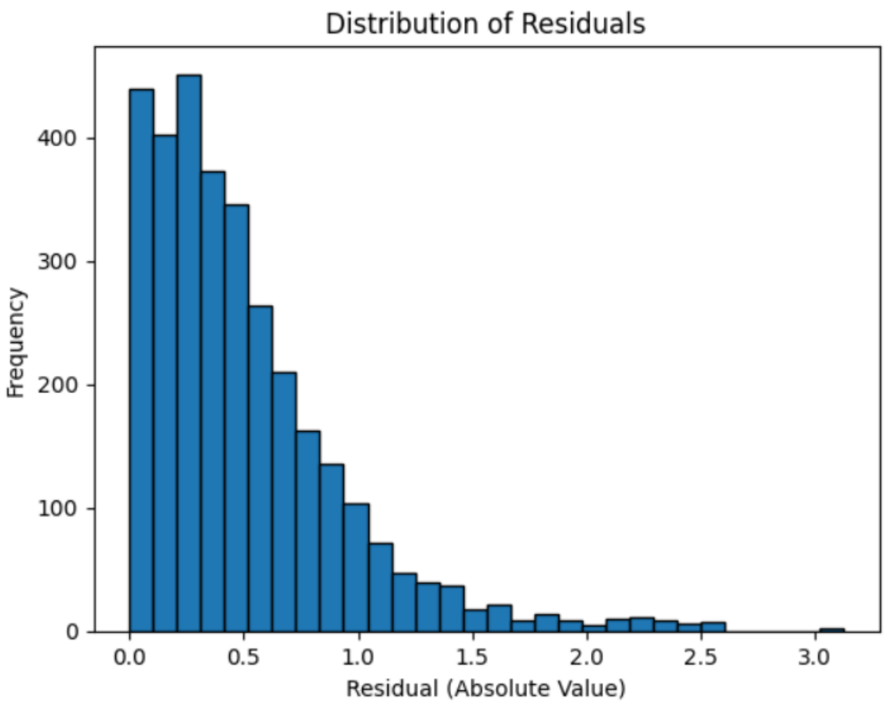
Homework III – Group 004

(ist1102695, ist1102703)

II. Programming and critical analysis

1)

```python
import numpy as np
import pandas as pd
from sklearn.model_selection import train_test_split
from sklearn.neural_network import MLPRegressor
import matplotlib.pyplot as plt
import warnings
from sklearn.exceptions import ConvergenceWarning

data = pd.read_csv("winequality-red.csv", delimiter=";")
X = data.drop("quality", axis=1)
y = data["quality"]

warnings.filterwarnings("ignore", category=ConvergenceWarning)

random_seeds = range(1, 11)

residuals = []

for random_seed in random_seeds:
    X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=0)

    mlp = MLPRegressor(hidden_layer_sizes=(10, 10), activation='relu', early_stopping=True, random_state=random_seed)
    mlp.fit(X_train, y_train)

    y_pred = mlp.predict(X_test)

    residual = np.abs(y_pred - y_test)
    residuals.extend(residual)

plt.hist(residuals, bins=30, edgecolor='k')
plt.title("Distribution of Residuals")
plt.xlabel("Residual (Absolute Value)")
plt.ylabel("Frequency")
plt.show()
```



Homework III – Group 004

(ist1102695, ist1102703)

2)

```
#2

mae = []
mae_rounded = []

for random_state in random_seeds:

    mae.append(np.mean(np.abs(y_test - y_pred)))

    y_pred_rounded = np.round(y_pred).clip(min=3, max=8)

    mae_rounded.append(np.mean(np.abs(y_test - y_pred_rounded)))

avg_mae_original = np.mean(mae)
avg_mae_rounded_bounded = np.mean(mae_rounded)

print(f"Average MAE (Unrouded and Unbounded Predictions): {avg_mae_original}")
print(f"Average MAE (Rounded & Bounded Predictions): {avg_mae_rounded_bounded}")
```

```
Average MAE (Unrouded and Unbounded Predictions): 0.49759676677876385
Average MAE (Rounded & Bounded Predictions): 0.415625
```

3)

```
#3
import pandas as pd
from sklearn.model_selection import train_test_split
from sklearn.neural_network import MLPRegressor
from sklearn.metrics import mean_squared_error
import numpy as np
import warnings
from math import sqrt


iterations = [20, 50, 100, 200]
num_runs = 10
rmse_scores = []

for num_iterations in iterations:
    mse_scores = []
    for random_state in range(1, num_runs + 1):
        mlp = MLPRegressor(hidden_layer_sizes=(10, 10),max_iter=num_iterations, activation='relu', random_state=random_state)

        mlp.fit(X_train, y_train)

        y_pred = mlp.predict(X_test)
        mse = sqrt(mean_squared_error(y_test, y_pred))
        mse_scores.append(mse)

    avg_rmse = np.mean(mse_scores)
    rmse_scores.append(avg_rmse)

for i, num_iterations in enumerate(iterations):
    print(f"Number of Iterations: {num_iterations}, Average RMSE: {rmse_scores[i]}")
```

```
Number of Iterations: 20, Average RMSE: 1.4039789509925442
Number of Iterations: 50, Average RMSE: 0.7996073631460566
Number of Iterations: 100, Average RMSE: 0.6940361469112144
Number of Iterations: 200, Average RMSE: 0.6554543932216474
```

4) No estado inicial do modelo, o número de iterações é 20, com um RMSE médio de 0,6922. Com um número limitado de iterações, o modelo não teve oportunidades suficientes para aprender os padrões subjacentes nos dados, levando a um RMSE relativamente mais elevado.

Quando o número de iterações é 50, ficamos com um RMSE médio de 0,6539. Com mais iterações, o modelo tem mais chances de ajustar seus parâmetros e ajustar os dados de treino. Consequentemente, o RMSE diminui, indicando uma melhoria na precisão preditiva do modelo.

Agora, o número de iterações é 100 com RMSE médio de 0,6476. Isto mostra que aumentar ainda mais o número de iterações continua a melhorar o desempenho do modelo. O RMSE diminui ligeiramente, sugerindo que o modelo ainda está a aprender suas representações.

Finalmente para um número de iterações de 200, temos um RMSE médio de 0,6403. Com 200 iterações, o modelo atinge o RMSE médio mais baixo neste experimento. Isto implica que o modelo beneficia de um treino prolongado, refinando as suas representações e previsões internas, o que resulta de uma precisão maior.

Os resultados obtidos mostram uma tendência de diminuição do RMSE à medida que o número de iterações aumenta. Concluímos então que permitir que o modelo treine para mais iterações geralmente leva a um melhor desempenho.

Quanto ao early stopping, pode ser benéfico em situações em que o modelo começa a ter overfitting nos dados de treino. O overfitting ocorre quando o modelo se torna muito complexo e começa a ajustar nos dados de treino, levando a um desempenho insatisfatório em dados não vistos. O early stopping evita o overfitting, interrompendo o processo de treinamento quando o desempenho em um conjunto de dados de validação começa a diminuir. Isso garante que o modelo generalize bem para dados invisíveis.

Porém, neste caso específico, a técnica de early stopping piora o desempenho. A razão para esta conclusão é que o RMSE diminui consistentemente à medida que o número de iterações aumenta, indicando que o modelo continua a melhorar com mais treino. Utilizar early stopping neste cenário pode impedir que o modelo atinja todo o seu potencial e resulta numa solução abaixo do ideal.

END