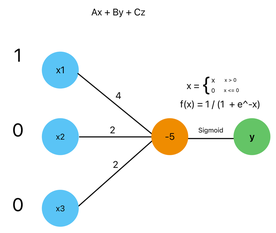
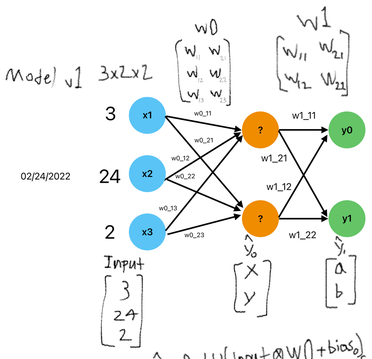


Model v1 3x2x2

02/24/2022



$$W0 = \begin{bmatrix} w_{11} & w_{21} \\ w_{12} & w_{22} \\ w_{13} & w_{23} \end{bmatrix} \qquad W1 = \begin{bmatrix} w_{11} & w_{21} \\ w_{12} & w_{22} \end{bmatrix}$$

$$\text{Input} = \begin{bmatrix} 3 \\ 24 \\ 2 \end{bmatrix} \qquad \hat{y}_0 = \begin{bmatrix} x \\ y \end{bmatrix} \qquad \hat{y}_1 = \begin{bmatrix} a \\ b \end{bmatrix}$$

$$\text{Loss} \begin{bmatrix} c \\ d \end{bmatrix} = \frac{1}{2}(\hat{y}_1 - \text{target})^2 = \begin{bmatrix} (a-\hat{a})^2 \\ (b-\hat{b})^2 \end{bmatrix}$$

$$\hat{y}_0 = \text{ReLU}(\text{input} \otimes W0 + \text{bias}_0)$$

$$\text{sum} = \text{Input} \otimes W0 = \begin{bmatrix} 3 \\ 24 \\ 2 \end{bmatrix} \otimes \begin{bmatrix} w_1 & w_2 \\ w_3 & w_4 \\ w_5 & w_6 \end{bmatrix} = \begin{bmatrix} 3w_1 + 24w_3 + 2w_5 \\ 3w_2 + 24w_4 + 2w_6 \end{bmatrix}$$

$$z_1 = \text{sum} + \text{bias}_0 = \begin{bmatrix} 3w_1 + 24w_3 + 2w_5 + b \\ 3w_2 + 24w_4 + 2w_6 + b \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix}$$

$$\hat{y}_1 = \text{ReLU}(\hat{y}_0 \otimes w1 + \text{bias}_1)$$

$$\hat{y}_0 \otimes w1 = \begin{bmatrix} x \\ y \end{bmatrix} \begin{bmatrix} w_1 & w_2 \\ w_3 & w_4 \end{bmatrix} = \begin{bmatrix} xw_1 + yw_3 \\ xw_2 + yw_4 \end{bmatrix}$$

$$z_2 = \hat{y} \otimes w_1 + \text{bias}_1 = \begin{bmatrix} xw_1 + yw_3 + b \\ xw_2 + yw_4 + b \end{bmatrix} = \begin{bmatrix} a \\ b \end{bmatrix}$$

$$\text{target} = \begin{bmatrix} \hat{a} \\ \hat{b} \end{bmatrix}$$

$$\frac{d}{d\text{Loss}} = \hat{y}_1 - \text{target} = \begin{bmatrix} a \\ b \end{bmatrix} - \begin{bmatrix} \hat{a} \\ \hat{b} \end{bmatrix} = \begin{bmatrix} c \\ d \end{bmatrix}$$

$$\text{ReLU} = \begin{cases} x & x > 0 \\ 0 & x <= 0 \end{cases} \qquad \text{ReLU}'(\hat{y}_1) = \begin{bmatrix} a' \\ b' \end{bmatrix}$$

$$\text{ReLU}' = \begin{cases} 1 & x > 0 \\ 0 & x <= 0 \end{cases} \qquad \text{ReLU}'(\hat{y}_0) = \begin{bmatrix} x' \\ y' \end{bmatrix}$$

Hadamard Product

$$\begin{cases} \partial\text{Loss} \otimes \text{ReLU}'(\hat{y}_1) = \begin{bmatrix} c \\ d \end{bmatrix} \otimes \begin{bmatrix} a' \\ b' \end{bmatrix} = \begin{bmatrix} ca' \\ db' \end{bmatrix} = L_1' \\ \partial\text{Loss} \otimes \text{ReLU}'(\hat{y}_0) = \begin{bmatrix} c \\ d \end{bmatrix} \otimes \begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} cx' \\ dy' \end{bmatrix} = L_0' \end{cases}$$

$$\frac{\partial L}{\partial w_1} = L_1' \otimes \hat{y}_0 = \begin{bmatrix} ca' \\ db' \end{bmatrix} \otimes \begin{bmatrix} xw_1 + yw_3 \\ xw_2 + yw_4 \end{bmatrix} = \begin{bmatrix} (xw_1 + yw_3)(ca') \\ (xw_2 + yw_4)(db') \end{bmatrix} = \begin{bmatrix} \hat{w}_{11} \\ \hat{w}_{12} \end{bmatrix}$$

$$\frac{\partial L}{\partial w_0} = L_0' \otimes \text{Input} = \begin{bmatrix} ca' \\ db' \end{bmatrix} \overset{\text{outer}}{\otimes} \begin{bmatrix} 3 \\ 24 \\ 2 \end{bmatrix} = \begin{bmatrix} 3ca' & 3db' \\ 24ca' & 24db' \\ 2ca' & 2db' \end{bmatrix} = \begin{bmatrix} \hat{w}_{11} & \hat{w}_{21} \\ \hat{w}_{12} & \hat{w}_{22} \\ \hat{w}_{13} & \hat{w}_{23} \end{bmatrix}$$

adjust weights

$$w_0 = w_0 - \alpha \frac{\partial L}{\partial w_0} = \begin{bmatrix} w_{11} - \alpha\hat{w}_{11} & w_{21} - \alpha\hat{w}_{21} \\ w_{12} - \alpha\hat{w}_{12} & w_{22} - \alpha\hat{w}_{22} \\ w_{13} - \alpha\hat{w}_{13} & w_{23} - \alpha\hat{w}_{23} \end{bmatrix}$$

$$w_1 = w_1 - \alpha \frac{\partial L}{\partial w} = \begin{bmatrix} w_{11} - \alpha\hat{w}_{11} & w_{21} - \alpha\hat{w}_{21} \\ w_{12} - \alpha\hat{w}_{12} & w_{22} - \alpha\hat{w}_{22} \end{bmatrix}$$

outer Prod

$$L_1' = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

$$L_1' \otimes \text{input} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

W_0 update

$$\frac{\partial L}{\partial R} \Rightarrow 2\times1$$

$$\frac{\partial L}{\partial R} \otimes \text{input} = \begin{bmatrix} a \\ b \end{bmatrix} [x_1 \; x_2 \; x_3] = \begin{bmatrix} x_1 a & x_1 b \\ x_2 a & x_2 b \\ x_3 a & x_3 b \end{bmatrix}$$

Layer 1

$$\begin{bmatrix} 3 \\ 24 \\ 2 \end{bmatrix} \begin{bmatrix} w_{11} & w_{21} \\ w_{12} & w_{22} \\ w_{13} & w_{23} \end{bmatrix}^T = \begin{bmatrix} 3 \\ 24 \\ 2 \end{bmatrix} \begin{bmatrix} w_{11} & w_{12} & w_{13} \\ w_{21} & w_{22} & w_{23} \end{bmatrix} \Rightarrow \begin{bmatrix} 3 \\ 24 \\ 2 \end{bmatrix} \begin{bmatrix} 1.124 & 0.760 & 2.402 \\ 0.687 & -0.311 & 1.673 \end{bmatrix} = \begin{bmatrix} 26.614 & -2.057 \end{bmatrix}$$

Layer 2

$$\begin{bmatrix} 26.614 & -2.057 \end{bmatrix} \begin{bmatrix} 3.637 & 2.511 \\ 1.838 & 1.284 \end{bmatrix} = \begin{bmatrix} 96.074 + (-3.780) & 66.827 + (-4.492) \end{bmatrix}$$

$$= \begin{bmatrix} 92.294 & 62.395 \end{bmatrix}$$

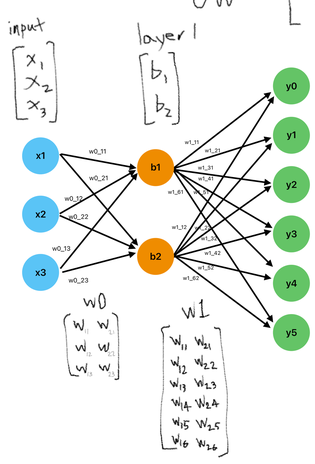
Layer 1 with bias

$$\begin{bmatrix} 3 \\ 24 \\ 2 \end{bmatrix} \begin{bmatrix} w_{11} & w_{21} \\ w_{12} & w_{22} \\ w_{13} & w_{23} \end{bmatrix}^T - b_1 = \begin{bmatrix} 3 \\ 24 \\ 2 \end{bmatrix} \begin{bmatrix} w_{11} & w_{12} & w_{13} \\ w_{21} & w_{22} & w_{23} \end{bmatrix} - 1.205 \Rightarrow \begin{bmatrix} 3 \\ 24 \\ 2 \end{bmatrix} \begin{bmatrix} 1.124 & 0.760 & 2.402 \\ 0.687 & -0.311 & 1.673 \end{bmatrix} - 1.205 = \begin{bmatrix} 26.614 & -2.057 \end{bmatrix} - 1.205 = \begin{bmatrix} 25.409 & -3.262 \end{bmatrix}$$

Layer 2 with bias

$$\begin{bmatrix} 25.409 & -3.262 \end{bmatrix} \begin{bmatrix} 3.637 & 2.511 \\ 1.838 & 1.284 \end{bmatrix} - 1.207 = \begin{bmatrix} 86.416 & 59.613 \end{bmatrix} - 1.207 = \begin{bmatrix} 85.209 & 58.406 \end{bmatrix}$$

Model v2 3x2x6



$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \qquad \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$$

$$W0 = \begin{bmatrix} w_{11} & w_{21} \\ w_{12} & w_{22} \\ w_{13} & w_{23} \end{bmatrix} \qquad W1 = \begin{bmatrix} w_{11} & w_{21} \\ w_{12} & w_{22} \\ w_{13} & w_{23} \\ w_{14} & w_{24} \\ w_{15} & w_{25} \\ w_{16} & w_{26} \end{bmatrix}$$

Output top 3 predictions

$$\begin{bmatrix} y_0 \\ y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \end{bmatrix} \begin{matrix} \} [\text{lat}_1, \text{long}_1] \\ \\ \} [\text{lat}_2, \text{long}_2] \\ \\ \} [\text{lat}_3, \text{long}_3] \\ \\ \end{matrix}$$

Loss vector needs to be 6D in order to subtract from output

all data

200 km²

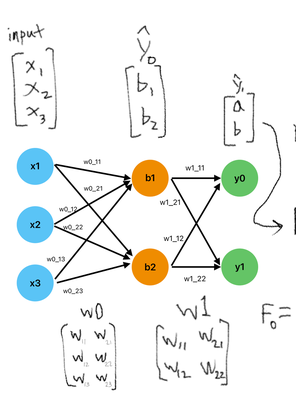
Region

only data within region of output

precise coordinates

Model v1.4 3x2x2 → 4x3x2

train model using samples within this coordinate range



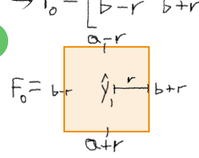
$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \qquad \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} \qquad \begin{bmatrix} a \\ b \end{bmatrix}$$

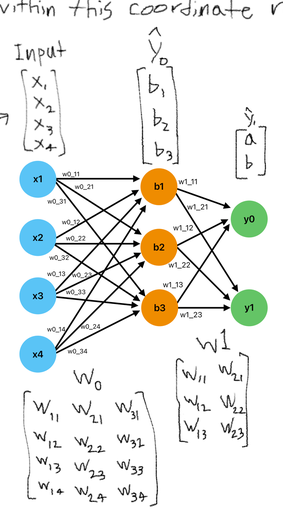
$$W0 = \begin{bmatrix} w_{11} & w_{21} \\ w_{12} & w_{22} \\ w_{13} & w_{23} \end{bmatrix} \qquad W1 = \begin{bmatrix} w_{11} & w_{21} \\ w_{12} & w_{22} \end{bmatrix}$$

Radius (m)

$r = 100000$

$$F_0 = \begin{bmatrix} a-r & a+r \\ b-r & b+r \end{bmatrix}$$

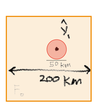




$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} \qquad \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} \qquad \begin{bmatrix} a \\ b \end{bmatrix}$$

$$W_0 = \begin{bmatrix} w_{11} & w_{21} & w_{31} \\ w_{12} & w_{22} & w_{32} \\ w_{13} & w_{23} & w_{33} \\ w_{14} & w_{24} & w_{34} \end{bmatrix} \qquad W1 = \begin{bmatrix} w_{11} & w_{21} \\ w_{12} & w_{22} \\ w_{13} & w_{23} \end{bmatrix}$$

$F_1 =$



$$W_0^T \cdot \text{Input} = \begin{bmatrix} w_{11} & w_{12} & w_{13} & w_{14} \\ w_{21} & w_{22} & w_{23} & w_{24} \\ w_{31} & w_{32} & w_{33} & w_{34} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} x_1 w_{11} + x_2 w_{12} + x_3 w_{13} + x_4 w_{14} \\ x_1 w_{21} + x_2 w_{22} + x_3 w_{23} + x_4 w_{24} \\ x_1 w_{31} + x_2 w_{32} + x_3 w_{33} + x_4 w_{34} \end{bmatrix} = \hat{y}_0 \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$$

$$\hat{y}_0 \cdot W_1 = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} \begin{bmatrix} w_{11} & w_{21} \\ w_{12} & w_{22} \end{bmatrix} = \begin{bmatrix} b_1 w_{11} + b_2 w_{21} \\ b_1 w_{12} + b_2 w_{22} \end{bmatrix} = \begin{bmatrix} y_0 \\ y_1 \end{bmatrix} = \hat{y}_1$$