

MIA-07 Redes Neuronales y Aprendizaje Profundo

Sección A, Grupo 2

Integrantes:

Bernie Hans, Benitez A. **Ramirez Ucañay, Barbarita P. J.** **Siu Siu Ting, Aldo D.**
bernie.benitez.a@uni.pe barbarita.ramirez.u@uni.pe aldo.siu.s@uni.pe

Diaz Cabrera, Alexander G. **Morales Ccasa, Geyson D.**
alexandergabrieldiazcabrera@gmail.com geyson.morales.c@uni.pe

Problema 1:

1. Hacer las operaciones de forward y backward propagation de forma manual para un MLP (perceptron multicapa) que tome 3 entradas (3x1), tenga 2 hidden layers de tamaño 4, y una salida de 1x1. El cálculo lo deben hacer para la siguiente data:

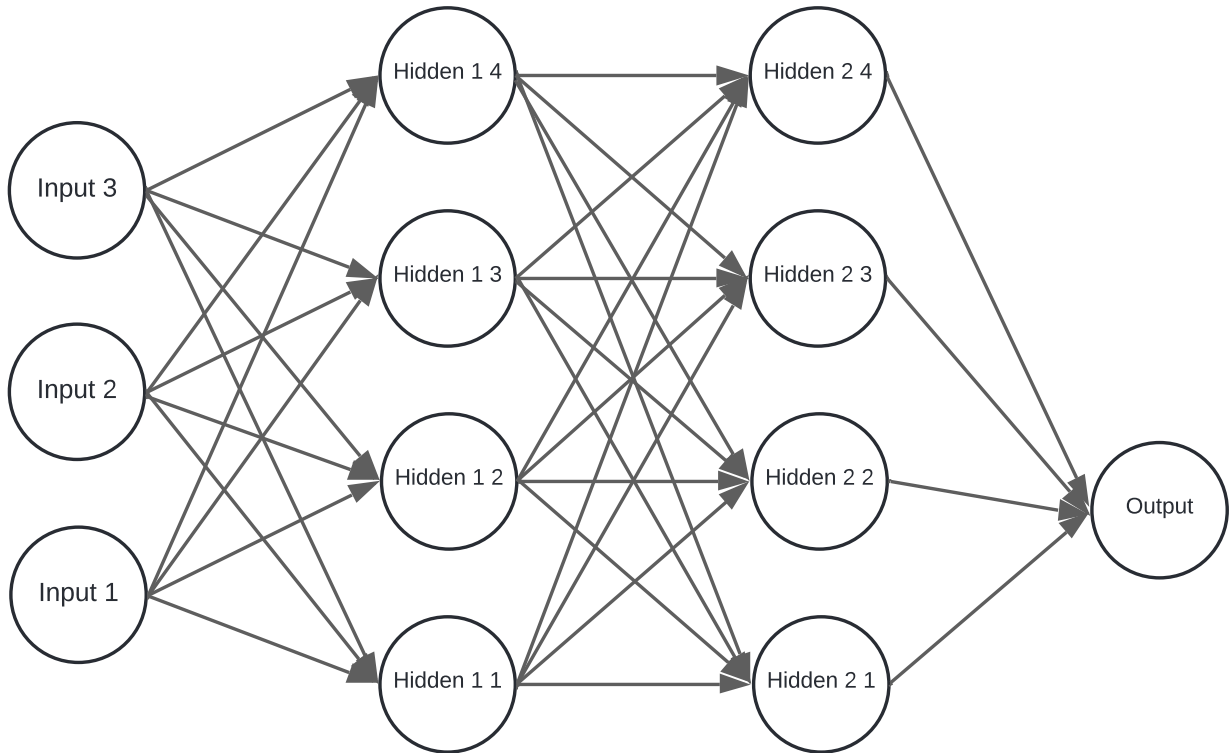
La data de entrada \mathbf{X}_s y el vector deseado es el siguiente y son:

$$\mathbf{X}_s = \begin{bmatrix} 2.5 & 3.0 & -2.5 \\ 4.0 & -1.0 & 0.5 \\ 0.5 & 1.5 & 1.0 \\ 3.0 & 2.0 & -2.0 \end{bmatrix} \quad \mathbf{y}_s = \begin{bmatrix} 1.0 \\ -1.0 \\ 1.0 \\ 1.0 \end{bmatrix} \quad (1)$$

Solución:

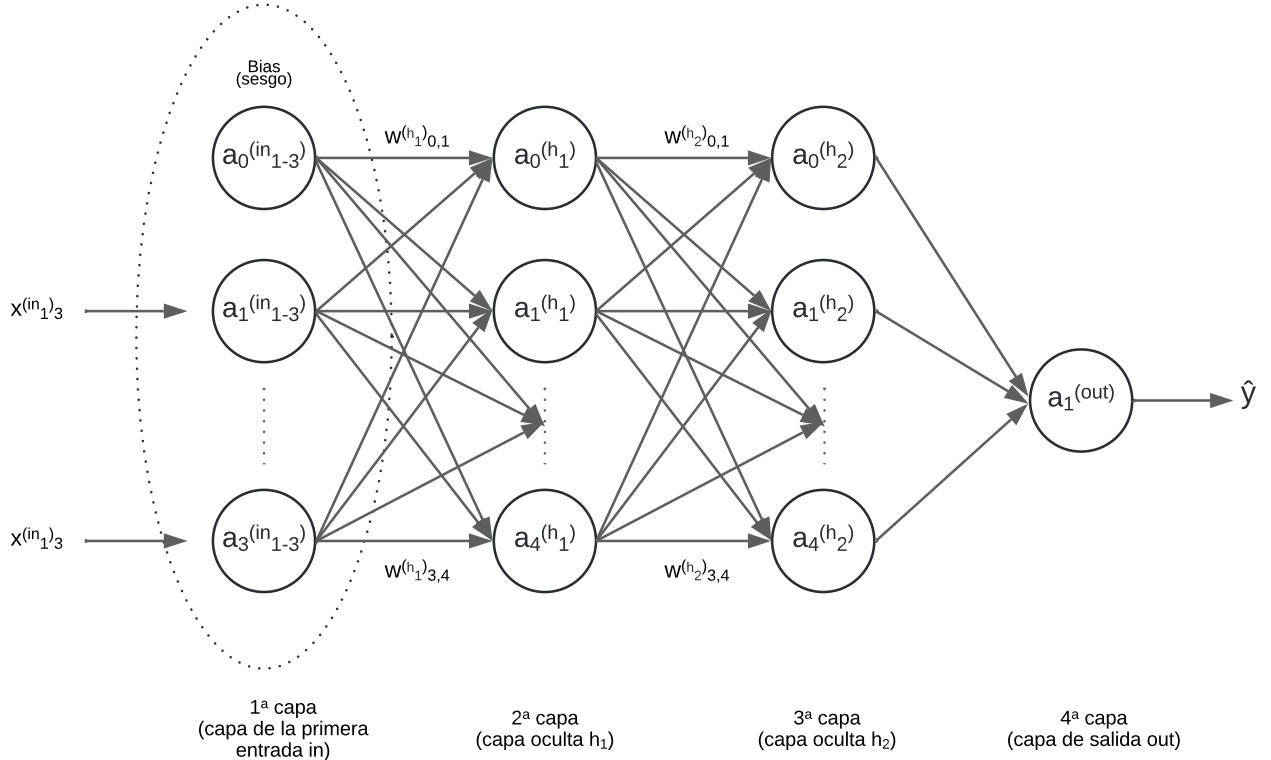
Según los datos y salida esperada, el MLP (Multilayer Perceptron) es representado en la siguiente imagen:

Figure 1: Arquitectura de la red neuronal artificial profunda (MPL) con 3 entradas, 2 capas de 4 nodos, y 1 salida (3-4-4-1).



Fuente: Elaboración propia.

Figure 2: Arquitectura de la red neuronal artificial profunda.



Fuente: Elaboración propia.

Forward Propagation:

Inicialización de pesos y sesgos:

Definimos los pesos y sesgos de las capas. Como no se especifican en el plantiamiento del problema, los inicializaremos de forma aleatoria (el generado en la implementación del del problema).

Capa 1 (de entrada a primera capa oculta):

- Tamaño de W_1 : 4×3 (4 neuronas, 3 entradas).
- Tamaño de b_1 : 4×1 (un sesgo para cada neurona de la primera capa).

Pesos de la primera capa oculta (W_1):

$$W_1 = \begin{bmatrix} -0.7731 & 0.1601 & 0.4321 \\ -0.0346 & -0.2008 & 0.3647 \\ -0.7577 & 0.4841 & -0.0354 \\ 0.1601 & -0.2008 & 0.4841 \end{bmatrix}$$

Sesgos de la primera capa oculta (b_1):

$$b_1 = \begin{bmatrix} -0.0132 \\ 0.7161 \\ -0.4399 \\ 0.5371 \end{bmatrix} \quad (2)$$

Capa 2 (de primera capa oculta a segunda capa oculta):

- Tamaño de W_2 : 4×4 (4 neuronas en ambas capas).
- Tamaño de b_1 : 4×1 .

Pesos de la segunda capa oculta (W_2):

$$W_2 = \begin{bmatrix} -0.4812 & -0.3530 & -0.4004 & -0.0734 \\ 0.7681 & -0.3188 & 0.5793 & 0.9179 \\ 0.1169 & -0.2848 & -0.4631 & 0.1716 \\ 0.8809 & 0.6788 & 0.2416 & 0.3609 \end{bmatrix}$$

Sesgos de la primera capa oculta (b_2):

$$b_2 = \begin{bmatrix} 0.1427 \\ -0.7043 \\ -0.8546 \\ 0.4936 \end{bmatrix}$$

Capa de salida:

Pesos de la capa de salida (W_3):

$$W_3 = [-0.2537 \quad -0.5915 \quad -0.0293 \quad -0.3593]$$

Sesgo de la capa de salida (b_3):

$$W_3 = [0.3446] \quad (3)$$

Paso 1: Cálculo en la primera capa oculta

$$Z_1 = X_s \cdot W_1^T + b_1 \quad , A_1 = \text{activación}(Z_1) \quad (4)$$

$$\begin{aligned} Z_1 &= \begin{bmatrix} 2.5 & 3.5 & -0.5 \\ 4.0 & -1.0 & 0.5 \\ 0.5 & 1.5 & 1.0 \\ 3.0 & 2.0 & -1.5 \end{bmatrix} \cdot \begin{bmatrix} -0.7731 & 0.1601 & 0.4321 \\ -0.0346 & -0.2008 & 0.3647 \\ -0.7577 & 0.4841 & -0.0354 \end{bmatrix}^T \\ &+ \begin{bmatrix} -0.0132 \\ 0.7161 \\ -0.4399 \\ 0.5371 \end{bmatrix} = \begin{bmatrix} -1.6017 & -0.2556 & -0.6211 & -0.0075 \\ -3.0496 & 0.9609 & -3.9725 & 1.6204 \\ 0.2725 & 0.7623 & -0.1280 & 0.8001 \\ -2.6605 & -0.3363 & -1.6917 & -0.1104 \end{bmatrix} \end{aligned}$$

Aplicando la función de activación ReLU:

$$\text{ReLU}(x) = \max(0, x) \quad (5)$$

Esto significa que:

- Si el valor de entrada x es mayor que 0, la salida es x .
- Si el valor de entrada x es menor o igual a 0, la salida es 0.

$$A_1 = \text{ReLU}(Z_1) = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0.9609 & 0 & 1.6204 \\ 0.2725 & 0.7623 & 0 & 0.8001 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Paso 2: Cálculo en la segunda capa oculta

- Ahora multiplicamos la salida de la primera capa A_1 (4x4) con la matriz de pesos de la segunda capa W_2 (4x4).
- Al igual que en la primera capa, realizamos una multiplicación de matrices estándar: cada fila de A_1 se multiplica por la columna correspondiente de W_2 , sumando los productos.

$$Z_2 = A_1 \cdot W_2^T + b_2 \quad , A_2 = \text{activación}(Z_2) \quad (6)$$

$$Z_2 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0.9609 & 0 & 1.6204 \\ 0.2725 & 0.7623 & 0 & 0.8001 \\ 0 & 0 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} -0.4812 & -0.3530 & -0.4004 & -0.0734 \\ 0.7681 & -0.3188 & 0.5793 & 0.9179 \\ 0.1169 & -0.2848 & -0.4631 & 0.1716 \\ 0.8809 & 0.6788 & 0.2416 & 0.3609 \end{bmatrix}^T + \begin{bmatrix} 0.1427 \\ -0.7043 \\ -0.8546 \\ 0.4936 \end{bmatrix}$$

$$Z_2 = \begin{bmatrix} 0.1427 & -0.7043 & -0.8546 & 0.4936 \\ -0.3155 & 0.4770 & -0.8506 & 1.7305 \\ 0.1427 & -0.7043 & -0.8546 & 0.4936 \\ 0.1427 & -0.7043 & -0.8546 & 0.4936 \end{bmatrix}$$

$$A_2 = \text{ReLU}(Z_2) = \begin{bmatrix} 0.1427 & 0 & 0 & 0.4936 \\ 0 & 0.4770 & 0 & 1.7305 \\ 0 & 0 & 0 & 1.5402 \\ 0.1427 & 0 & 0 & 0.4936 \end{bmatrix}$$

Paso 3: Cálculo para la capa de salida:

$$Z_3 = A_2 \cdot W_3^T + b_3 \quad , A_3 = \text{activación}(Z_2) \quad (7)$$

$$Z_3 = Z_{\text{out}} = \begin{bmatrix} 0.1427 & 0 & 0 & 0.4936 \\ 0 & 0.4770 & 0 & 1.7305 \\ 0 & 0 & 0 & 1.5402 \\ 0.1427 & 0 & 0 & 0.4936 \end{bmatrix} \cdot \begin{bmatrix} -0.2537 \\ -0.5915 \\ -0.0293 \\ -0.3593 \end{bmatrix}^T + 0.3446$$

$$Z_3 = \begin{bmatrix} 0.1310 \\ -0.5592 \\ -0.2089 \\ 0.1310 \end{bmatrix}$$

Para regresión:

$$y_{\text{pred}} = Z_{\text{out}}$$

$$y_{\text{pred}} = \begin{bmatrix} 0.1310 \\ -0.5592 \\ -0.2089 \\ 0.1310 \end{bmatrix}$$