



3. Dado el modelo lineal:

$$g_{\theta}(x) = \theta_1 x + \theta_0, \quad (2)$$

y la función de pérdida:

$$\mathcal{L}(\theta, x_i, y_i) = \frac{1}{n} \sum_{k=1}^n (g_{\theta}(x_i) - y_i)^2 \quad (3)$$

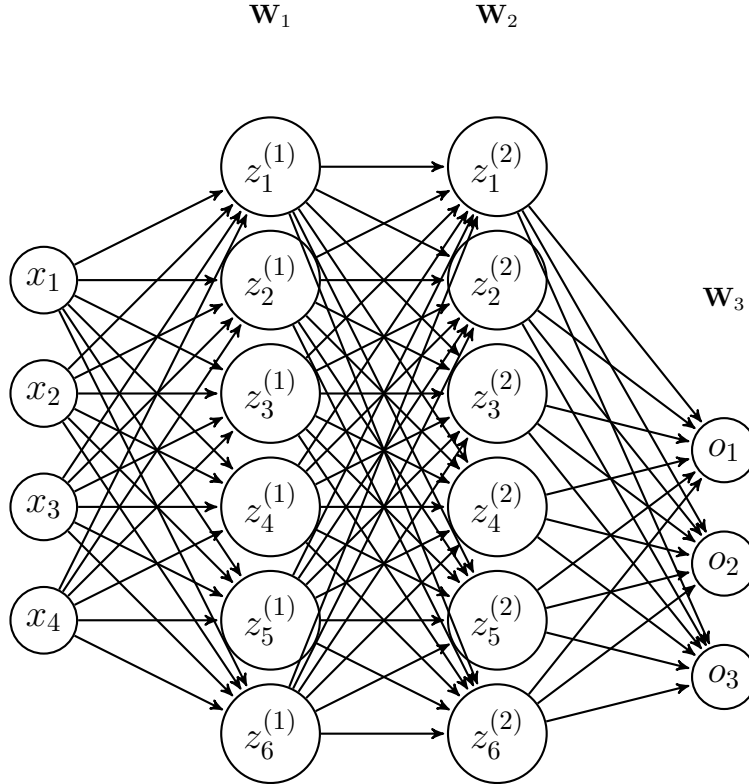
y dados los siguientes parámetros de inicialización:

- $\alpha = 0.1$
- $\theta_0 = 0$
- $\theta_1 = 0$
- (1pt) Escriba las ecuaciones de Gradiente Descendiente, calcule las derivadas de forma explícita.
- (1pt) Realice una iteración del modelo de gradiente descendente para todo el conjunto de entrenamiento (no estocástico):

x	y
1	1
2	0.5
3	2

Tabela 1: Datos de entrenamiento

4. Dado el siguiente diagrama de una red neuronal donde  $x_j$  con  $j = 1, \dots, 4$  siendo las componentes de entrada o información de entrada y además, donde el sesgo o bias es igual a cero para cada capa, es decir  $\mathbf{b}_i = 0, \forall i = 1, \dots, \text{numero de capas}$ :



Responda las siguientes preguntas:

- (a) (2pt) Cuántas capas tiene este modelo.
- (b) (2pt) Cuántos parámetros tiene este modelo.
- (c) (2pt) Sea  $z_i^{(j)}$  la salida de la neurona  $i$  en la capa  $j$ . Cuál es la dimensión de las matrices involucradas si los modelos dentro de cada potencial de activación son modelos lineales de la forma  $W_i \mathbf{x} + \mathbf{b}_i$ .
- (d) (2pt) Escriba las ecuaciones que representan a este modelo (ya sea de forma anidada o como un sistema de ecuaciones).

5. En relación a las redes convolucionales:

- (a) (2pt) Describa la arquitectura de una red neuronal convolucional (CNN) utilizada para la clasificación de imágenes. Utilice un dibujo de ser necesario.
- (b) (2pt) Describa brevemente los componentes clave como capas convolucionales, capas de pooling, capas completamente conectadas y el propósito de cada una. Además, explique cómo el tamaño de la imagen de entrada afecta a la arquitectura.
- (c) (1pt) Explique con un ejemplo el concepto fundamental de convolución en una CNN.
- (d) (1pt) ¿Por qué sirven para trabajar con audio?