



Universidad Nacional Autónoma de México

Facultad de Ciencias

Redes Neuronales

Examen 01

Carlos Emilio Castañón Maldonado



1 De los siguientes mecanismos para procesar información en una neurona biológica, indique cuales inspiraron qué elemento de las redes neuronales artificiales:

- Potenciales de acción ➤ Uso de una función de activación para decidir si se envía información a las siguientes neuronas.
- Cono axónico ➤ Cálculo de combinaciones lineales a partir de los valores de activación de las neuronas entrantes.
- Campos receptivos ➤ Diseño de arquitecturas de redes para problemas específicos, como reconocimiento de imágenes.
- Neuroplasticidad ➤ Uso de un algoritmo de entrenamiento que permitiera incrementar las magnitudes de los pesos que favorecen la obtención de respuestas correctas y reducción de los que producen errores.

2 Asocie la medida de rendimiento de un algoritmo de aprendizaje de máquina con su definición.

- Exactitud (Accuracy) ➤ Indica la proporción de respuestas correctas que dio la red.
- Precisión ➤ Medición que denunciaría la proporción de personas que recibieron un tratamiento riesgoso, sin que lo necesitaran, contra quienes sí lo necesitaban y lo recibieron correctamente.
- Recuperación (Recall) ➤ Medición que denunciaría la proporción de pacientes que lograron recibir el tratamiento que necesitaban por haber sido diagnosticados, contra quienes se quedaron sin él.
- Especificidad ➤ Dado el total de ejemplares negativos, refleja cuántos fueron reconocidos correctamente.
- F ➤ Indica la capacidad promedio para identificar positivos correctamente sin aceptar elementos de más, o dejar de detectar ejemplares por ser demasiado estricto.

3 Dados los ejemplares siguientes de aprendizaje de máquina

Indicar de qué tipo de aprendizaje se trata:

Tip: Esto tiene que ver con el *clustering*.

Supervisado	Clasificación	Entrenar a un sistema cibernético para mantener la temperatura indicada en una habitación, debiendo elegir entre encender la calefacción o la ventilación.
Por refuerzo	Regresión	Entrenar a un submarino para que encuentre géisers otorgándole puntos cuando los encuentra. El vehículo debe elegir en qué dirección moverse y con qué velocidad de forma continua.
No supervisado	Agrupación (clustering)	Organizar las facturas de una empresa, si no se cuenta con información a priori.
Supervisado	Clasificación	Dado un dibujo, identificar qué objeto representa.

4 Indica la veracidad o falsedad de los siguientes enunciados

	Verdadero	Falso
Las conexiones definidas entre los perceptrones no influyen en nuestro espacio de hipótesis		★
La modificación de los pesos en una red multicapa modifica el espacio de hipótesis		★
La elección del método de entrenamiento para una red multicapa no afecta el espacio de hipótesis, siempre y cuando la arquitectura de la red y la función de pérdida se mantengan fijas.	★	
La cantidad de perceptrones que se decide usar en la capa oculta modifica el espacio de hipótesis	★	

5 Contesta las siguientes preguntas:

- 5 a) ¿Cuál es el conjunto que se utiliza para seleccionar el nivel de regularización que se aplicará en la función de error?
➤ Validación
- 5 b) ¿Cuál es el conjunto que se utiliza para seleccionar una hipótesis de nuestro espacio del conjunto de hipótesis?
➤ Entrenamiento
- 5 c) ¿Cuál es el conjunto que se utiliza para evaluar a la red simulando un ambiente de producción con datos nunca antes vistos?
➤ Prueba

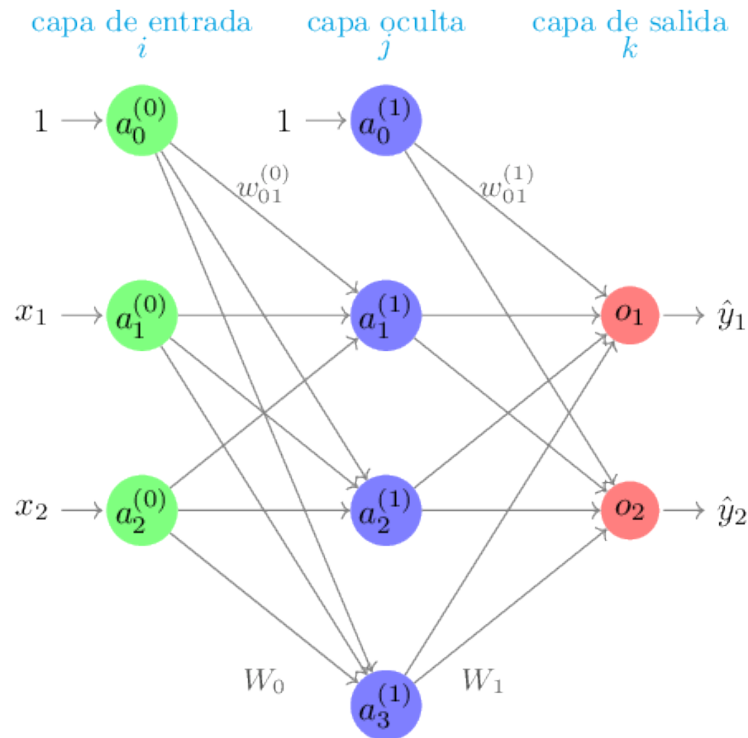
6 Indique el o los elementos de la red correspondientes.

Datos de entrada X y valores deseados Y	Constantes de las cuales depende el valor de la función de error durante un paso de entrenamiento.
Pesos Θ	Parámetros que se modificarán para entrenar a la red.
Hipótesis $H_{\Theta}(X)$	Es la función que calcula la red.
El error o pérdida $J(\Theta)$	Es la función que se desea optimizar cuando se entrena una red neuronal.
El gradiente $\nabla J_{\Theta}(X)$	Es la función calculada por el algoritmo de propagación hacia atrás.

7 Indique si las afirmaciones siguientes sobre redes neuronales son verdaderas o falsas.

	Verdadero	Falso
Se sabe que una red neuronal de tres capas puede aproximar cualquier función si tiene suficientes neuronas en la capa oculta, pero no se sabe cómo calcular los pesos.		★
Los sesgos se utilizan cuando la función que se desea aproximar puede tener un valor distinto de cero incluso si todos los valores de las características de entrada son ceros.	★	
Cuando los sesgos se añaden como una neurona extra se entrenan más parámetros que cuando se suman después de haber calculado la combinación lineal de los valores de entrada a cada neurona.		★
Una red neuronal no se puede usar/evaluar sin el algoritmo de retropropagación.		★

- 8 Dada la red siguiente, coloca qué peso se coloca en qué posición de las matrices (en las matrices intencionalmente sobran espacios, pon el guión bajo en las casillas vacías):



Para W_0 y B_0 :

$$W_0 = \begin{bmatrix} w_{11} & w_{12} & w_{13} \\ w_{21} & w_{22} & w_{23} \\ - & - & - \end{bmatrix}$$

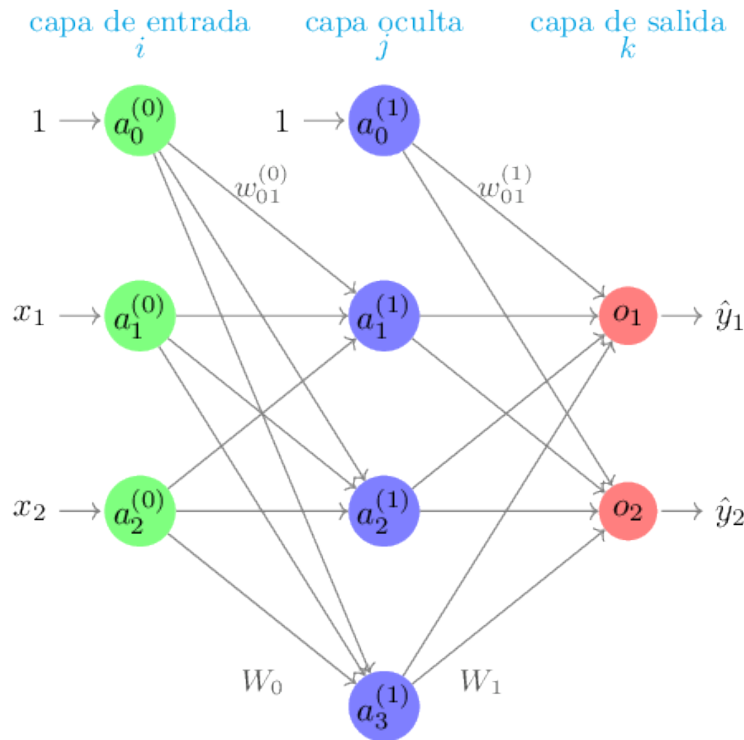
$$B_0 = \begin{bmatrix} w_{01} & w_{02} & w_{03} \\ - & - & - \\ - & - & - \end{bmatrix}$$

Para W_1 y B_1 :

$$W_1 = \begin{bmatrix} w_{11} & w_{12} & - \\ w_{21} & w_{22} & - \\ w_{31} & w_{32} & - \end{bmatrix}$$

$$B_1 = \begin{bmatrix} w_{01} & w_{02} & - \\ - & - & - \\ - & - & - \end{bmatrix}$$

9 Dada la siguiente red neuronal:



Calcula lo siguiente:

- 9 a) Dada la red neuronal, calcule sus valores de activación si $x_1 = 1$ y $x_2 = 0$, utilizando la función de activación sigmoide.

Pesos de la primera capa:

$$w_{01} = 0.4$$

$$w_{02} = 0.6$$

$$w_{03} = 0.4$$

$$w_{11} = 0.9$$

$$w_{12} = 0.2$$

$$w_{13} = 0.7$$

$$w_{21} = 0.8$$

$$w_{22} = 0.3$$

$$w_{23} = 0.6$$

Dado lo anterior, las entradas son:

$$x_1 = 1 \quad \text{y} \quad x_2 = 0$$

Los pesos de la primera capa son:

$$w_{01} = -0.4, \quad w_{02} = -0.6, \quad w_{03} = 0.4$$

$$w_{11} = -0.9, \quad w_{12} = 0.2, \quad w_{13} = -0.7$$

$$w_{21} = 0.8, \quad w_{22} = 0.3, \quad w_{23} = 0.6$$

Calculamos los valores de z y a para cada neurona en la capa oculta.

Cálculo de $z_1^{(1)}$ y $a_1^{(1)}$:

$$z_1^{(1)} = w_{01} \cdot 1 + w_{11} \cdot x_1 + w_{21} \cdot x_2$$

$$z_1^{(1)} = (-0.4) \cdot 1 + (-0.9) \cdot 1 + 0.8 \cdot 0$$

$$z_1^{(1)} = -0.4 - 0.9 + 0 = -1.3$$

Aplicando la función de activación sigmoide:

$$a_1^{(1)} = \sigma(z_1^{(1)}) = \frac{1}{1 + e^{-z_1^{(1)}}}$$

$$a_1^{(1)} = \frac{1}{1 + e^{1.3}} \approx \frac{1}{1 + 3.67} \approx \frac{1}{4.67} \approx 0.214$$

Cálculo de $z_2^{(1)}$ y $a_2^{(1)}$:

$$z_2^{(1)} = w_{02} \cdot 1 + w_{12} \cdot x_1 + w_{22} \cdot x_2$$

$$z_2^{(1)} = (-0.6) \cdot 1 + 0.2 \cdot 1 + 0.3 \cdot 0$$

$$z_2^{(1)} = -0.6 + 0.2 + 0 = -0.4$$

Aplicando la función de activación sigmoide:

$$a_2^{(1)} = \sigma(z_2^{(1)}) = \frac{1}{1 + e^{-z_2^{(1)}}}$$

$$a_2^{(1)} = \frac{1}{1 + e^{0.4}} \approx \frac{1}{1 + 1.49} \approx \frac{1}{2.49} \approx 0.401$$

Cálculo de $z_3^{(1)}$ y $a_3^{(1)}$:

$$z_3^{(1)} = w_{03} \cdot 1 + w_{13} \cdot x_1 + w_{23} \cdot x_2$$

$$z_3^{(1)} = 0.4 \cdot 1 + (-0.7) \cdot 1 + 0.6 \cdot 0$$

$$z_3^{(1)} = 0.4 - 0.7 + 0 = -0.3$$

Aplicando la función de activación sigmoide:

$$a_3^{(1)} = \sigma(z_3^{(1)}) = \frac{1}{1 + e^{-z_3^{(1)}}}$$

$$a_3^{(1)} = \frac{1}{1 + e^{0.3}} \approx \frac{1}{1 + 1.35} \approx \frac{1}{2.35} \approx 0.426$$

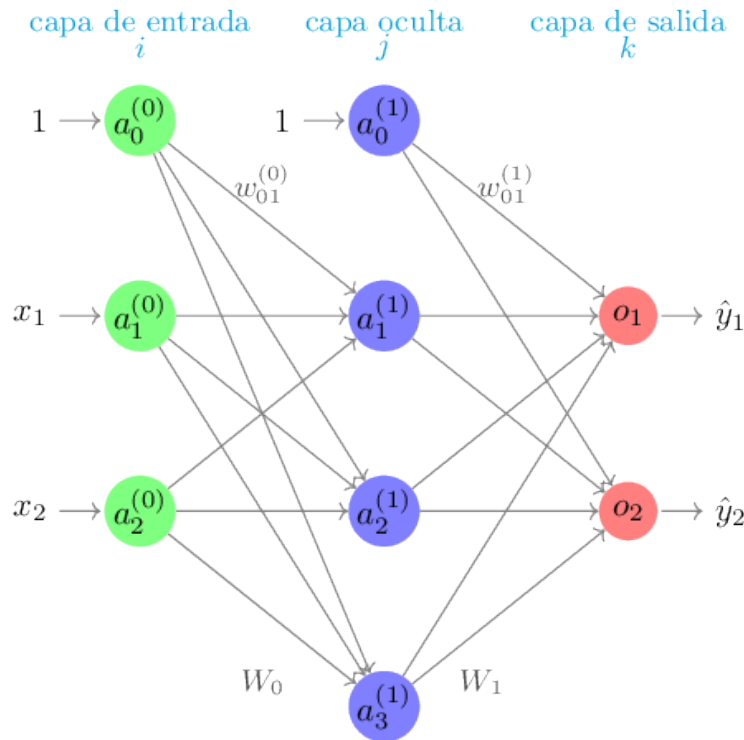
∴ Los valores de activación son:

$$z_1^{(1)} = -1.3, \quad a_1^{(1)} = 0.214$$

$$z_2^{(1)} = -0.4, \quad a_2^{(1)} = 0.401$$

$$z_3^{(1)} = -0.3, \quad a_3^{(1)} = 0.426$$

10 Dada la siguiente red neuronal:



Calcula lo siguiente:

10 a) Dada la red neuronal, calcula el gradiente si $x_1 = 1$ y $x_2 = 1$ y las salidas deseadas son $y_1 = 0$ y $y_2 = 1$ utilizando la función de activación sigmoide, calcula el error.

Pesos de la primera capa:

$$w_{01} = 0.6$$

$$w_{02} = 0.7$$

$$w_{03} = 0.2$$

$$w_{11} = 0.4$$

$$w_{12} = 0.2$$

$$w_{13} = 0.1$$

$$w_{21} = 0.2$$

$$w_{22} = 0.9$$

$$w_{23} = 0.7$$

Pesos de la segunda capa:

$$w_{01} = 0.5$$

$$w_{02} = 0.9$$

$$w_{11} = 0$$

$$w_{12} = 0.4$$

$$w_{21} = 0.4$$

$$w_{22} = 0.2$$

$$w_{311} = 0.2$$

$$w_{32} = 0.3$$

TIP: Recuerda que evaluar un solo ejemplar es un caso particular de evaluar para varios ejemplares, además de usar entropía binaria como función de error.

Realizamos la propagación hacia adelante

Tomando en cuenta la, Función de activación sigmoide:

$$\sigma(z) = \frac{1}{1 + e^{-z}}$$

Realizamos el cálculo de las activaciones en la capa oculta:

$$a_0^{(0)} = 1, \quad x_1 = 1, \quad x_2 = 1$$

Calculamos $z_1^{(1)}$:

$$\begin{aligned} z_1^{(1)} &= w_{01}^{(1)} a_0^{(0)} + w_{11}^{(1)} x_1 + w_{21}^{(1)} x_2 \\ &= -0.6 \cdot 1 + 0.4 \cdot 1 + 0.2 \cdot 1 \\ &= -0.6 + 0.4 + 0.2 \\ &= 0 \end{aligned}$$

Activación $a_1^{(1)}$:

$$a_1^{(1)} = \sigma(z_1^{(1)}) = \sigma(0) = 0.5$$

Calculamos $z_2^{(1)}$:

$$\begin{aligned} z_2^{(1)} &= w_{02}^{(1)} a_0^{(0)} + w_{12}^{(1)} x_1 + w_{22}^{(1)} x_2 \\ &= 0.7 \cdot 1 + (-0.2) \cdot 1 + (-0.9) \cdot 1 \\ &= 0.7 - 0.2 - 0.9 \\ &= -0.4 \end{aligned}$$

Activación $a_2^{(1)}$:

$$a_2^{(1)} = \sigma(z_2^{(1)}) = \sigma(-0.4) = \frac{1}{1 + e^{0.4}} \approx 0.401$$

Calculamos $z_3^{(1)}$:

$$\begin{aligned} z_3^{(1)} &= w_{03}^{(1)} a_0^{(0)} + w_{13}^{(1)} x_1 + w_{23}^{(1)} x_2 \\ &= 0.2 \cdot 1 + 0.1 \cdot 1 + 0.7 \cdot 1 \\ &= 0.2 + 0.1 + 0.7 \\ &= 1 \end{aligned}$$

Activación $a_3^{(1)}$:

$$a_3^{(1)} = \sigma(z_3^{(1)}) = \sigma(1) = \frac{1}{1 + e^{-1}} \approx 0.731$$

Cálculo de las activaciones en la capa de salida:

Calculamos $z_1^{(2)}$:

$$\begin{aligned} z_1^{(2)} &= w_{01}^{(2)} a_0^{(1)} + w_{11}^{(2)} a_1^{(1)} + w_{21}^{(2)} a_2^{(1)} + w_{31}^{(2)} a_3^{(1)} \\ &= -0.5 \cdot 1 + 0 \cdot 0.5 + (-0.4) \cdot 0.401 + 0.2 \cdot 0.731 \\ &= -0.5 + 0 - 0.1604 + 0.1462 \\ &= -0.5142 \end{aligned}$$

Activación $a_1^{(2)}$:

$$a_1^{(2)} = \sigma(z_1^{(2)}) = \sigma(-0.5142) = \frac{1}{1 + e^{0.5142}} \approx 0.374$$

Calculamos $z_2^{(2)}$:

$$\begin{aligned} z_2^{(2)} &= w_{02}^{(2)} a_0^{(1)} + w_{12}^{(2)} a_1^{(1)} + w_{22}^{(2)} a_2^{(1)} + w_{32}^{(2)} a_3^{(1)} \\ &= 0.9 \cdot 1 + 0.4 \cdot 0.5 + (-0.2) \cdot 0.401 + 0.3 \cdot 0.731 \\ &= 0.9 + 0.2 - 0.0802 + 0.2193 \\ &= 1.2391 \end{aligned}$$

Activación $a_2^{(2)}$:

$$a_2^{(2)} = \sigma(z_2^{(2)}) = \sigma(1.2391) = \frac{1}{1 + e^{-1.2391}} \approx 0.775$$

Cálculo del Error: Usamos la función de error de entropía binaria:

$$y_1 = 0, \quad y_2 = 1$$

$$J = - \left[y_1 \log(a_1^{(2)}) + (1 - y_1) \log(1 - a_1^{(2)}) + y_2 \log(a_2^{(2)}) + (1 - y_2) \log(1 - a_2^{(2)}) \right]$$

$$J = - [0 \cdot \log(0.374) + (1 - 0) \cdot \log(1 - 0.374) + 1 \cdot \log(0.775) + (1 - 1) \cdot \log(1 - 0.775)]$$

$$J = - [\log(0.626) + \log(0.775)]$$

$$J = - [-0.467 + (-0.255)]$$

$$J \approx 0.722$$

Propagación hacia atrás:

Calculamos $\delta^{(2)}$:

$$\delta_k^{(2)} = a_k^{(2)} - y_k$$

$$\delta_1^{(2)} = 0.374 - 0 = 0.374$$

$$\delta_2^{(2)} = 0.775 - 1 = -0.225$$

Calculamos $\delta^{(1)}$:

$$\delta_j^{(1)} = \left(\sum_k \delta_k^{(2)} w_{jk}^{(2)} \right) \sigma'(z_j^{(1)})$$

Para $\sigma'(z)$:

$$\sigma'(z) = \sigma(z)(1 - \sigma(z))$$

$$\begin{aligned} \delta_1^{(1)} &= \left(\delta_1^{(2)} w_{11}^{(2)} + \delta_2^{(2)} w_{12}^{(2)} \right) \sigma'(z_1^{(1)}) \\ &= (0.374 \cdot 0 + (-0.225) \cdot 0.4) \sigma'(0) \\ &= (-0.090) \cdot 0.25 \\ &= -0.0225 \end{aligned}$$

$$\begin{aligned} \delta_2^{(1)} &= \left(\delta_1^{(2)} w_{21}^{(2)} + \delta_2^{(2)} w_{22}^{(2)} \right) \sigma'(z_2^{(1)}) \\ &= (0.374 \cdot (-0.4) + (-0.225) \cdot (-0.2)) \sigma'(-0.4) \\ &= (-0.1496 + 0.045) \cdot 0.241 \\ &= -0.1046 \cdot 0.241 \\ &= -0.025 \end{aligned}$$

$$\begin{aligned} \delta_3^{(1)} &= \left(\delta_1^{(2)} w_{31}^{(2)} + \delta_2^{(2)} w_{32}^{(2)} \right) \sigma'(z_3^{(1)}) \\ &= (0.374 \cdot 0.2 + (-0.225) \cdot 0.3) \sigma'(1) \\ &= (0.075 - 0.067) \cdot 0.196 \\ &= 0.008 \cdot 0.196 \\ &= 0.0015 \end{aligned}$$

Gradientes de la primera capa:

$$\frac{\partial J}{\partial w_{ij}^{(1)}} = a_i^{(0)} \delta_j^{(1)}$$

Para $w_{11}^{(1)}$:

$$\frac{\partial J}{\partial w_{11}^{(1)}} = x_1 \cdot \delta_1^{(1)} = 1 \cdot -0.0225 = -0.0225$$

Resultados finales:

$$J \approx 0.722$$

$$\delta_1^{(2)} \approx 0.374$$

$$\frac{\partial J}{\partial w_{11}^{(1)}} \approx -0.0225$$