

El aprendizaje competitivo

El aprendizaje competitivo, en el contexto de las redes neuronales, se refiere a una técnica de entrenamiento en la que las neuronas “compiten” entre sí para activarse y aprender representaciones de los datos de entrada. Este tipo de red no supervisada es monocapa (ver Fig 1). Como en el entrenamiento no se presentan salidas objetivas la red, a través de un proceso de autoorganización, debe descubrir por sí misma regularidades y asociaciones en los datos.

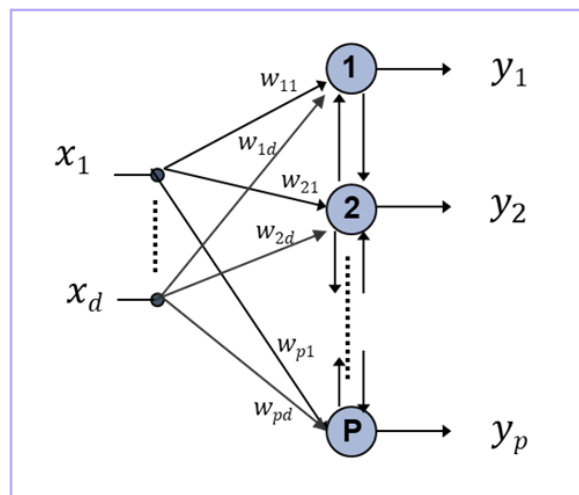


Fig. 1. Red de aprendizaje competitivo.

¿Cómo funciona? El proceso es el siguiente:

1. Cada neurona i calcula la similitud entre el vector de entrada \mathbf{x}_i y su vector de pesos \mathbf{w}_i , utilizando una medida de distancia o criterio de similitud.
2. A continuación, se determina la neurona cuyo vector de pesos es más similar al vector de entrada.

$$d(\mathbf{w}_g, \mathbf{x}) = \min \{d(\mathbf{w}_i, \mathbf{x})\}, i = 1..p$$

Esta se conoce como “neurona ganadora” (denotada por g) y es la que ofrece un mayor nivel de activación para el vector de entrada. La información del nivel de activación se “pasa” a través de las conexiones laterales a través de la función de activación competitiva.

3. Como es una competición, la neurona ganadora recibe un “premio”: el refuerzo de sus conexiones. Para otorgar el “premio” se utiliza la siguiente regla de aprendizaje:

$$w_j(t+1) = w_j(t) + \alpha (x_i - w_j(t)), \text{ si } j = g$$

Con la regla, se “mueve” el vector de pesos de la ganadora hacia el vector de entrada, en una pequeña cantidad que depende de la constante de aprendizaje. La idea es actualizar los pesos de la neurona ganadora para aproximarla más al patrón de entrada. Los pesos del resto de las neuronas no se actualizan.

Veamos entonces cómo sería el entrenamiento de una red de aprendizaje competitivo. Utilizaremos los datos mostrados en la Fig 2. La red para utilizar tiene dos entradas y una capa de salida con tres neuronas.

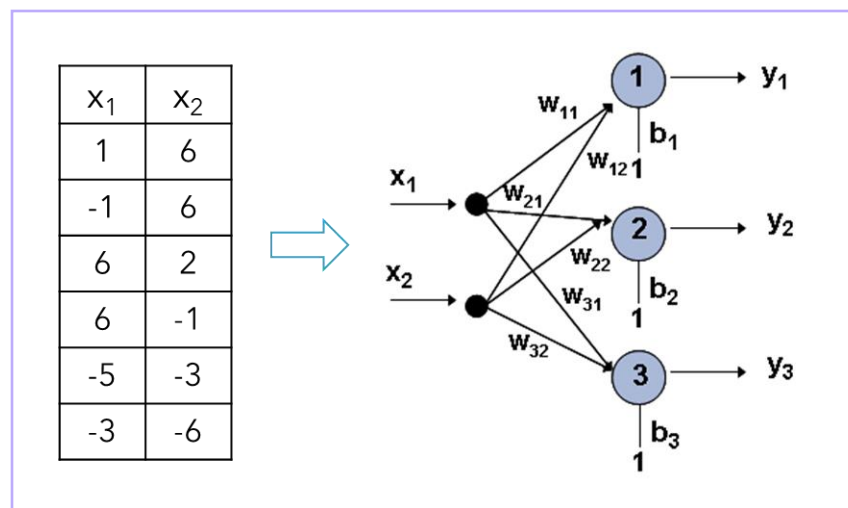


Fig. 2. Conjunto de datos de dos dimensiones y la red competitiva que será utilizada para el aprendizaje.

Como primer paso se inicializa la red con valores aleatorios y se fija la constante de aprendizaje. Se utilizarán los siguientes valores:

$$\mathbf{W}(0) = \begin{bmatrix} 4 & -6 \\ -6 & 0 \\ 4 & 5 \end{bmatrix}, \quad \alpha = 0.5$$

A continuación, se sigue el siguiente procedimiento:

1. Se toma aleatoriamente un vector de entrada (Ejemplo: \mathbf{x}_2).
2. Se calcula la distancia del vector de entrada a cada vector de pesos (se calcula h).

$$d_1(\mathbf{w}_1, \mathbf{x}_2) = \sqrt{(-1 - 4)^2 + (6 - (-6))^2} = 13.0 = h_1$$

$$d_2(\mathbf{w}_2, \mathbf{x}_2) = \sqrt{(-1 - (-6))^2 + (6 - 0)^2} = 7.81 = h_2$$

$$d_3(\mathbf{w}_3, \mathbf{x}_2) = \sqrt{(-1 - 4)^2 + (6 - 5)^2} = 5.09 = h_3$$

3. Se aplica la función de activación competitiva:

$$C \begin{bmatrix} h_1 \\ h_2 \\ h_3 \end{bmatrix} = C \begin{bmatrix} 13.0 \\ 7.81 \\ 5.09 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

El mayor h se presenta en la neurona 3, por lo tanto ella es la ganadora.

4. Se aplica la regla de aprendizaje:

$$\mathbf{w}_j(t+1) = \mathbf{w}_j(t) + \alpha (\mathbf{x}_i - \mathbf{w}_j(t)), \text{ si } i = g$$

$$\mathbf{w}_j(t+1) = \mathbf{w}_j, \text{ si } i \neq g$$

Para la actualización de los pesos de la red:

$$\mathbf{w}_1(1) = \mathbf{w}_1(0)$$

$$\mathbf{w}_2(1) = \mathbf{w}_2(0)$$

$$\mathbf{w}_3(1) = \mathbf{w}_3(0) + \alpha (\mathbf{x}_2 - \mathbf{w}_3(0)) = [4 \ 5] + 0.5([-1 \ 6] - [4 \ 5]) = [1.5 \ 5.5]$$

Así, se pasa de una configuración de pesos de la iteración 0 a la iteración 1:

$$W(0) = \begin{bmatrix} 4 & -6 \\ -6 & 0 \\ 4 & 5 \end{bmatrix}$$

$$W(1) = \begin{bmatrix} 4 & -6 \\ -6 & 0 \\ 1.5 & 5.5 \end{bmatrix}$$

Este proceso (pasos 1 a 4) se repite muchas veces (la red requiere un cierto grado de redundancia). Después de varias épocas la red converge a:

$$W(\text{final}) = \begin{bmatrix} 5.99 & 1.41 \\ 0.17 & 5.99 \\ -3.66 & -4.99 \end{bmatrix}$$

Como se puede observar en la Fig. 3, cada vector de pesos se ha especializado en una región del espacio de entrada y puede ser visto como un prototipo. De esta forma, cada neurona actúa como un detector de rasgos, y la neurona ganadora indica el tipo de rasgo o patrón detectado en el vector de entrada.

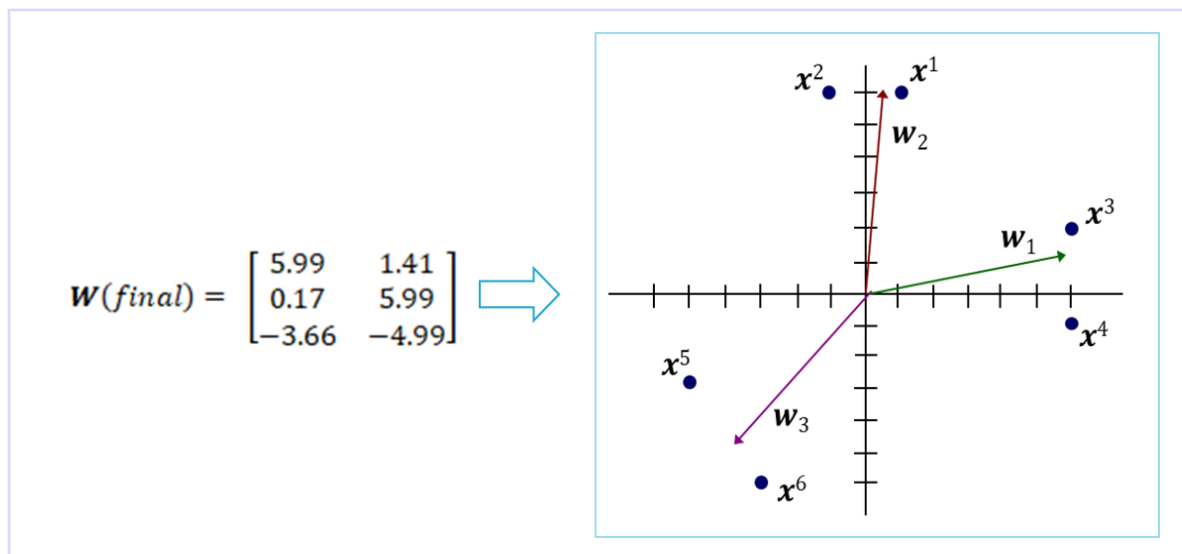


Fig. 3. Convergencia de la red competitiva para el ejemplo.

Bibliografía

Hagan, M., Demuth, H., Beale, M., De Jesús, O. (2014). *Neural Network Design*. IBN13 9780971732117

© - **Derechos Reservados:** la presente obra, y en general todos sus contenidos, se encuentran protegidos por las normas internacionales y nacionales vigentes sobre propiedad intelectual, por lo tanto su utilización parcial o total, reproducción, comunicación pública, transformación, distribución, alquiler, préstamo público e importación, total o parcial, en todo o en parte, en formato impreso o digital y en cualquier formato conocido o por conocer, se encuentran prohibidos, y solo serán lícitos en la medida en que se cuente con la autorización previa y expresa por escrito de la Universidad de los Andes.

De igual manera, la utilización de la imagen de las personas, docentes o estudiantes, sin su previa autorización está expresamente prohibida. En caso de incumplirse con lo mencionado, se procederá de conformidad con los reglamentos y políticas de la universidad, sin perjuicio de las demás acciones legales aplicables.
