



Universidad
Tecmilenio®



Redes neuronales

Modelo Fukushima
Modelo general de una
neurona artificial

Semana 2





Una de las razones por las cuales los humanos y las computadoras se necesitan mutuamente es la diferencia fundamental entre ellos. Las computadoras pueden hacer muy bien las tareas que los humanos realizan torpemente, por ejemplo, las computadoras son lógicas y precisas, mientras que los humanos no siempre son así.

No obstante, los humanos pueden procesar información de diferentes y múltiples fuentes en una escala que alcanza la generalidad, la incertidumbre, la ambigüedad, el posicionamiento y la integración en las conclusiones. ¿Cómo simularlo?

La suma de un conjunto de redes neuronales, las interfases de entrada/salida y los módulos lógicos adicionales constituyen un sistema neuronal artificial.



Características y funcionamiento del modelo Fukushima

Tienen un método de entrenamiento supervisado, en donde la red presenta el par de modos: el modo de entrada se empareja con el modo de salida deseado. Para cada representación, los pesos se ajustan para reducir el error entre la salida deseada y la respuesta de la red.

Asimismo, el algoritmo de aprendizaje incluye una etapa de propagación hacia adelante y una etapa de propagación hacia atrás, realizando dos etapas para cada patrón (Caterini y Chang, 2018).

De acuerdo con Namigtle (2017), en el modelo Fukushima los pesos sinápticos pueden ser generalmente positivos, nulos o negativos. Se denominan excitantes si son positivos () o inhibitorios si son negativos (). Entonces, todos los pesos sinápticos (fuerzas) y todas las señales de entrada y salida son no negativos.

Ventajas y desventajas.

Ventajas.

Su principal ventaja son sus capacidades de mapeo de modo general, ya que la red puede aprender una variedad de relaciones de mapeo de patrones. Por ende, no requiere conocimientos matemáticos sobre la función de los patrones de entrada y los patrones de salida. Asimismo, la propagación hacia atrás solo necesita ejemplos de mapeo para aprender.

Desventajas.

La mayor desventaja es el tiempo de convergencia, ya que, aunque puede haber miles de ejemplos en el conjunto de formación de una aplicación real, requiere varios días de tiempo de cálculo. Además, es propensa a fallar en el entrenamiento, es decir, es posible que la red nunca converja.



La salida del modelo Fukushima se representa por la siguiente expresión:

$$y_j = \varphi \left(\frac{1 + \sum_{i=1}^n a_{ji} x_i}{1 + \sum_{i=1}^n b_{ji} v_i} - 1 \right)$$

En donde:

φ	Es la función de activación.
a_{ji}	Son los pesos sinápticos $\in \mathbb{R}_{\geq 0}$.
b_{ji}	Son los pesos sinápticos $\in \mathbb{R}_{\geq 0}$.
x_i	Son las entradas $\in \mathbb{R}_{\geq 0}$.
v_i	Son las entradas $\in \mathbb{R}_{\geq 0}$.
n	Es el número de entradas.
y_j	Son las salidas.
j	Indica la j -ésima neurona.

φ está definida como:

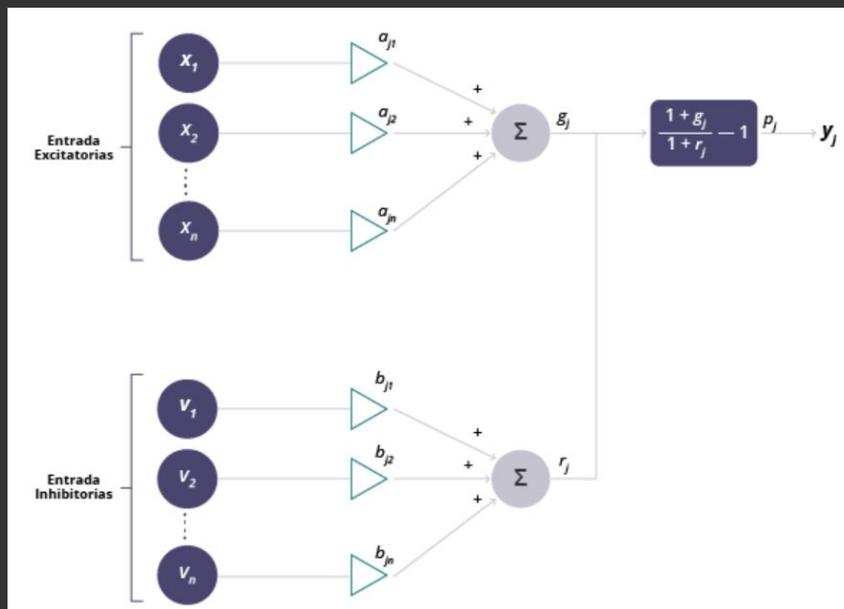
$$y_j = \varphi(p_j) = \begin{cases} p_j & \text{si } p_j \geq 0 \\ 0 & \text{otro caso} \end{cases}$$

En esta expresión, las " a_{ji} " significan los pesos sinápticos excitatorios (realizados normalmente como conductancias no negativas) y las " b_{ji} " son las sinapsis inhibitorias (también realizadas como conductancias no negativas).



El modelo Fukushima y la ley de Hebb:

Los pesos sinápticos son generalmente variables (están sintonizados) y se refuerzan durante el proceso de autoorganización de una red neuronal. Por ende, el proceso reforzado se basa principalmente en la ley de Hebb o regla de Hebb, que puede enunciarse sencillamente como sigue:



Modelo Fukushima.

Fuente: Caterini, A., y Chang, D. (2018). *Deep Neural Networks in a Mathematical Framework*. Suiza: Springer.

El neuropsicólogo Donald Hebb (Chéron, 2018) propuso que cuando dos o más neuronas se activan continuamente en el tiempo y el espacio, las conexiones sinápticas se fortalecen. Por tanto, si dos neuronas están activas, su conexión se fortalecerá. Asimismo, Hebb mencionó que, si el axón de la neurona A está lo suficientemente cerca de la célula B, promoviendo repetidamente su activación, ciertos cambios metabólicos aumentarán la eficiencia de esta sinapsis.

Esto dará como resultado la expansión de botones existentes en los axones, que estarán en contacto directo con la célula postsináptica. Las asociaciones entre diferentes neuronas pueden producir impresiones o ensamblaje de células "redes neuronales".

Por lo tanto, cuanto mayor sea la probabilidad entre la activación neuronal y un estímulo, es más probable que la red neuronal active el impulso cuando el estímulo vuelva a ocurrir. Hebb sugirió que para hacer esto la primera neurona debe activarse inmediatamente antes que la segunda neurona.

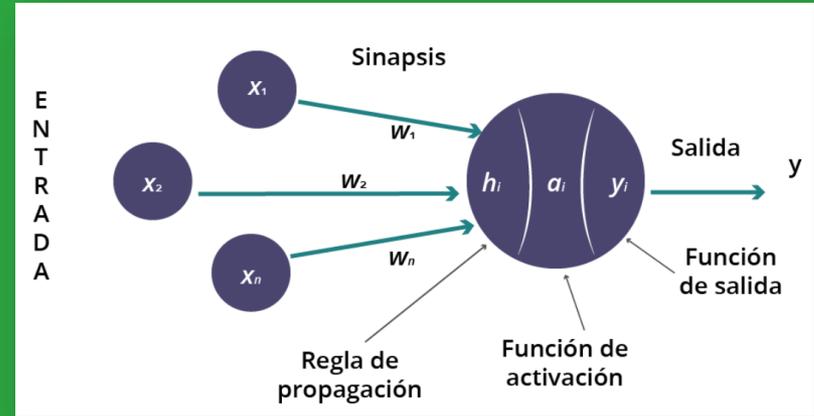


Modelo general de una neurona artificial.

Funciones de activación

- La neurona artificial es un mecanismo de procesamiento simple que, en función de un vector de entradas, elabora una salida única. Comúnmente, existen tres tipos de neuronas artificiales y cada una tiene su contraparte en el sistema nervioso (Acevedo, Serna y Serna, 2017).

Elementos que componen una red artificial



1. Neuronas de entrada: toman información desde el exterior.

2. Neuronas ocultas: reciben información de otras neuronas artificiales y en sus sinapsis se lleva a cabo la representación de los datos acumulados.

3. Neuronas de salida: reciben información procesada y la regresan al exterior.





Neurona artificial según el modelo general

Conjunto de entradas x_j proceden del exterior o de otras neuronas artificiales.

Pesos sinápticos W_{ij} : comunicación entre la neurona artificial j y la neurona artificial i . Pueden ser excitadores o inhibidores.

Regla de propagación h_i : compone la información derivada de las diferentes neuronas y provee el valor del potencial postsináptico de la neurona i .

Función de activación a_i : establece el estado de activación real de la neurona i .

Función de salida y_i representa la salida actual de la neurona i .

La neurona cuenta con **entradas y salidas** que se catalogan en dos conjuntos:

Binarias
Continuas

La elección de usar uno u otro prototipo de neurona resulta de la aplicación y del modelo a construir (García, 2021).





La regla de propagación establece el potencial que resulta de la interacción de la neurona i con las N neuronas vecinas. El potencial resultante h_i se puede enunciar de la siguiente forma:

$$h_i = \sigma_i(w_{ij}, x_j)$$

La regla de propagación más sencilla y aplicada radica en elaborar una suma de las entradas ponderadas con sus respectivos pesos sinápticos:

:

$$h_i = \sum_{j=1}^n w_{ij} * x_j$$

La función de activación establece el estado de activación real de la neurona con base en el potencial resultante h_i y al estado de activación previo de la neurona $a_i(t-1)$ (García, 2021).



El peso sináptico w_{ij} establece la fuerza de una conexión sináptica entre la neurona presináptica i y la postsináptica j . Este peso puede alcanzar valores positivos, negativos o cero. De ser una entrada positiva, un peso positivo opera como excitador, en tanto que un peso negativo opera como inhibidor. Si el peso es cero, no existe comunicación entre ambas neuronas. La red está preparada para adaptarse al contexto y ejecutar una tarea definida mediante el ajuste de los pesos sinápticos (García, 2021).



Las **funciones de activación** más empleadas en los diferentes modelos de redes neuronales artificiales son las siguientes:

Función	Fórmula	Rango
Identidad	$y = x$	$[-\infty, \infty]$
Escalón	$y = \begin{cases} 1, & \text{si } x \geq 0 \\ 0, & \text{si } x < 0 \end{cases}$	$[0, 1]$
Escalón	$y = \begin{cases} 1, & \text{si } x \geq 0 \\ -1, & \text{si } x < 0 \end{cases}$	$[-1, 1]$
Lineal a tramos	$y = \begin{cases} 1, & \text{si } x \geq 0 \\ x, & \text{si } -1 \leq x \leq 1 \\ -1, & \text{si } x < -1 \end{cases}$	$[-1, 1]$
Sigmoidea	$y = \frac{1}{1 + e^{-x}}$	$[0, 1]$
Sigmoidea	$y = \tanh(x)$	$[-1, 1]$
Sinusoidal	$y = A \sin(wx + \varphi)$	$[-1, 1]$





Función de salida

Por otro lado, con respecto al rango de la neurona artificial, una forma sistemática es analizando con la variable v , y posteriormente a y . Debido a que las entradas son binarias, los pesos sinápticos son reales y el umbral θ es cero, se puede construir el conjunto de valores posibles de v y de y . Hay muchos artificios para construir este conjunto, pero ahora se utilizará una tabla para mostrar los pormenores.

x_1	x_2	θ	$v = w_1x_1 + w_2x_2 + \theta$	$y = \varphi(v)$
0	0	0	0	0
0	1	0	w_2	$\tanh(w_2)$
1	0	0	w_1	$\tanh(w_1)$
1	1	0	$w_1 + w_2$	$\tanh(w_1 + w_2)$





Conjuntos difusos y función de pertenencia

En la teoría de conjuntos difusos, la pertenencia de un elemento a un conjunto difuso está generalmente evaluada con una función que indica en qué grado dicho elemento pertenece al conjunto.

Iniciemos pues, por definir textualmente el conjunto difuso de las personas altas.

Asimismo, el algoritmo de aprendizaje incluye una etapa de propagación hacia adelante y una etapa de propagación hacia atrás, realizando dos etapas para cada patrón (Caterini y Chang, 2018).

$$\Phi = \{ \text{personas altas} \}$$

Antes de expresar matemáticamente el conjunto difuso anterior, definamos una función de pertenencia, la cual expresa el grado de pertenencia de cada elemento del universo de discurso.

$$\mu_{\Phi}(\text{persona}) = \begin{cases} 0.30, & \text{persona} = \text{Paty} \\ 0.60, & \text{persona} = \text{Lupe} \\ 0.90, & \text{persona} = \text{José} \\ 0.95, & \text{persona} = \text{Toño} \\ 1.00, & \text{persona} = \text{Monse} \end{cases}$$

De manera que puedes observar que:

Considera un conjunto de personas que trabajan en una determinada oficina, en la cual están Paty, Monse, Lupe, José y Toño. Sea U el conjunto personas que laboran en una oficina, donde: $U = \{ \text{Paty}, \text{Monse}, \text{José}, \text{Lupe}, \text{Toño} \}$

Persona	Estatura
Paty	1.6
Lupe	1.7
José	1.78
Toño	1.83
Monse	1.85

$$\mu_{\Phi}(\text{Paty}) = 0.3,$$

Esto expresa que la estatura de Paty, que mide 1.6 m., pertenece al conjunto difuso Φ de las personas altas en un grado de 0.3.



Con base en lo descrito en el tema, reflexiona sobre las siguientes preguntas.



01

En el modelo Fukushima, ¿Qué diferencia tienen los dos tipos de entradas? ¿Existirá un máximo de entradas? ¿Qué tendrá de diferente este modelo con otros existentes?

02

En el modelo general, ¿Cómo identificar un conjunto difuso? ¿Cómo encontrar la función de pertenencia?





En este tema se desarrolló el funcionamiento de la red neuronal Fukushima, el cual depende de dos entradas (excitatorias e inhibitorias), en donde sus pesos sinápticos pueden ser positivos, nulos o negativos. Asimismo, se presentó el diagrama del modelo de Fukushima junto a su ecuación general, expresando que la base de su diseño fue la ley de Hebb.



Además, se desarrolló la función de activación donde se utilizaron los conjuntos difusos, y con ello identificar la función de pertenencia.





- Acevedo, E., Serna, A., y Serna, E. (2017). Principios y características de las redes neuronales artificiales. En M.E. Serna (coord.), *Desarrollo e innovación en ingeniería*. Colombia: Instituto Antioqueño de Investigación.
- Caterini, A., y Chang, D. (2018). *Deep Neural Networks in a Mathematical Framework*. Suiza: Springer.
- Chéron, G. (2018). Neurofisiología del movimiento. Oscilaciones neuronales y aprendizaje motor. *EMC-Kinesiterapia-Medicina Física*, 39. Recuperado de <https://www.sciencedirect.com/science/article/abs/pii/S1293296518898127>
- García, J. (2021). Redes neuronales artificiales: aplicación a la regionalización de la precipitación y temperaturas diarias. Recuperado de https://repositorio.aemet.es/bitstream/20.500.11765/12707/1/NT_34_Neural_networks_artificial.pdf
- Namigtle, J. (2017). *Identificación de sistemas no lineales mediante redes neuronales, equilibrio entre costo computacional y precisión* (Tesis de maestría). Centro Nacional de Investigación y Desarrollo Tecnológico, México.

