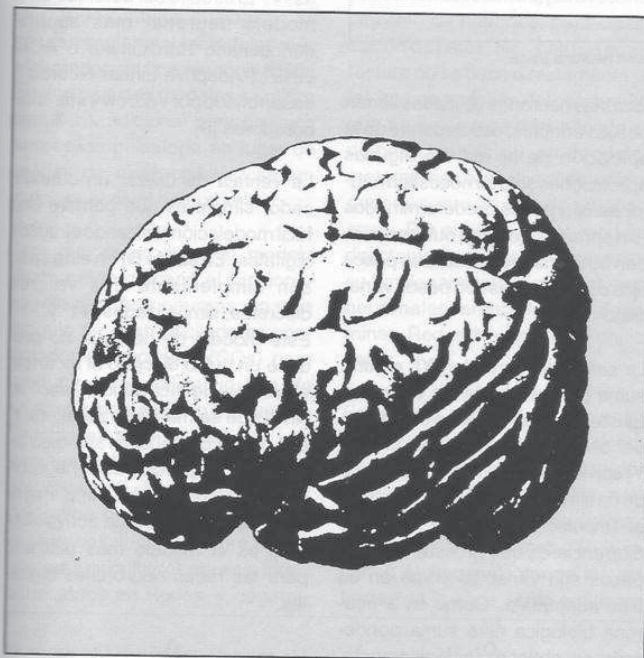


Introducción a las Redes Neuronales Artificiales

(Segunda Parte -Final-)



□ Gabriel J. Gómez R., M.Sc.
Ingeniero Electricista
Profesor Auxiliar
Departamento de Electricidad
Universidad del Valle.

5. IMPLEMENTACIONES DE REDES NEURONALES ARTIFICIALES

Tal como se mencionó, existen dos tipos de redes neuronales: las redes neuronales modeladas y las redes neuronales funcionales. Las implementaciones para cada tipo son muy diferentes (siendo bastante

numerosas en cada caso) por lo que se tratarán independientemente.

5.1. Redes neuronales modeladas.

Las implementaciones de este tipo de redes, que tratan de emular físicamente las redes neuronales biológicas, son bastante complicadas para explicarlas en este artículo introductorio, ya que para una sola función neuronal (visión en la retina, audición en la cóclea, etc.) deben modelarse varios tipos de neuronas y sus intrincadas interconexiones. Además el diseño de este tipo de redes exige conocimientos avanzados de la teoría de circuitos integrados analógicos.

En este artículo nos limitaremos a mencionar que aunque no son muchos, se han implementado en la práctica algunos sistemas neuronales locales, con algún éxito. Tal como se mencionó anteriormente, uno de los grupos de investigación líderes en este tipo de diseños es el de Carver Mead y sus colaboradores en CalTech, en Los Angeles, California (USA), donde se han diseñado varios sistemas como el *detector de Movimiento*, la *Retina Electrónica*, la *Cóclea* y otros más.

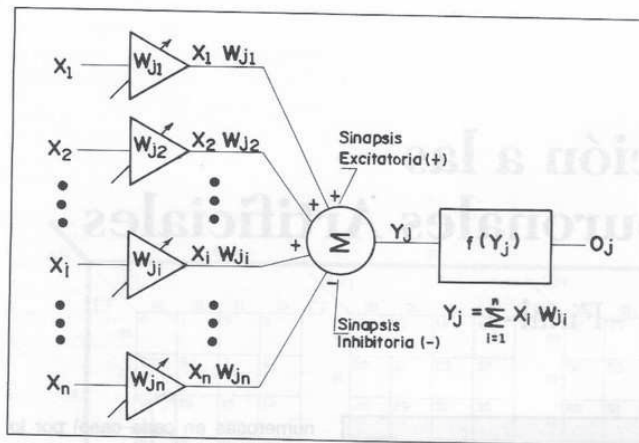


Fig. 3 Modelo general de una neurona típica

Aunque hay un gran interés por estos tipos de redes, especialmente por su importancia fisiológica y por el gran aporte que hacen a la investigación neuronal, su utilidad práctica es muy poca (por ahora), y las aplicaciones comerciales apenas están surgiendo. Por esta razón se le dedicará más espacio a la explicación del otro tipo de redes neuronales: Las redes neuronales funcionales.

5.2. Redes neuronales funcionales.

Modelo neuronal general

Las implementaciones de redes neuronales funcionales se basan en su gran mayoría en modelos de la neurona típica vista en la Fig. 1. El modelo general más utilizado se muestra en la Fig. 3.

En este modelo, los W_{ji} son los "pesos" que permiten la ponderación de las entradas X_i . El valor de estos pesos se hace variable, para permitir el "aprendizaje" de las neuronas al ir variando los pesos en forma dinámica. El subíndice j indica la posibilidad de variación de los pesos con cada iteración en el tiempo, durante el aprendizaje. No

todas las neuronas utilizadas tienen pesos variables, eso depende de la aplicación de la red; en algunas aplicaciones solo se necesitan neuronas con pesos predeterminados (sin aprendizaje), y en otras se pueden combinar neuronas con pesos fijos con neuronas con pesos variables.

La salida del sumador, Y_j , es una suma ponderada de las entradas. Existe una gran similitud de la parte del circuito utilizada para obtener Y_j con la implementación canónica de un filtro digital de respuesta finita al Impulso (Filtro FIR); la única diferencia es que en este caso los pesos son variables como en un filtro adaptativo. Como en la neurona biológica ésta suma ponderada se obtiene analógicamente, se logra un modelo más exacto usando circuitos analógicos, pero también pueden utilizarse valores cuantificados (discretos) para los X_i y los W_{ji} permitiendo así la utilización de circuitos digitales.

Función de Activación

La función de activación $f(Y_j)$ permite obtener una salida similar a la

producida por una neurona biológica, dependiendo del valor de Y_j . La idea es obtener una salida discreta, o un pulso, o una serie de pulsos modulados en frecuencia, de acuerdo con el modelo neuronal que se desee obtener.

La función más simple (Fig. 4) es un cuantificador que realiza una decisión entre los valores $+V$ y $-V$ así: si Y_j es mayor que un umbral predeterminado, la salida es $+V$; si Y_j es menor que el umbral, la salida es $-V$. El circuito así obtenido es el modelo neuronal más simple, denominado "Hard Limiter" o "ADALINE" ("Adaptive Linear Neuron"), desarrollado por Widrow y sus colaboradores [9].

La ventaja de utilizar un cuantificador simple es que permite una fácil modelación utilizando circuitos digitales. La salida O_j en este caso son simplemente dos valores discretos, simbolizados $+1$ y -1 . Este modelo de neurona no produce un pulso discreto al alcanzar el voltaje de activación, sino un Nivel que se mantiene mientras Y_j sea mayor que el umbral (no es un modelo muy preciso de la neurona real). Debido a la fácil implementación de la función de activación, éste es el modelo más utilizado para las redes neuronales digitales.

Un modelo mejorado de neurona utiliza la función de activación *Sigmoidal*, la cual no produce una salida discreta, sino analógica. Este modelo trata de modelar mejor la región en la cual la neurona tiene una respuesta no lineal a la suma ponderada de entradas, para valores menores a la tensión de umbral. La función sigmoideal se basa en la forma exponencial: $(1 + e^{-x})^{-1}$, y produce una característica

entrada-salida como la mostrada en la Fig. 4b. Esta función de activación es mucho más difícil de ser modelada digitalmente (aunque a veces se hace, utilizando modelos lineales por tramos, o aproximaciones, como una función rampa), pero es fácilmente modelada utilizando circuitos analógicos (la característica entrada-salida de un amplificador de transconductancia es muy similar a la función sigmoide).

Modelos mejorados

Utilizando circuitos analógicos se pueden implementar modelos más elaborados de una neurona típica. En uno de estos modelos se utiliza circuitería adicional para producir una salida pulsatoria en lugar de salidas de nivel, una vez se sobrepasa el nivel de umbral.

Otra mejora produce un modelo mucho más cercano al funcionamiento real de la neurona; en este modelo se usan osciladores controlados por voltaje (VCOs) para producir como salida un tren de impulsos modulado en frecuencia. Gracias al VCO se obtiene una frecuencia de salida que depende del voltaje Y_j . Estos últimos modelos son denominados *modelos oscilatorios*, los cuales presentan mayores dificultades para ser interconectados en Redes, y general-

mente son implementados utilizando circuitos analógicos.

Interconexiones entre neuronas: las redes neuronales

Una vez se ha escogido el modelo de neurona a utilizar, se debe también elegir un modelo de la red neuronal que produzca buenos resultados para aplicaciones específicas. Se han desarrollado modelos de redes para implementar memorias asociativas, para reconocimiento de patrones, para clasificación de datos, etc., pero su arquitectura no se basa directamente en las interconexiones del cerebro (ya que no se conoce aún a fondo su disposición), sino en modelos matemáticos o intuitivos, obtenidos generalmente a base de prueba y error. Por el hecho de no modelar directamente una red biológica sino su funcionamiento, estos sistemas neuronales artificiales se denominan *Redes Neuronales Funcionales*.

En una red dada se colocan una o varias "capas" de neuronas, donde las posibilidades de interconexión son innumerables. Una de las redes más utilizada se conforma de tres o más capas de neuronas ("Multi-layer" o Multi-capas), donde las entradas en la capa i sólo provienen

de las salidas de la capa $i-1$, o sea solamente hay alimentación "hacia adelante", por lo que la red se denomina "feed-forward"; en la Fig. 5 se ilustra una red de este tipo con tres capas de neuronas.

En la red mostrada se utiliza un interconexión completo (lo cual no es necesario, puede hacerse también en forma parcial), en el cual la salida de cada neurona va a todas las neuronas de la capa siguiente. La primera capa se denomina "capa de entrada"; la segunda, de gran importancia en el aprendizaje se denomina la "capa Escondida" ("Hidden Layer") pues no puede ser accesada desde el exterior, y la tercera capa es la de salida.

Se han desarrollado muchos otros tipos de redes, en los que se usan realimentaciones laterales (entre neuronas de la misma capa), o realimentaciones hacia capas anteriores ("feed-backward"), o combinaciones de ellas. Es imposible enumerar todos los tipos de redes que se han desarrollado, pues cada investigador clama que su diseño es el más óptimo para determinada aplicación. El número de capas utilizado depende también de la aplicación y del diseñador, y es determinado en general utilizando el método de prueba y error. Algunos de los tipos de redes más conocidos y utilizados además del visto, son el de Hamming, Hopfield, Anderson, etc. (ver [5], [6], [7], [8]).

Algoritmos para el aprendizaje (Adaptación de los pesos)

Tal como se mencionó, cada entrada a una neurona es ponderada por un "peso" W_{ji} . El valor de W_{ji} en general no es fijo (aunque existen redes diseñadas con todos

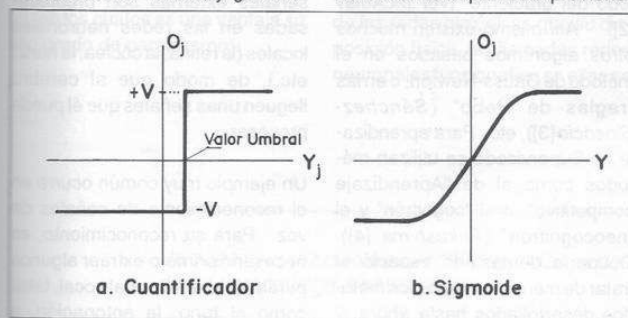


Fig. 4 Funciones de activación

los pesos fijos), sino que puede adaptarse, de acuerdo con un *algoritmo de aprendizaje*. El aprendizaje puede ser *supervisado* o *no-supervisado*.

Para el aprendizaje supervisado se sigue más o menos el proceso descrito a continuación: Para unas entradas de control se conoce la salida deseada. Se presenta una de esas entradas en $(x_1, \dots, x_n)$, y por medio de un algoritmo interactivo se adaptan los pesos hasta lograr que el Error entre la salida deseada (conocida) y la salida presente ($O_1, \dots, O_m$) sea mínimo en algún sentido (en general se minimiza el error medio cuadrado - MSE), para esa entrada. Luego se presentan más entradas con salidas conocidas, y se permite la adaptación de los pesos nuevamente; poco a poco la red "aprende" qué salidas debe presentar a una entrada dada.

En el aprendizaje no supervisado se presentan solamente las entradas, y se espera a que los pesos se adapten a cada entrada, sin comparar con ninguna salida. En este caso debe existir un algoritmo que decida cuáles pesos deben ser reforzados y cuáles anulados, para una combinación de entradas dada.

Lo importante del aprendizaje es que después de "aprender" su comportamiento para una serie de entradas de control, la red debe ser capaz de *Generalizar*. Porejemplo, si se le enseña a reconocer el carácter "A" en cuatro posiciones diferentes (0, 90, 180, 270 grados), la red debe ser capaz de reconocerlo si se le presenta en una nueva posición, o en presencia de ruido, etc. Muchas redes implementadas realizan fácilmente labores como la descrita. Si una red con apren-

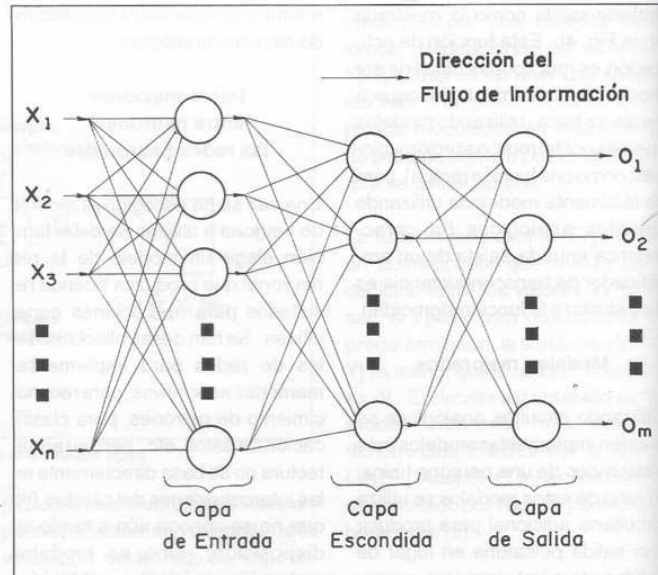


Fig. 5 Red neuronal multicapa en conexión "Feed-forward"

dizaje no es capaz de generalizar, su utilidad es prácticamente nula.

Se han desarrollado numerosos algoritmos para el aprendizaje; uno de los algoritmos de aprendizaje supervisado más conocidos es el de "backpropagation" (propagación hacia atrás), el cual modifica primero los pesos de la capa de salida y luego comienza a modificar los de las capas anteriores, siempre tratando de minimizar el error mínimo cuadrado, utilizando el "método del gradiente" (ver McAulay [2]). Así mismo existen muchos otros algoritmos basados en el método de Gauss-Newton, o en las "reglas de Hebb" (Sánchez-Sinencio [3]), etc. Para aprendizaje no-Supervisado se utilizan métodos como el de "Aprendizaje competitivo", o el "cognitrón" y el "neocognitrón" (Fukushima [4]). Ocuparía demasiado espacio el tratar de mencionar todos los métodos desarrollados hasta ahora, y más aún el tratar de explicarlos;

para mayor información al respecto, consultar las referencias dadas al final del artículo.

Pre-procesamiento

En general, antes de utilizar las redes neuronales para procesar una señal física, es necesario realizar un preproceso de esas señales para convertirlas en señales que la red "entienda". En ese sentido, el funcionamiento de la red es análogo al del cerebro: las señales externas son preprocesadas en las redes neuronales locales (la retina, la cóclea, la nariz, etc.), de modo que al cerebro lleguen unas señales que él pueda procesar.

Un ejemplo muy común ocurre en el reconocimiento de señales de voz. Para su reconocimiento, es necesario primero extraer algunos parámetros de la señal vocal, tales como el tono, la entonación, y principalmente las frecuencias com-

ponentes (*formantes*). Utilizando estos parámetros simplificados puede intentarse reconocer el vocablo pronunciado, cosa que sería muy difícil, sino imposible, utilizando la señal vocal directamente. El trabajo de preprocesamiento puede ser una aplicación futura muy importante para las redes neuronales modeladas, ya que ellas en general modelan redes neuronales locales, que son las que realizan ese preproceso en el sistema neurológico humano.

NOTA FINAL: Tal como se ha mencionado varias veces, las redes neuronales artificiales pueden implementarse ya sea utilizando circuitos analógicos (se dificulta más el aprendizaje) o circuitos digitales (se dificulta la generación de la función de activación), o mejor aún, utilizando circuitos analógico-digitales (donde se combinan las ventajas de cada tecnología). También puede utilizarse un computador digital para simular su funcionamiento, a costo de un mayor tiempo de procesamiento. Pero la implementación no tiene que ser necesariamente utilizando circuitos electrónicos; se ha intentado modelar redes neuronales utilizando otros sistemas de procesamiento, entre los cuales ha tenido algún éxito la modelación usando Sistemas de procesamiento ópticos [2] (en los cuales es una ventaja su alto grado de paralelismo).

6. CONCLUSIONES

En este artículo se presentó una introducción a un tema de tanta actualidad como son las Redes Neuronales Artificiales, intentando desarrollar en la forma más completa posible un tema demasiado extenso para ser explicado en unas cuantas páginas.

El campo de investigación en redes neuronales requiere en general del trabajo de un equipo interdisciplinar (neurofisiólogos, ingenieros de sistemas, ingenieros electrónicos, etc.). Aún existen demasiados enigmas por resolver en cuanto al funcionamiento de las redes neuronales biológicas, en particular de las encontradas en el cerebro humano, trabajo que mantendrá ocupados a los neurólogos y neurofisiólogos por bastantes años más.

Las redes Neuronales Artificiales desarrolladas hasta ahora, son sólo burdas imitaciones de las biológicas. Sólo se han podido emular con cierta precisión algunas redes locales cuyos principios fisiológicos son conocidos más o menos bien (la retina, la cóclea, etc.), utilizando las denominadas redes neuronales "modeladas". También se ha desarrollado otro tipo de redes que trata de emular más el funcionamiento de las redes biológicas que su disposición física, las llamadas redes neuronales funcionales; en ellas se

modela lo mejor posible una neurona biológica, pero luego la forma de interconectar estas neuronas artificiales se basa en algoritmos y arquitecturas extraídos de modelos matemáticos o intuitivos, que en general no tienen nada que ver con las redes biológicas.

Aunque se han obtenido buenos resultados en algunos temas específicos como memorias asociativas o reconocimiento de patrones, falta aún mucha investigación y desarrollo, antes de obtener modelos realmente comerciales. Se han obtenido unos resultados a veces sorprendentes en procesamiento de señales, los cuales sería muy difícil o hasta imposible obtener utilizando computadores digitales convencionales; al comparar estos resultados con los que obtiene el cerebro, nos damos cuenta de lo lejos que aún estamos de lograr resultados que tan siquiera se acerquen a los obtenidos por esa extraordinaria máquina biológica desarrollada por la evolución.

Lo mejor de todo es que debido a lo mucho que aún se puede lograr, éste es un campo en donde la investigación es bienvenida, y cualquiera puede aportar su granito de arena. Mejor aún, para comenzar sólo es necesario tener acceso a un computador, trabajar con muchos deseos, y poner a funcionar nuestra red neuronal más importante: *el cerebro*. ▣

Un buen oyente no solo es popular en todas partes sino que después de un tiempo llega a saber algo.

JW. Marriott, Sr.
