

## INTRODUCCIÓN

Las líneas de investigación que se han establecido en la facultad de Ingeniería Eléctrica han permitido que tanto profesores como estudiantes encuentren las herramientas necesarias para sus procesos de investigación, adicional a esto se han presentado tendencias alternativas que fortalecen el desarrollo de nuevas tecnologías no sólo en la facultad de Ingeniería Eléctrica, sino en acuerdos con otras facultades, e inclusive con la industria regional y nacional; en este aspecto hay aún mucho que aportar y es precisamente ese el objetivo de este proyecto con el cual se ha pretendido introducir los conceptos y herramientas básicas sobre Redes Neuronales Artificiales, para que cualquier persona interesada en conocer esta teoría y aplicarla en la solución de un problema específico, esté en capacidad de decidir si este método es el mejor camino y en tal caso de escoger el tipo de Red Neuronal Artificial que le permita dar solución a su interés particular.

Dando prioridad a las aplicaciones que pueda tener esta teoría en el campo de la Ingeniería Eléctrica, este proyecto se ha dividido en tres secciones principales o capítulos que progresivamente llevarán al lector a tener una idea clara de lo que es la teoría de las redes neuronales y cuales son los requisitos para aplicarla.



En la primera sección o capítulo uno, se encontrará un bosquejo general sobre el desarrollo de la teoría de las redes neuronales artificiales desde su nacimiento hasta los últimos aportes que se han hecho; se encontrará también un resumen especializado acerca del funcionamiento del cerebro en general y de cómo ocurre el proceso neuronal, para terminar haciendo la analogía que llevó a la transición del modelo biológico al modelo artificial. Este primer capítulo es fundamental para el entendimiento de las Redes Neuronales Artificiales, pues el lector descubrirá allí la importancia de esta teoría, además entrará en el ambiente matemático y gráfico que le permitirá comprender con facilidad las secciones posteriores.

El siguiente capítulo es dedicado a la explicación de los principales tipos de Redes Neuronales Artificiales, allí se profundiza en los aspectos fundamentales de cada tipo de red, sus antecedentes y evolución histórica, la estructura general de la red y el desarrollo matemático de la regla de aprendizaje. Aunque en este capítulo se trataron especialmente las redes con base en las cuales se desarrollaron las aplicaciones del capítulo tres, la idea es que el lector encuentre una motivación inicial que lo lleve a profundizar en otros tipos de redes de igual importancia que las estudiadas y muchas de ellas de mayor novedad y alcance.

El capítulo tres es la recopilación de todos los conocimientos adquiridos desde el capítulo uno y ratificados en el capítulo dos, aplicados a problemas típicos de Ingeniería Eléctrica; en la elección de las aplicaciones de este



capítulo se trató de escoger aquellos problemas representativos en el área de eléctrica, que permitieran ilustrar ampliamente el comportamiento de la red neuronal escogida en cada caso. En este capítulo se destacan las bondades de las redes neuronales para solucionar problemas para los cuales no se había encontrado una respuesta satisfactoria mediante métodos tradicionales, o cuya solución involucraba procesos muy complejos que son enormemente simplificados al resolverse utilizando el método de las redes neuronales artificiales. Para una buena comprensión de este capítulo, cada aplicación se ha subdividido en tres partes, la descripción teórica del problema, una justificación del uso de la red neuronal escogida y una fase final que explica el proceso de entrenamiento de la red, en la cual se puede encontrar el código fuente para que el lector actualice los datos y realice sus propias aplicaciones.

Tratando de que esta herramienta esté al alcance del mayor número de personas posible, este proyecto ha tomado la forma de tutorial sobre Redes Neuronales Artificiales y se ha implementado en un sitio Web hospedado en el servidor de la Universidad Tecnológica de Pereira, cuya dirección es: <http://ohm.utp.edu.co/neuronales>, al cual tendrán acceso no sólo las personas vinculadas a la universidad, sino todas las personas interesadas en encontrar una primera herramienta en su estudio de Redes Neuronales Artificiales.



Para terminar, es nuestro deber aclarar que el estudio de esta teoría involucra además de tiempo y dedicación una fundamentación matemática al nivel de cálculo diferencial para que pueda ser apreciado su real aporte; esperamos que este proyecto sea solo el primer paso en una cadena de proyectos de investigación que permitan a la facultad de Ingeniería Eléctrica y a la Universidad Tecnológica en general seguir siendo pioneros en el desarrollo de nuevas tecnologías y de sus herramientas de construcción.



# **1. CARACTERÍSTICAS PRINCIPALES DE LAS REDES NEURONALES**

## **1.1 INTRODUCCIÓN A LAS REDES NEURONALES**

Resulta irónico pensar que máquinas de cómputo capaces de realizar 100 millones de operaciones en coma flotante por segundo, no sean capaces de entender el significado de las formas visuales o de distinguir entre distintas clases de objetos. Los sistemas de computación secuencial, son exitosos en la resolución de problemas matemáticos o científicos, en la creación, manipulación y mantenimiento de bases de datos, en comunicaciones electrónicas, en el procesamiento de textos, gráficos y auto edición, incluso en funciones de control de electrodomésticos, haciéndolos más eficientes y fáciles de usar, pero definitivamente tienen una gran incapacidad para interpretar el mundo.

Esta dificultad de los sistemas de cómputo que trabajan bajo la filosofía de los sistemas secuenciales, desarrollados por Von Neuman, ha hecho que un gran número de investigadores centre su atención en el desarrollo de nuevos sistemas de tratamiento de la información, que permitan solucionar problemas cotidianos, tal como lo hace el cerebro humano; este órgano biológico cuenta con varias



características deseables para cualquier sistema de procesamiento digital, tales como:

1. Es robusto y tolerante a fallas, diariamente mueren neuronas sin afectar su desempeño.
2. Es flexible, se ajusta a nuevos ambientes por medio de un proceso de aprendizaje, no hay que programarlo.
3. Puede manejar información difusa, con ruido o inconsistente.
4. Es altamente paralelo
5. Es pequeño, compacto y consume poca energía.

El cerebro humano constituye una computadora muy notable, es capaz de interpretar información imprecisa suministrada por los sentidos a un ritmo increíblemente veloz. Logra discernir un susurro en una sala ruidosa, un rostro en un callejón mal iluminado y leer entre líneas un discurso; lo más impresionante de todo, es que el cerebro aprende sin instrucciones explícitas de ninguna clase, a crear las representaciones internas que hacen posibles estas habilidades.

Basados en la eficiencia de los procesos llevados a cabo por el cerebro, e inspirados en su funcionamiento, varios investigadores han desarrollado desde hace más de 30 años la teoría de las Redes Neuronales Artificiales (RNA), las cuales emulan el comportamiento de las redes neuronales biológicas, y que se han utilizado para aprender estrategias de solución basadas en ejemplos de



comportamiento típico de patrones; estos sistemas no requieren que la tarea a ejecutar se programe, ellos generalizan y aprenden de la experiencia.

La teoría de las RNA ha brindado una alternativa a la computación clásica, para aquellos problemas, en los cuales los métodos tradicionales no han entregado resultados muy convincentes, o poco convenientes. Las aplicaciones más exitosas de las RNA son:

1. Procesamiento de imágenes y de voz
2. Reconocimiento de patrones
3. Planeamiento
4. Interfaces adaptivas para sistemas Hombre/máquina
5. Predicción
6. Control y optimización
7. Filtrado de señales

Los sistemas de cómputo tradicional procesan la información en forma secuencial; un computador serial consiste por lo general de un solo procesador que puede manipular instrucciones y datos que se localizan en la memoria, el procesador lee y ejecuta una a una las instrucciones en la memoria; este sistema serial es secuencial, todo sucede en una sola secuencia determinística de operaciones. Las RNA no ejecutan instrucciones, responden en paralelo a las entradas que se les presenta; el resultado no se almacena en una posición de memoria, este es el estado de la red para el cual se logra equilibrio. El conocimiento de una red



neuronal no se almacena en instrucciones, el poder de la red está en su topología y en los valores de las conexiones (pesos) entre neuronas.

Las RNA son una teoría que aún está en proceso de desarrollo, su verdadera potencialidad no se ha alcanzado todavía; aunque los investigadores han desarrollado potentes algoritmos de aprendizaje de gran valor práctico, las representaciones y procedimientos de que se sirve el cerebro, son aún desconocidas. Tarde o temprano los estudios computacionales del aprendizaje con RNA acabarán por converger a los métodos descubiertos por evolución, cuando eso suceda, muchos datos empíricos concernientes al cerebro comenzarán súbitamente a adquirir sentido y se tornarán factibles muchas aplicaciones desconocidas de las redes neuronales.

## 1.2 FUNCIONAMIENTO DE UNA NEURONA BIOLÓGICA

El cerebro consta de un gran número (aproximadamente  $10^{11}$ ) de elementos altamente interconectados (aproximadamente  $10^4$  conexiones por elemento), llamados neuronas. Estas neuronas tienen tres componentes principales, las dendritas, el cuerpo de la célula o soma y el axón. Las dendritas, son el árbol receptor de la red, son como fibras nerviosas que cargan de señales eléctricas el cuerpo de la célula. El cuerpo de la célula, realiza la suma de esas señales de entrada. El axón es una fibra larga que lleva la señal desde el cuerpo de la célula hacia otras neuronas. El punto de contacto entre un axón de una célula y una





dendrita de otra célula es llamado sinapsis, la longitud de la sinápsis es determinada por la complejidad del proceso químico que estabiliza la función de la red neuronal. Un esquema simplificado de la interconexión de dos neuronas biológicas se observa en la figura 1.2.1

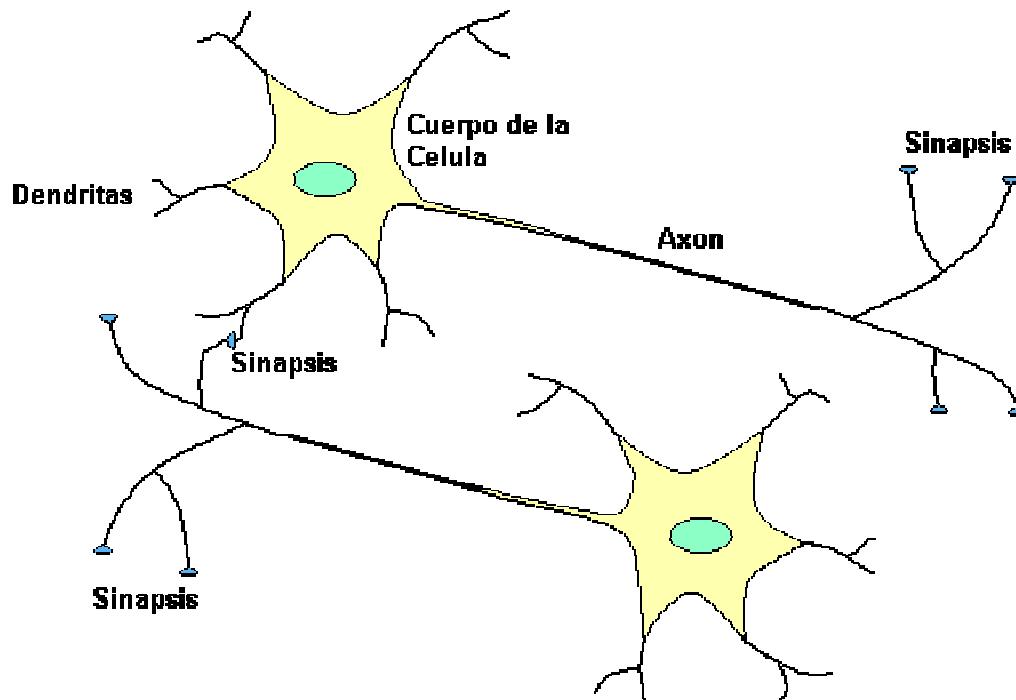


Figura 1.2.1 Neuronas Biológicas

Algunas de las estructuras neuronales son determinadas en el nacimiento, otra parte es desarrollada a través del aprendizaje, proceso en que nuevas conexiones neuronales son realizadas y otras se pierden por completo. El desarrollo neurológico se hace crítico durante los primeros años de vida, por ejemplo está demostrado que si a un cachorro de gato, se le impide usar uno de sus ojos durante un período corto de tiempo, el nunca desarrollara una visión normal en ese ojo.



Las estructuras neuronales continúan cambiando durante toda la vida, estos cambios consisten en el refuerzo o debilitamiento de las uniones sinápticas, por ejemplo se cree que nuevas memorias son formadas por la modificación de esta intensidad entre sinapsis, así el proceso de recordar el rostro de un nuevo amigo, consiste en alterar varias sinapsis.

Como consecuencia de los primeros estudios sobre la base neural de los sistemas mnémicos (relacionados con la memoria), se creía que el almacenamiento de la memoria asociativa, requería de un circuito neuronal muy complejo. Entre quienes comenzaron a oponerse a este enfoque se hallaba Donald O. Hebb, profesor de la Universidad de Milner; Hebb sugirió que el aprendizaje asociativo podría ser producido por un mecanismo celular sencillo y propuso que las asociaciones podrían formarse por una actividad neuronal coincidente: “Cuando un axón de la célula A excita la célula B y participa en su activación, se produce algún proceso de desarrollo o cambio metabólico en una o en ambas células, de suerte que la eficacia de A, como célula excitadora de B, se intensifica”. Según la regla Hebbiana de aprendizaje, el que coincida la actividad de las neuronas presinápticas (suministran el impulso de entrada) con la de las postsinápticas (reciben el impulso) es muy importante para que se refuerce la conexión entre ellas, este mecanismo es llamado pre-postasociativo, del cual puede observarse un ejemplo en la figura 1.2.2



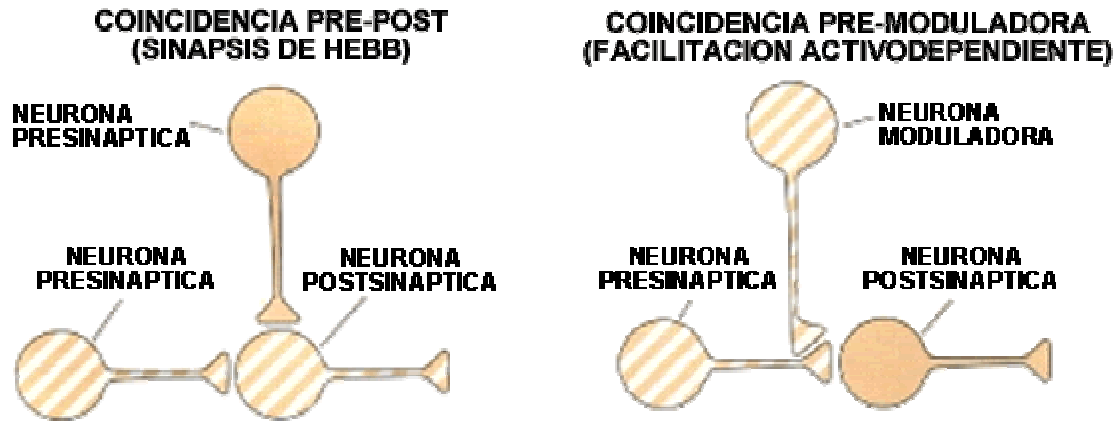


Figura 1.2.2 Cambios asociativos de las fuerzas sinápticas durante el aprendizaje

Todas las neuronas conducen la información de forma similar, ésta viaja a lo largo de axones en breves impulsos eléctricos, denominados potenciales de acción; los potenciales de acción que alcanzan una amplitud máxima de unos 100 mV y duran 1 ms, son resultado del desplazamiento a través de la membrana celular de iones de sodio dotados de carga positiva, que pasan desde el fluido extracelular hasta el citoplasma intracelular; la concentración extracelular de sodio supera enormemente la concentración intracelular.

La membrana en reposo mantiene un gradiente de potencial eléctrico de  $-70\text{mV}$ , el signo negativo se debe a que el citoplasma intracelular está cargado negativamente con respecto al exterior; los iones de sodio no atraviesan con facilidad la membrana en reposo, los estímulos físicos o químicos que reducen el gradiente de potencial, o que despolaricen la membrana, aumentan su permeabilidad al sodio y el flujo de este ión hacia el exterior acentúa la despolarización de la membrana, con lo que la permeabilidad al sodio se incrementa más aún.



Alcanzado un potencial crítico denominado “umbral”, la realimentación positiva produce un efecto regenerativo que obliga al potencial de membrana a cambiar de signo. Es decir, el interior de la célula se torna positivo con respecto al exterior, al cabo de 1 ms, la permeabilidad del sodio decae y el potencial de membrana retorna a  $-70\text{mv}$ , su valor de reposo. Tras cada explosión de actividad iónica, el mecanismo de permeabilidad del sodio se mantiene refractario durante algunos milisegundos; la tasa de generación de potenciales de acción queda así limitada a unos 200 impulsos por segundo, o menos.

Aunque los axones puedan parecer hilos conductores aislados, no conducen los impulsos eléctricos de igual forma, como hilos eléctricos no serían muy valiosos, pues su resistencia a lo largo del eje es demasiado grande y la resistencia de la membrana demasiado baja; la carga positiva inyectada en el axón durante el potencial de acción queda disipada uno o dos milímetros más adelante, para que la señal recorra varios centímetros es preciso regenerar frecuentemente el potencial de acción a lo largo del camino, la necesidad de reforzar repetidamente esta corriente eléctrica limita a unos 100 metros por segundo la velocidad máxima de viaje de los impulsos, tal velocidad es inferior a la millonésima de la velocidad de una señal eléctrica por un hilo de cobre.

Los potenciales de acción, son señales de baja frecuencia conducidas en forma muy lenta, éstos no pueden saltar de una célula a otra, la comunicación entre neuronas viene siempre mediada por transmisores químicos que son liberados en las sinapsis.



Un ejemplo de comunicación entre neuronas y del proceso químico de la liberación de neurotransmisores se ilustra en la figura 1.2.3.

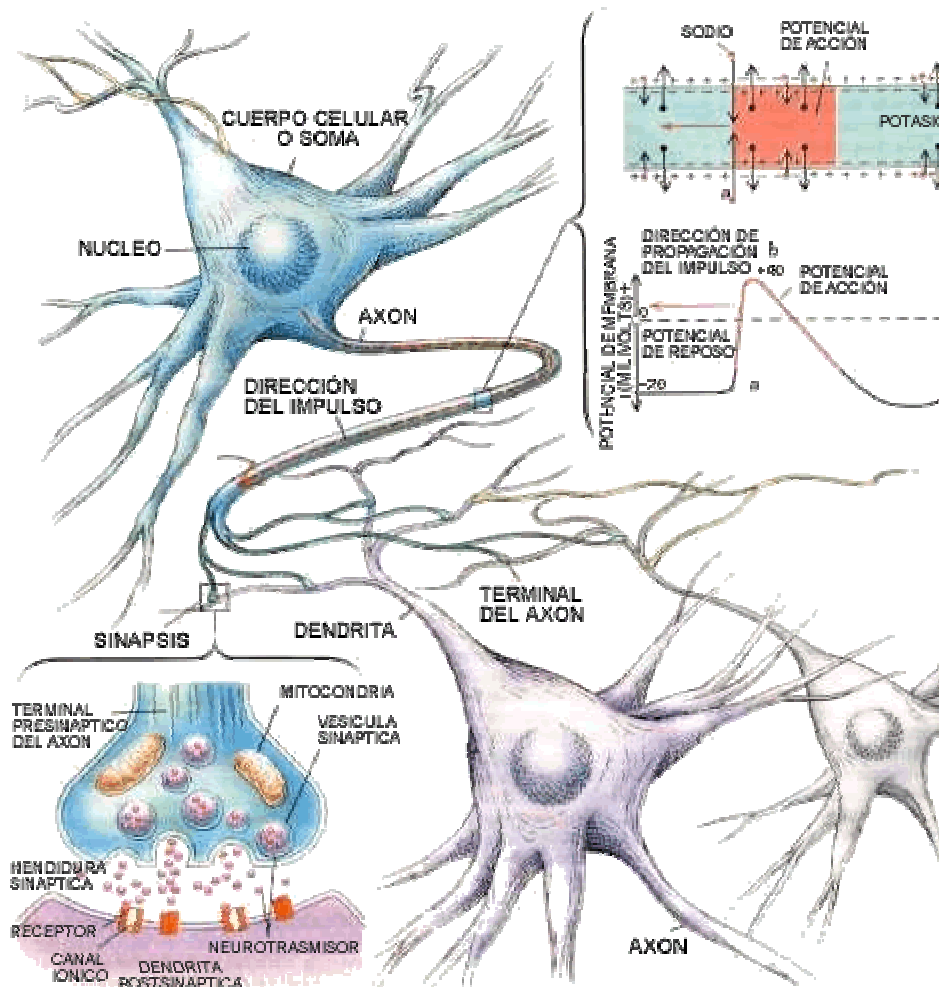


Figura 1.2.3 Comunicación entre neuronas

Cuando un potencial de acción llega al terminal de un axón son liberados transmisores alojados en diminutas vesículas, que después son vertidos en una hendidura de unos 20 nanómetros de anchura que separa la membrana presináptica de la postsináptica; durante el apogeo del potencial de acción, penetran iones de calcio en el terminal nervioso, su movimiento constituye la señal determinante de la exocitosis sincronizada, esto es la liberación coordinada de



moléculas neurotransmisoras. En cuanto son liberados, los neurotransmisores se enlazan con receptores postsinápticos, instando el cambio de la permeabilidad de la membrana.

Cuando el desplazamiento de carga hace que la membrana se aproxime al umbral de generación de potenciales de acción, se produce un efecto excitador y cuando la membrana resulta estabilizada en la vecindad el valor de reposo se produce un efecto inhibitor. Cada sinapsis produce sólo un pequeño efecto, para determinar la intensidad (frecuencia de los potenciales de acción) de la respuesta cada neurona ha de integrar continuamente hasta unas 1000 señales sinápticas, que se suman en el soma o cuerpo de la célula.

En algunas neuronas los impulsos se inician en la unión entre el axón y el soma, y luego se transmiten a lo largo del axón a otras células nerviosas. Cuando el axón está cerca de sus células destino, se divide en muchas ramificaciones que forman sinapsis con el soma o axones de otras células. Las sinapsis pueden ser excitatorias o inhibitorias según el neurotransmisor que se libere, cada neurona recibe de 10.000 a 100.000 sinapsis y su axón realiza una cantidad similar de sinapsis.

Las sinapsis se clasifican según su posición en la superficie de la neurona receptora en tres tipos: axo-somática, axo-dendrítica, axo-axónica. Los fenómenos que ocurren en la sinapsis son de naturaleza química, pero tienen efectos eléctricos laterales que se pueden medir.



En la figura 1.2.4 se visualiza el proceso químico de una sinapsis y los diferentes elementos que hacen parte del proceso tanto en la neurona presináptica, como en la postsináptica.

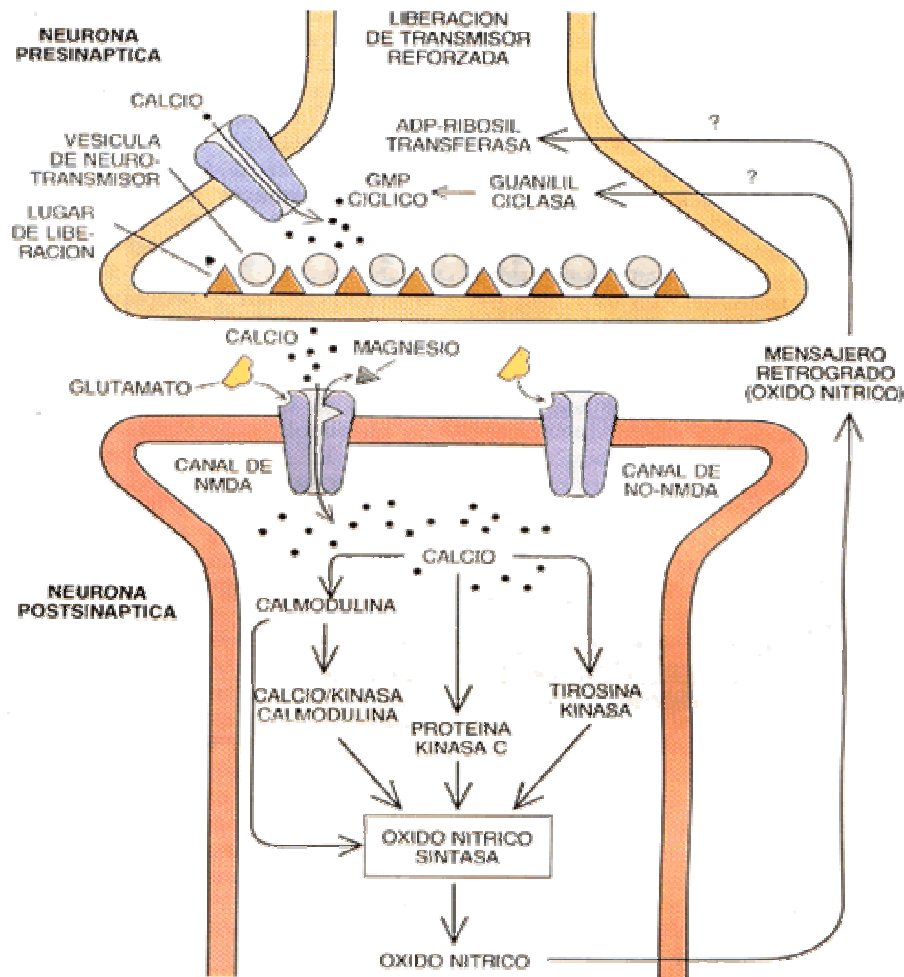


Figura 1.2.4 Proceso químico de una sinapsis

Las RNA no alcanzan la complejidad del cerebro, sin embargo hay dos aspectos similares entre redes biológicas y artificiales, primero los bloques de construcción de ambas redes son sencillos elementos computacionales (aunque las RNA son mucho más simples que las biológicas) altamente interconectados; segundo, las conexiones entre neuronas determinan la función de la red.



### 1.3 CARACTERÍSTICAS DE UNA RED NEURONAL ARTIFICIAL

El modelo de una neurona artificial es una imitación del proceso de una neurona biológica, puede también asemejarse a un sumador hecho con un amplificador operacional tal como se ve en la figura 1.3.1.

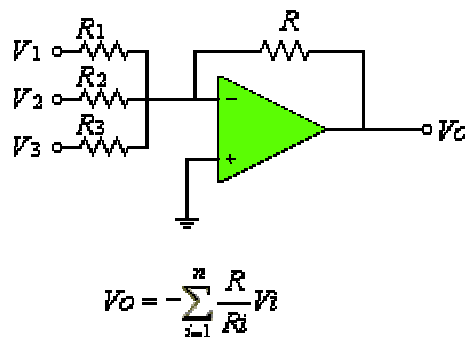


Figura 1.3.1 Neurona Artificial

Existen varias formas de nombrar una neurona artificial, es conocida como nodo, neuronodo, celda, unidad o elemento de procesamiento (PE); En la figura 1.3.1 se observa un PE en forma general y su similitud con una neurona biológica

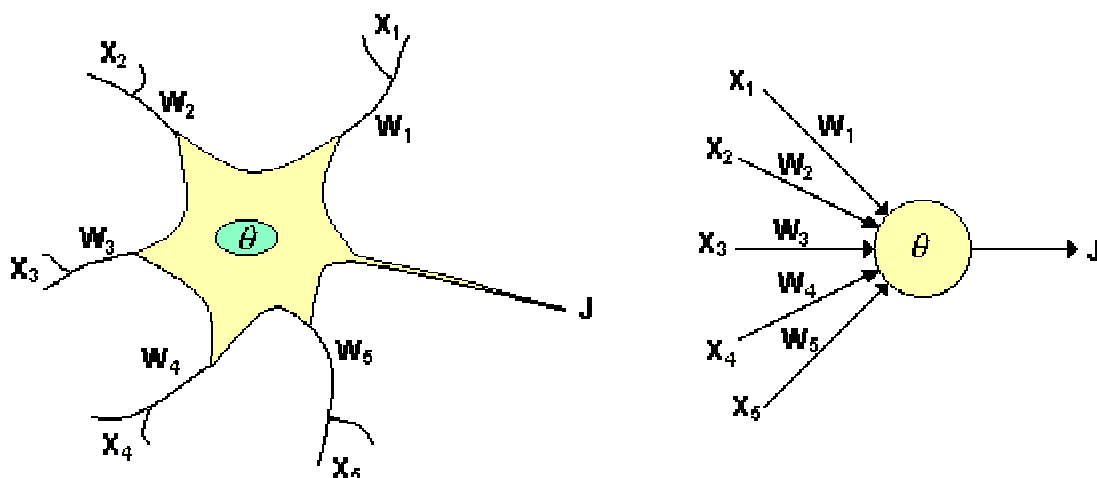


Figura 1.3.2 De la neurona biológica a la neurona artificial





De la observación detallada del proceso biológico se han hallado los siguientes análogos con el sistema artificial:

- Las entradas  $X_i$  representan las señales que provienen de otras neuronas y que son capturadas por las dendritas.
- Los pesos  $W_i$  son la intensidad de la sinapsis que conecta dos neuronas; tanto  $X_i$  como  $W_i$  son valores reales.
- $\theta$  es la función umbral que la neurona debe sobrepasar para activarse; este proceso ocurre biológicamente en el cuerpo de la célula.

Las señales de entrada a una neurona artificial  $X_1, X_2, \dots, X_n$  son variables continuas en lugar de pulsos discretos, como se presentan en una neurona biológica. Cada señal de entrada pasa a través de una ganancia o peso, llamado peso sináptico o fortaleza de la conexión cuya función es análoga a la de la función sináptica de la neurona biológica. Los pesos pueden ser positivos (excitatorios), o negativos (inhibitorios), el nodo sumatorio acumula todas las señales de entradas multiplicadas por los pesos o ponderadas y las pasa a la salida a través de una función umbral o función de transferencia. La entrada neta a cada unidad puede escribirse de la siguiente manera



$$neta_i = \sum_{i=1}^n W_i X_i = \vec{X} \vec{Y} \quad (1.3.1)$$

Una idea clara de este proceso se muestra en la figura 1.3.3, en donde puede observarse el recorrido de un conjunto de señales que entran a la red.

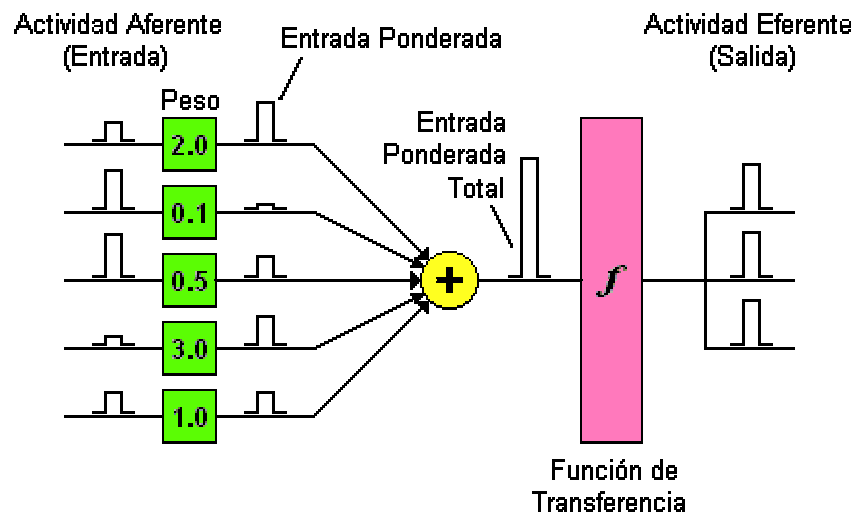


Figura 1.3.3 Proceso de una red neuronal

Una vez que se ha calculado la activación del nodo, el valor de salida equivale a

$$x_i = f_i(neta_i) \quad (1.3.2)$$

Donde  $f_i$  representa la función de activación para esa unidad, que corresponde a la función escogida para transformar la entrada  $neta_i$  en el valor de salida  $x_i$ , y que depende de las características específicas de cada red.



**1.3.1 Notación.** Una notación matemática estándar no ha sido aún establecida para las redes neuronales, ya que sus aplicaciones son útiles en muchos campos, Ingeniería, Física, Psicología y Matemáticas. En este trabajo se adoptó la siguiente convención para identificar las variables, de manera que fuera compatible con las diferentes áreas, siendo lo más sencilla posible:

- Valores escalares: se representarán por medio de letra minúscula itálica
- Vectores: se representarán con letra itálica minúscula en negrilla.
- Matrices: se representarán con letra mayúscula itálica en negrilla.

Para redes multicapa, los parámetros adoptarán la siguiente forma:

$$W_{s^c, s^c}^c$$

Donde  $c$ , es el número de la capa a la que corresponde dicho peso, y  $s$  representa las neuronas que participan en proceso.

Así  $W_{1,1}^2$  representa el peso de la segunda capa que comunica la primera neurona de dicha capa con la primera neurona de la primera capa. De igual manera el peso que representa la conexión desde la última neurona de la capa dos a la última neurona de la capa uno estará representado por:  $W_{s^2, s^1}^2$



Esta convención es adoptada para todos los parámetros de la red.

**1.3.2 Funciones de Transferencia.** Un modelo más académico que facilita el estudio de una neurona, puede visualizarse en la figura 1.3.4

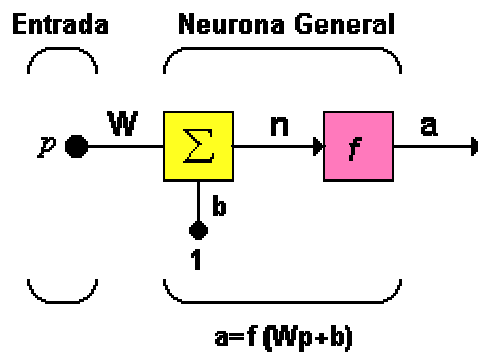


Figura 1.3.4 Neurona de una sola entrada

Las entradas a la red serán ahora presentadas en el vector  $p$ , que para el caso de una sola neurona contiene sólo un elemento,  $w$  sigue representando los pesos y la nueva entrada  $b$  es una ganancia que refuerza la salida del sumador  $n$ , la cual es la salida neta de la red; la salida total está determinada por la función de transferencia  $f$ , la cual puede ser una función lineal o no lineal de  $n$ , y que es escogida dependiendo de las especificaciones del problema que la neurona tenga que resolver; aunque las RNA se inspiren en modelos biológicos no existe ninguna limitación para realizar modificaciones en las funciones de salida, así que se encontrarán modelos artificiales que nada tienen que ver con las características del sistema biológico.



**1.3.2.1 Limitador fuerte (Hardlim).** La figura 1.3.5, muestra como esta función de transferencia acerca la salida de la red a cero, si el argumento de la función es menor que cero y la lleva a uno si este argumento es mayor que uno. Esta función crea neuronas que clasifican las entradas en dos categorías diferentes, característica que le permite ser empleada en la red tipo Perceptrón

$$a = \begin{cases} 1 & \text{si } n \geq 0 \\ 0 & \text{si } n < 0 \end{cases} \quad (1.3.3)$$

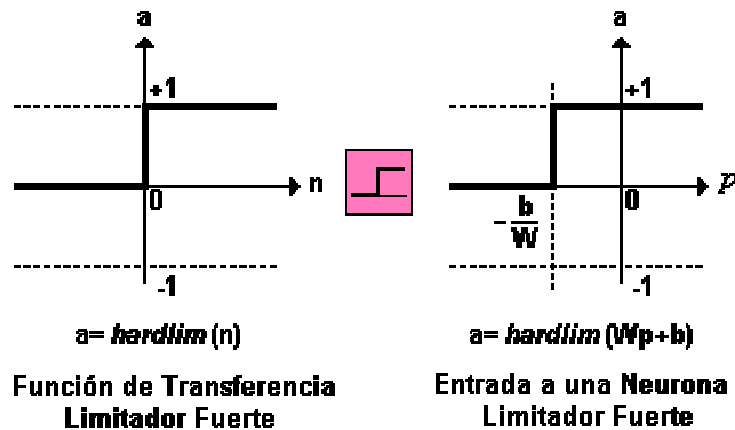


Figura 1.3.5 Función de transferencia Hardlim

El icono para la función Hardlim reemplazará a la letra  $f$  en la expresión general, cuando se utilice la función Hardlim.

Una modificación de esta función puede verse en la figura 1.3.6, la que representa la función de transferencia Hardlims que restringe el espacio de salida a valores entre 1 y -1.



$$a = \begin{cases} 1 & \text{si } n \geq 0 \\ -1 & \text{si } n < 0 \end{cases} \quad (1.3.4)$$

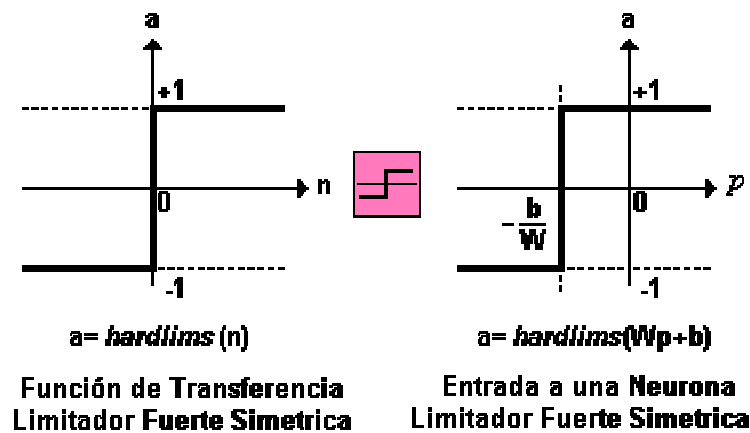


Figura 1.3.6 Función de transferencia Hardlims

**1.3.2.2** Función de transferencia lineal (purelin): La salida de una función de transferencia lineal es igual a su entrada,

$$a = n \quad (1.3.5)$$

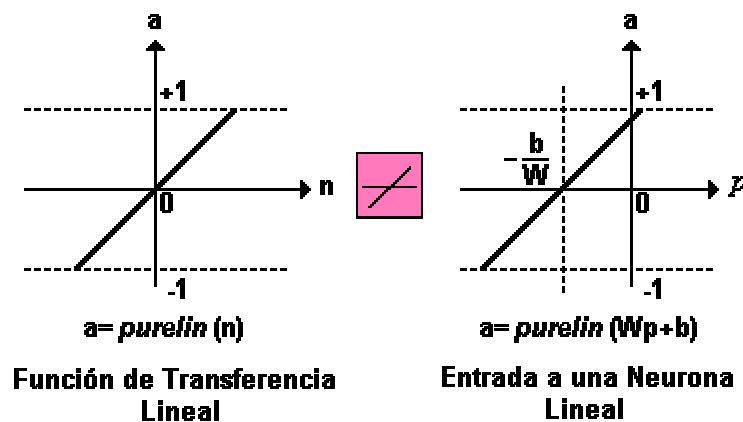


Figura 1.3.7 Función de transferencia lineal

En la gráfica del lado derecho de la figura 1.3.7, puede verse la relación entre la salida  $a$  de la red, y la entrada  $p$ , teniendo en cuenta el valor de ganancia  $b$ ;



neuronas que emplean esta función de transferencia son utilizadas en la red tipo Adaline.

**1.3.2.3 Función de transferencia sigmoideal (logsig).** Esta función toma los valores de entrada, los cuales pueden oscilar entre más y menos infinito, y restringe la salida a valores entre cero y uno, de acuerdo a la expresión

$$a = \frac{1}{1 + e^{-n}} \quad (1.3.6)$$

Esta función es comúnmente usada en redes multicapa, como la Backpropagation, en parte porque la función logsig es diferenciable.

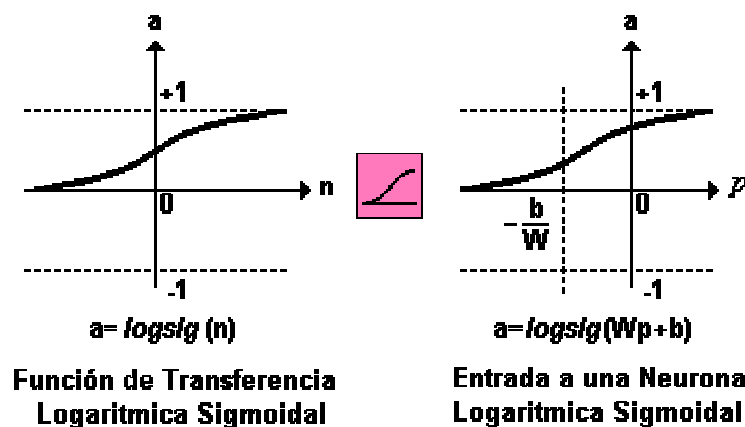


Figura 1.3.8 Función de transferencia sigmoideal

La tabla 1.3.1 hace una relación de las principales funciones de transferencia empleadas en redes neuronales.












Nombre	Relación Entrada /Salida	Icono	Función
Limitador Fuerte	$a = 0 \quad n < 0$ $a = 1 \quad n \geq 0$		<i>hardlim</i>
Limitador Fuerte Simétrico	$a = -1 \quad n < 0$ $a = +1 \quad n \geq 0$		<i>hardlims</i>
Lineal Positiva	$a = 0 \quad n < 0$ $a = n \quad 0 \leq n$		<i>poslin</i>
Lineal	$a = n$		<i>purelin</i>
Lineal Saturado	$a = 0 \quad n < 0$ $a = n \quad 0 \leq n \leq 1$ $a = 1 \quad n > 1$		<i>satlin</i>
Lineal Saturado Simétrico	$a = -1 \quad n < -1$ $a = n \quad -1 \leq n \leq 1$ $a = +1 \quad n > 1$		<i>satlins</i>
Sigmoidal Logarítmico	$a = \frac{1}{1 + e^{-n}}$		<i>logsig</i>
Tangente Sigmoidal Hiperbólica	$a = \frac{e^n - e^{-n}}{e^n + e^{-n}}$		<i>tansig</i>
Competitiva	$a = 1 \quad \text{Neurona con } n \text{ max}$ $a = 0 \quad \text{El resto de neuronas}$		<i>compet</i>

Tabla 1.3.1 Funciones de Transferencia

**1.3.3 Topología de una Red.** Típicamente una neurona tiene más de una entrada; en la figura 1.3.9 se observa una neurona con R entradas; las entradas individuales  $p_1, p_2, \dots, p_R$  son multiplicadas por los pesos correspondientes  $w_{1,1}, w_{1,2}, \dots, w_{1,R}$  pertenecientes a la matriz de pesos  $W$ .





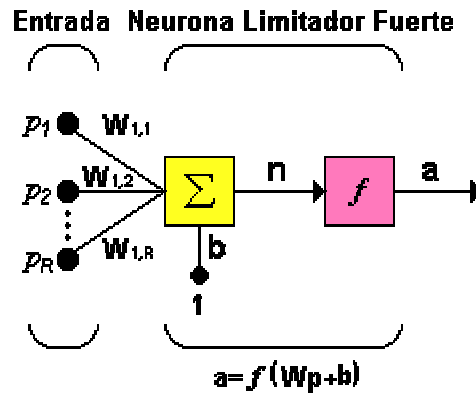


Figura 1.3.9 Neurona con múltiples entradas

La neurona tiene una ganancia  $b$ , la cual llega al mismo sumador al que llegan las entradas multiplicadas por los pesos, para formar la salida  $n$ ,

$$n = w_{1,1}p_1 + w_{1,2}p_2 + \dots + w_{1,R}p_R + b \quad (1.3.7)$$

Esta expresión puede ser escrita en forma matricial

$$n = Wp + b \quad (1.3.8)$$

Los subíndices de la matriz de pesos representan los términos involucrados en la conexión, el primer subíndice representa la neurona destino y el segundo representa la fuente de la señal que alimenta a la neurona. Por ejemplo, los índices de  $w_{1,2}$  indican que este peso es la conexión desde el segundo elemento del vector de entrada a la primera neurona. Esta convención se hace más útil cuando hay más de una neurona, o cuando se tiene una red con demasiados



parámetros, en este caso la notación de la figura 1.3.9 puede resultar inapropiada y se prefiere emplear la notación compacta de la figura 1.3.10

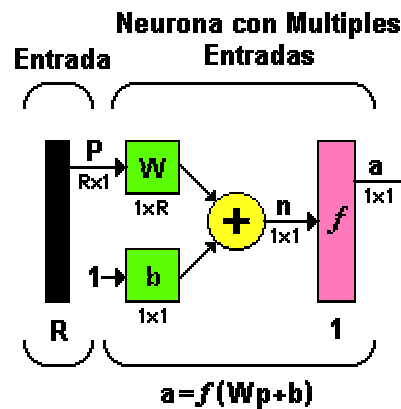


Figura 1.3.10 Neurona con múltiples entradas, notación abreviada

El vector de entrada  $p$  es representado por la barra sólida vertical a la izquierda, las dimensiones de  $p$  son mostradas en la parte inferior de la variable como  $R \times 1$  indicando que el vector de entrada es un vector fila de  $R$  elementos. Las entradas van a la matriz de pesos  $W$ , la cual tiene  $R$  columnas y sólo una fila para el caso de una sola neurona. Una constante 1 entra a la neurona multiplicada por la ganancia escalar  $b$ . La salida de la red  $a$  es en este caso un escalar, si la red tuviera más de una neurona  $a$ , sería un vector.

Dentro de una red neuronal, los elementos de procesamiento se encuentran agrupados por capas, una capa es una colección de neuronas; de acuerdo a la ubicación de la capa en la RNA, esta recibe diferentes nombres



**Capa de entrada:** Recibe las señales de la entrada de la red, algunos autores no consideran el vector de entrada como una capa pues allí no se lleva a cabo ningún proceso, en este trabajo se seguirá este criterio.

**Capas ocultas:** Estas capas son aquellas que no tienen contacto con el medio exterior, sus elementos pueden tener diferentes conexiones y son éstas las que determinan las diferentes topologías de la red.

**Capa de salida:** Recibe la información de la capa oculta y transmite la respuesta al medio externo.

Una red de una sola capa con un número  $S$  de neuronas, se observa en la figura 1.3.11 en la cual, cada una de las  $R$  entradas es conectada a cada una de las neuronas, la matriz de pesos tiene ahora  $S$  filas.

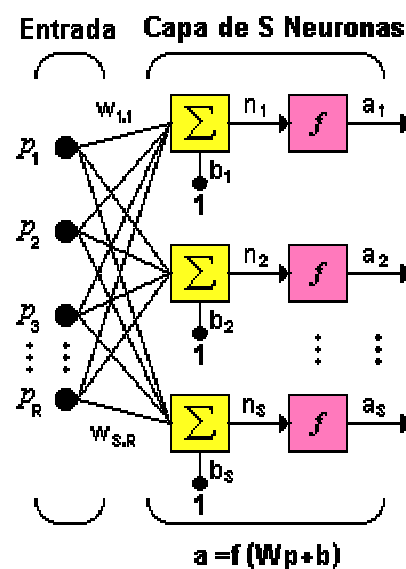


Figura 1.3.11 Capa de S neuronas



La capa incluye la matriz de pesos, los sumadores, el vector de ganancias, la función de transferencia y el vector de salida. Esta misma capa se observa en notación abreviada en la figura 1.3.12

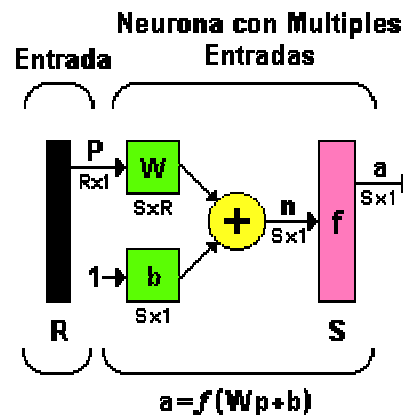


Figura 1.3.12 Capa de S neuronas con notación abreviada

En la figura 1.3.12 se han dispuesto los símbolos de las variables de tal manera que describan las características de cada una de ellas, por ejemplo la entrada a la red es el vector  $p$  cuya longitud  $R$  aparece en su parte inferior,  $W$  es la matriz de pesos con dimensiones  $S \times R$  expresadas debajo del símbolo que la representa dentro de la red,  $a$  y  $b$  son vectores de longitud  $S$  el cual, como se ha dicho anteriormente representa el número de neuronas de la red.

Ahora, si se considera una red con varias capas, o red multicapa, cada capa tendrá su propia matriz de peso  $W$ , su propio vector de ganancias  $b$ , un vector de entradas netas  $n$ , y un vector de salida  $a$ . La versión completa y la versión en notación abreviada de una red de tres capas, pueden ser visualizadas en las figuras 1.3.13 y 1.3.14, respectivamente.



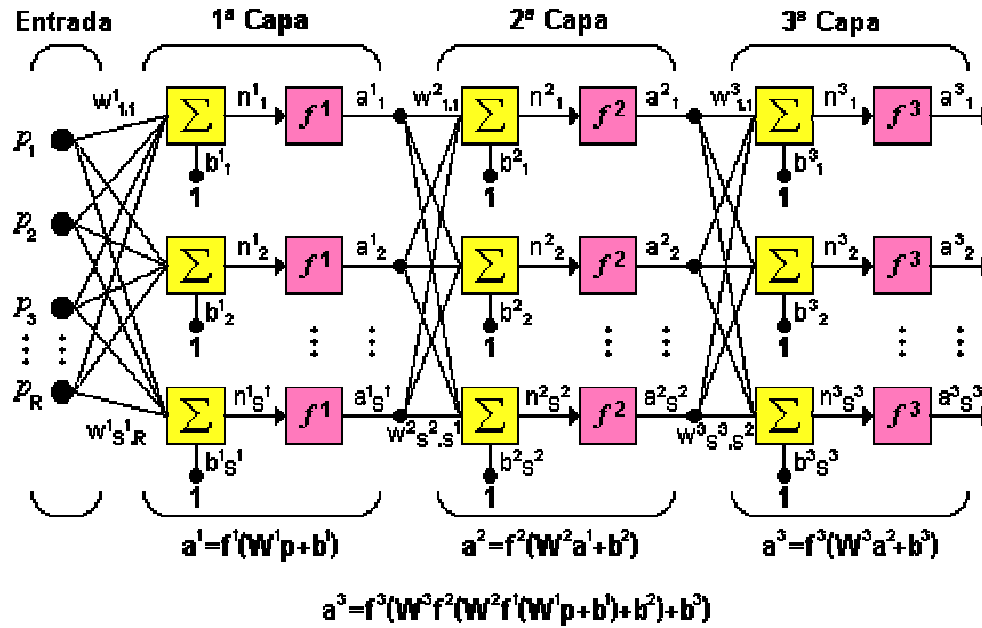


Figura 1.3.13 Red de tres capas

Para esta red se tienen  $R$  entradas,  $S^1$  neuronas en la primera capa,  $S^2$  neuronas en la segunda capa, las cuales pueden ser diferentes; las salidas de las capas 1 y 2 son las entradas a las capas 2 y 3 respectivamente, así la capa 2 puede ser vista como una red de una capa con  $R=S^1$  entradas,  $S_1=S^2$  neuronas y una matriz de pesos  $W^2$  de dimensiones  $S^1 \times S^2$

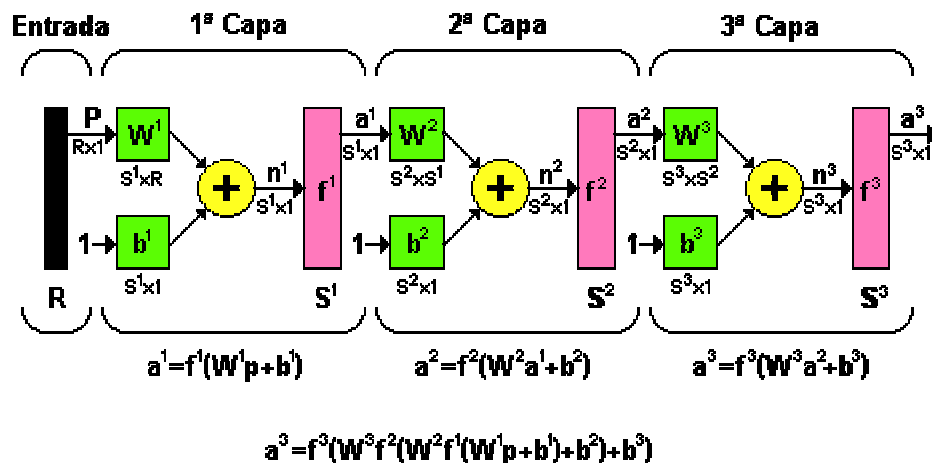


Figura 1.3.14 Red de tres capas con notación abreviada



Las redes multicapa son más poderosas que las redes de una sola capa, por ejemplo, una red de dos capas que tenga una función sigmoideal en la primera capa y una función lineal en la segunda, puede ser entrenada para aproximar muchas funciones de forma aceptable, una red de una sola capa no podría hacer esto como se verá en capítulos posteriores.

Un tipo de redes, un poco diferente a las que se han estudiado hasta el momento, son las redes recurrentes, estas contienen una realimentación hacia atrás o retroalimentación, es decir algunas de sus salidas son conectadas a sus entradas. Un tipo de red recurrente de tiempo discreto es mostrado en la figura 1.3.15.

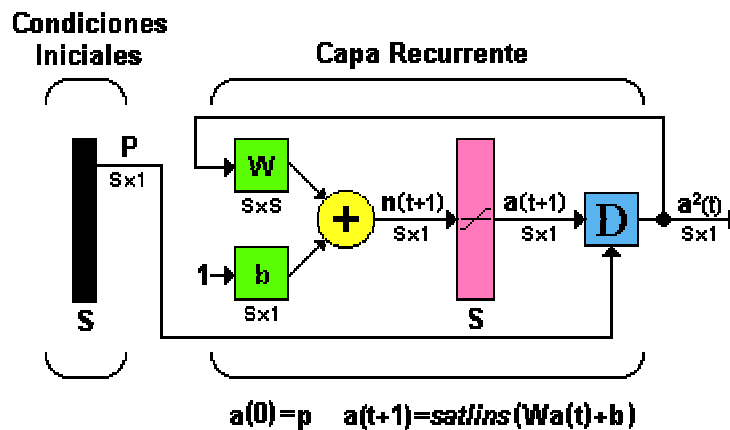


Figura 1.3.15 Redes Recurrentes

Para este tipo particular de red el vector  $p$  suple las condiciones iniciales ( $a(0) = p$ ), y la salida está determinada por:

$$a(1) = satlins(Wa(0) + b), \quad a(2) = satlins(Wa(1) + b) \quad (1.3.9)$$



Donde  $a(1)$  y  $a(2)$ , corresponden a la salida de la red para el primer y segundo intervalo de tiempo, respectivamente. La red alcanzará su estado estable cuando la salida para un instante de tiempo sea la misma salida del instante de tiempo anterior.

Las redes recurrentes son potencialmente más poderosas que las redes con realimentación hacia delante. En este tipo de redes se introducen también dos nuevos conceptos, el bloque de retardo de la figura 1.3.16 y el bloque integrador de la figura 1.3.17

- Retardo

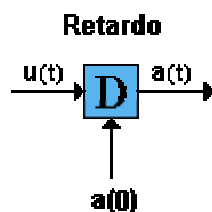


Figura 1.3.16 Bloque de retardo

$$a(t) = u(t-1) \quad (1.3.10)$$

La salida del bloque de retardo es el valor de entrada retrasado en un paso de tiempo, este bloque requiere que la salida sea inicializada con el valor  $a(0)$  para el tiempo  $t=0$ ;  $a(0)$  se convierte en la salida de la red para el instante de tiempo inicial.



- Integrador

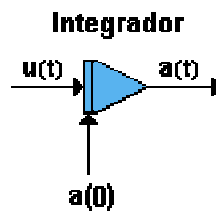


Figura 1.3.17 Bloque integrador

La salida del integrador es calculada de acuerdo a la expresión

$$a(t) = \int_0^t u(\tau) d\tau + a(0) \quad (1.3.11)$$

Una red recurrente cuya implementación necesita un bloque integrador se ilustra en la figura 2.6.4.

En general las redes neuronales se pueden clasificar de diversas maneras, según su topología, forma de aprendizaje (supervisado o no supervisado), tipos de funciones de activación, valores de entrada (binarios o continuos); un resumen de esta clasificación se observa en la figura 1.3.18

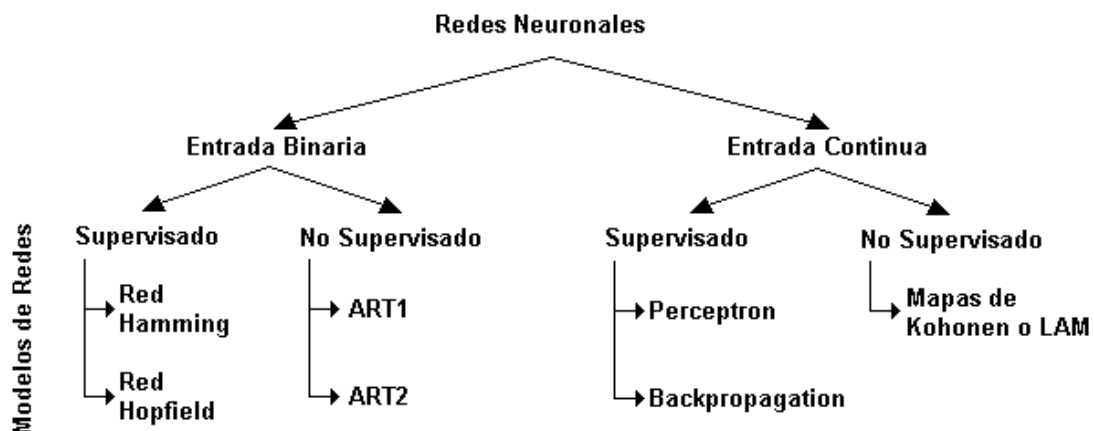


Figura 1.3.18 Clasificación de las Redes Neuronales

