

Instituto Tecnológico de Nuevo Laredo

Ingeniería en Sistemas Computacionales

Inteligencia Artificial

Ing. Bruno López Takeyas

Unidad I

Tema 2 Redes Neuronales

Sergio Alfredo Santos Ramírez	01100304
Gerson Antonio García Carrasco	01100208
Renato García Jiménez	01100209
Isidro Gómez Rodríguez	01100219
Gilberto Antonio Anwar Peña Nuño	01100275

Nuevo Laredo Tamaulipas a 24 de Agosto del 2005

INDICE

Introducción	3
El modelo Biológico	5
Ejemplos de Problemas de Interés para científicos e ingenieros	7
Elementos de una red neuronal artificial	9
Formas de conexión entre neuronas	14
Topologías de redes neuronales	15
Mecanismos de Aprendizaje	16
Descripción de las topologías ADALINE y Backpropagation	20
Características de las redes neuronales	24
Ventajas de las redes neuronales	24
Aplicaciones de las redes neuronales	25
Redes Neuronales y Control	26
Futuro de las redes neuronales	27
Bibliografía	28

Introducción

Las redes neuronales como su nombre lo indica pretenden imitar a pequeñísima escala la forma de funcionamiento de las neuronas que forman el cerebro humano. Todo el desarrollo de las redes neuronales tiene mucho que ver con la neurofisiología, no en vano se trata de imitar a una neurona humana con la mayor exactitud posible. Entre los pioneros en el modelado de neuronas se encuentra Warren McCulloch y Walter Pitts. Estos dos investigadores propusieron un modelo matemático de neurona. En este modelo cada neurona estaba dotada de un conjunto de entradas y salidas. Cada entrada está afectada por un peso. La activación de la neurona se calcula mediante la suma de los productos de cada entrada y la salida es una función de esta activación. La principal clave de este sistema se encuentra en los pesos de las diferentes entradas. Como se ha visto, las entradas son modificadas por el peso y las salidas son función de estas modificaciones. Esto nos lleva a concluir que los pesos influyen de forma decisiva en la salida y por lo tanto pueden ser utilizados para controlar la salida que se desea.

En realidad cuando se tienen interconectadas muchas de estas neuronas artificiales lo que se hace inicialmente es entrenar el sistema. El entrenamiento consiste en aplicar unas entradas determinadas a la red y observar la salida que produce. Si la salida que produce no se adecua a la que se esperaba, se ajustan los pesos de cada neurona para interactivamente ir obteniendo las respuestas adecuadas del sistema. A la red se le somete a varios ejemplos representativos, de forma que mediante la modificación de los pesos de cada neurona, la red va "aprendiendo".

Redes neuronales

Las redes neuronales son sistemas compuesto por un gran número de elementos básicos (Artificial Neurons), agrupados en capas (Layers) y que se encuentran altamente interconectados (Synapses); Esta estructura posee varias entradas y salidas, las cuales serán entrenadas para reaccionar (valores O), de una manera deseada, a los estímulos de entrada (valores I).

Estos sistemas emulan, de una cierta manera, al cerebro humano. Requieren aprender a comportarse (Learning) y alguien debe encargarse de enseñarles o entrenarles (Training), en base a un conocimiento previo del entorno del problema.

Redes neuronales artificiales son redes interconectadas masivamente en paralelo de elementos simples (usualmente adaptativos) y con organización jerárquica, las cuales intentan interactuar con los objetos del mundo real del mismo modo que lo hace el sistema nervioso biológico.

Un poco de historia

En 1956 se organizó en Dartmouth la primera conferencia sobre IA. Aquí se discutió el uso potencial de las computadoras para simular "todos los aspectos del aprendizaje o

cualquier otra característica de la inteligencia" y se presentó la primera simulación de una red neuronal, aunque todavía no se sabían interpretar los datos resultantes.

En 1959, Widrow (Widrow 1959) publica una teoría sobre la adaptación neuronal y unos modelos inspirados en esa teoría, el Adaline (Adaptative Linear Neuron) y el Madaline (Multiple Adaline). Estos modelos fueron usados en numerosas aplicaciones y permitieron usar, por primera vez, una red neuronal en un problema importante del mundo real: filtros adaptativos para eliminar ecos en las líneas telefónicas.

En 1962, Roseblatt (Roseblatt 1962) publica los resultados de un ambicioso proyecto de investigación, el desarrollo del Perceptrón, un identificador de patrones ópticos binarios, y salida binaria. Las capacidades del Perceptrón se extendieron al desarrollar la regla de aprendizaje delta, que permitía emplear señales continuas de entrada y salida.

1969, Minsky y Papert (Minsky & Papert 1969)) realizan una seria crítica del Perceptrón, revelando serias limitaciones, como su incapacidad para representar la función XOR, debido a su naturaleza lineal. Este trabajo creó serias dudas sobre las capacidades de los modelos conexionistas y provocó una caída en picado de las investigaciones.

Años 70: a pesar del duro golpe que supuso el trabajo de Minsky y Papert para las investigaciones en computación conexionista, un puñado de investigadores siguió trabajando y desarrollando nuevas ideas:

Anderson (Anderson, Silverstein, Ritz & Jomnes 1977) estudia y desarrolla modelos de memorias asociativas. Destaca el autoasociador lineal conocido como modelo brain-state-in-a-box (BSB).

Kohonen (Kohonen 1984) continua el trabajo de Anderson y desarrolla modelos de aprendizaje competitivo basados en el principio de inhibición lateral. Su principal aportación consiste en un procedimiento para conseguir que unidades físicamente adyacentes aprendieran a representar patrones de entrada similares; a las redes basadas en este procedimiento se las denomina redes de Kohonen.

Grossberg (Grossberg 1987) realizó un importante trabajo teórico - matemático tratando de basarse en principios fisiológicos; aportó importantes innovaciones con su modelo ART (Adaptative Resonance Theory) y, junto a Cohen, elabora un importante teorema sobre la estabilidad de las redes recurrentes en términos de una función de energía.

Años 80: En esta década se produce el renacimiento del interés por el campo gracias sobre todo al trabajo del grupo PDP y las aportaciones de Hopfield.

Rumelhart, McClelland & Hinton crean el grupo PDP (Parallel Distributed Processing). Como resultado de los trabajos de este grupo salieron los manuales (Rumelhart & McClelland 1986 y 1988) con más influencia desde la crítica de Minsky y Papert. Destaca el capítulo dedicado al algoritmo de retropropagación, que soluciona los

problemas planteados por Minsky y Papert y extiende enormemente el campo de aplicación de los modelos de computación conexionistas.

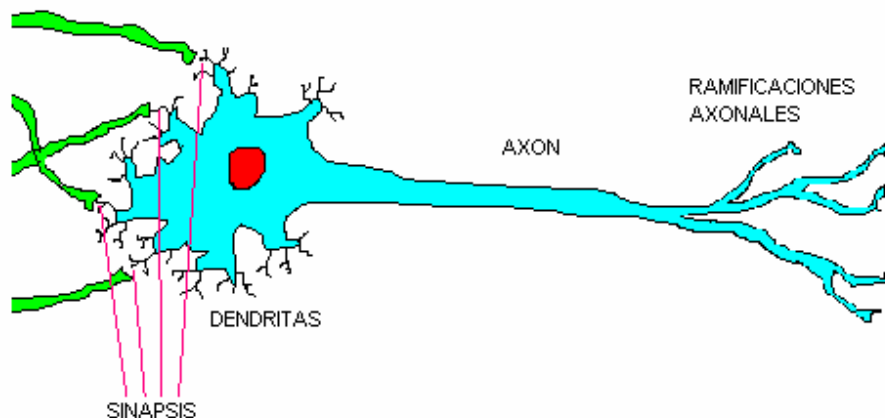
Hopfield (Hopfield 1982) elabora un modelo de red consistente en unidades de proceso interconectadas que alcanzan mínimos energéticos, aplicando los principios de estabilidad desarrollados por Grossberg. El modelo de Hopfield resultó muy ilustrativo sobre los mecanismos de almacenamiento y recuperación de la memoria. Su entusiasmo y claridad de presentación dieron un nuevo impulso al campo y provocaron el incremento de las investigaciones.

Otros desarrollos destacables de esta década son la máquina de Boltzmann (Hinton & Sejnowski 1986) y los modelos BAM (Kosko 1987)

El modelo biológico

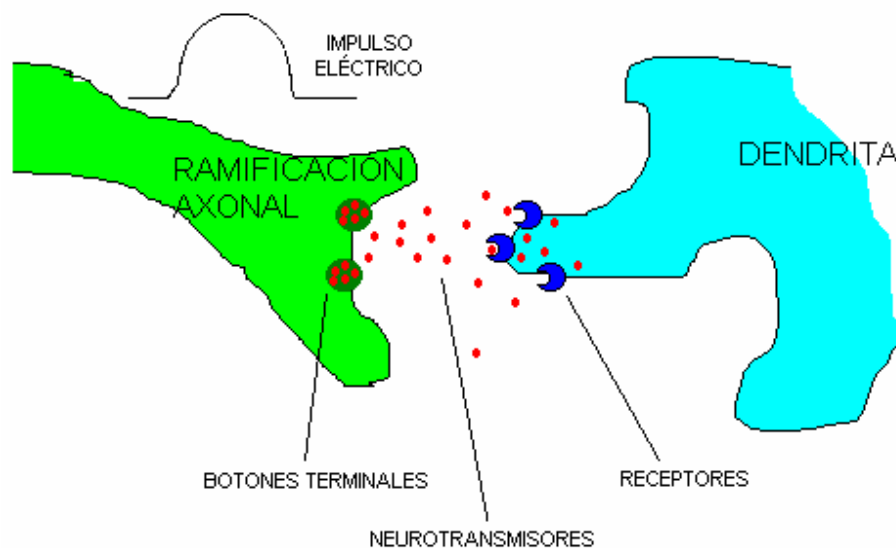
A grandes rasgos, recordemos que el cerebro humano se compone de decenas de billones de neuronas interconectadas entre sí formando circuitos o redes que desarrollan funciones específicas.

Una neurona típica recoge señales procedentes de otras neuronas a través de delicadas estructuras llamadas dendritas. La neurona emite impulsos de actividad eléctrica a lo largo de una fibra larga y delgada denominada axón.



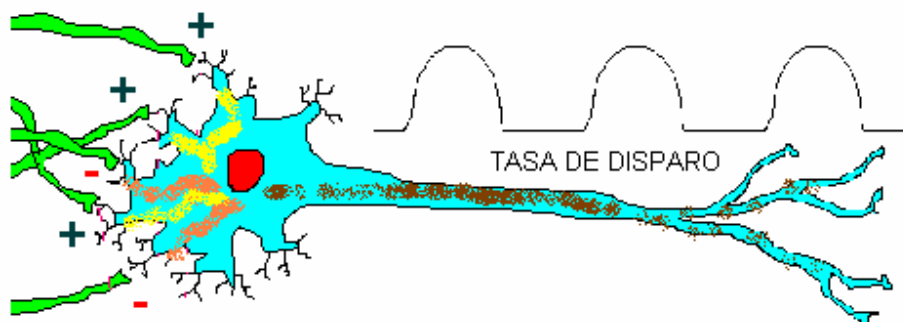
Neurona y conexiones sinápticas

Las extremidades de estas ramificaciones llegan hasta las dendritas de otras neuronas y establecen unas conexiones llamadas sinápsis, en las cuales se produce una transformación del impulso eléctrico en un mensaje neuroquímico, mediante la liberación de unas sustancias llamadas neurotransmisores.



Detalle de una sinápsis

El efecto de los neurotransmisores sobre la neurona receptora puede ser excitatorio o inhibitorio, y es variable (la intensidad del efecto depende de numerosos factores que no sería oportuno describir aquí), de manera que podemos hablar de la fuerza o efectividad de una sinápsis. Las señales excitatorias e inhibitorias recibidas por una neurona se combinan, y en función de la estimulación total recibida, la neurona toma un cierto nivel de activación, que se traduce en la generación de breves impulsos nerviosos con una determinada frecuencia o tasa de disparo, y su propagación a lo largo del axón hacia las neuronas con las cuales sinapta.



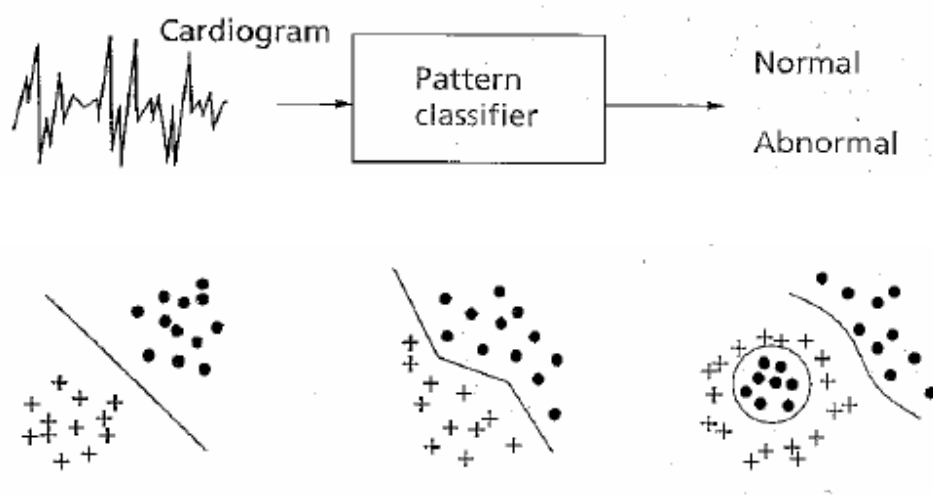
Activación y disparo de una neurona

De esta manera la información se transmite de unas neuronas a otras y va siendo procesada a través de las conexiones sinápticas y las propias neuronas. El aprendizaje de las redes neuronales se produce mediante la variación de la efectividad de las sinápsis, de esta manera cambia la influencia que unas neuronas ejercen sobre otras, de aquí se deduce que la arquitectura, el tipo y la efectividad de las conexiones en un momento dado, representan en cierto modo la memoria o estado de conocimiento de la red.

Ejemplos de Problemas de Interés para científicos e ingenieros:

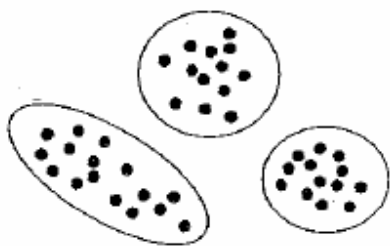
Clasificación de patrones:

La tarea de la clasificación de patrones (como una onda de voz o un símbolo manuscrito) es asignar un patrón de entrada representado por un vector a una de varias clases pre especificado. Algunas aplicaciones conocidas incluyen el reconocimiento de caracteres, reconocimiento de voz, Clasificación de Ondas EEG, clasificación de células sanguíneas, inspección de circuitos impresos.



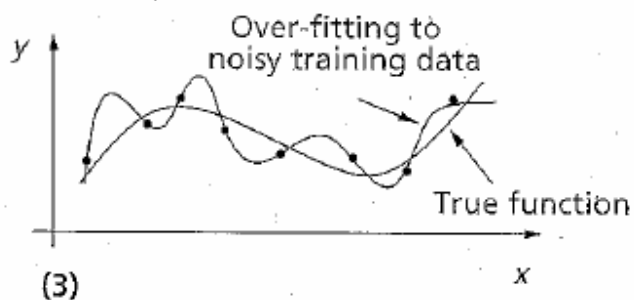
Clustering/Categorización.

En clustering, también es conocido como clasificación de patrón sin supervisor, no hay datos de entrenamiento con etiquetas de clases conocidas. Un algoritmo de clustering explora la similitud entre los patrones con patrones similares de aplicaciones de cluster, incluyendo la minería de datos, compresión de datos y análisis exploratorio de datos.



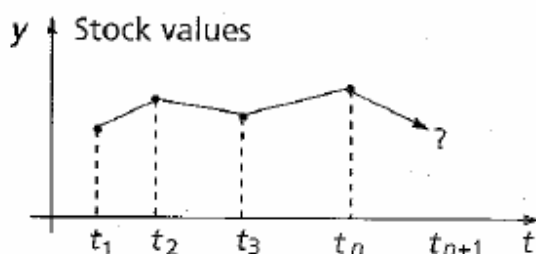
Aproximación de Función.

Suponga un conjunto de n patrones de entrenamiento $\{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}$ se han generado de una función desconocida x (sujeta a ruido). La tarea de la aproximación de funciones es encontrar un estimado, u , de una función x . Varios Modelos de problemas de ingeniería y científicos requieren aproximación de funciones.



Predicción

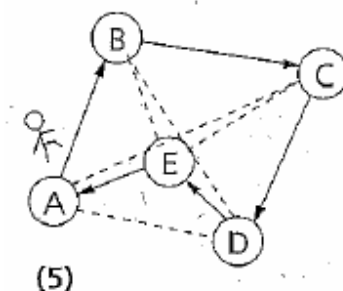
Dado un conjunto de n muestras $\{y(t_1), y(t_2), \dots, y(t_n)\}$ en una secuencia de tiempo, $t_1, t_2, \dots, t_{n+1}$. La predicción tiene un impacto significativo en la toma de decisiones de negocios, ciencia e ingeniería. Predicciones de Mercado y predicciones meteorológicas son aplicaciones típicas de técnicas de predicción.



Optimización

Una amplia variedad de problemas matemáticos, estadísticos, de ingeniería, ciencia, medicina y económicos pueden resolverse con la técnica de optimización.

La meta de un algoritmo de optimización es encontrar la solución satisfactoria a un conjunto de constantes como los objetivos de una función maximizados o minimizados. El problema de la agencia de viajes es un ejemplo clásico.



Elementos de una Red Neuronal Artificial

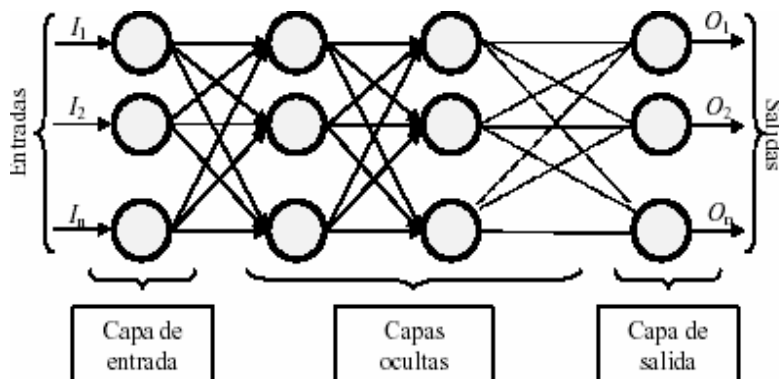
Las redes neuronales son modelos que intentan reproducir el comportamiento del cerebro. Los mismos constan de dispositivos elementales de proceso: las neuronas. A partir de ellas, se pueden generar representaciones específicas, de tal forma que un estado conjunto de ellas puede significar una letra, un número u otro objeto. Generalmente se pueden encontrar tres tipos de neuronas:

Aquellas que reciben estímulos externos relacionados con el aparato sensorial, que tomarán la información de entrada.

Dicha información se transmite a ciertos elementos internos que se ocupan de su procesamiento. Es en las sinapsis y neuronas correspondientes a este segundo nivel donde se genera cualquier tipo de representación interna de información. Como no tienen relación directa con la información de entrada ni con la salida, estos elementos se denominan unidades ocultas.

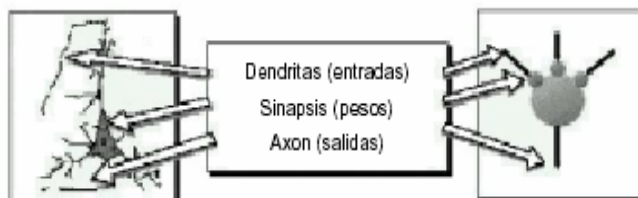
Una vez finalizado el período de procesado, la información llega a las unidades de salida, cuya misión es dar la respuesta al sistema.

A continuación se puede ver en la siguiente figura, un esquema de una red neuronal:



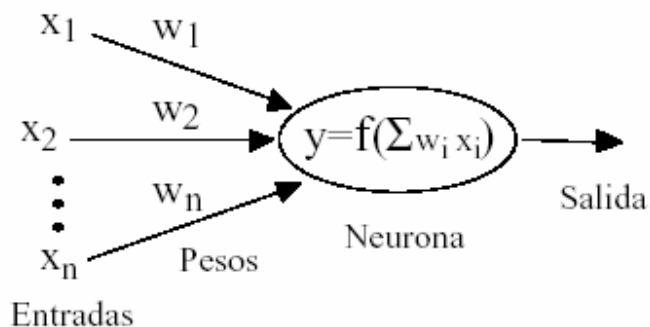
La misma está constituida por neuronas interconectadas y arregladas en tres capas (esto último puede variar). Los datos ingresan por medio de la “capa de entrada”, pasan a través de la “capa oculta” y salen por la “capa de salida”. Cabe mencionar que la capa oculta puede estar constituida por varias capas.

En la siguiente figura se compara una neurona biológica con una neurona artificial. En la misma se pueden observar las similitudes entre ambas (tienen entradas, utilizan pesos y generan salidas).



La neurona artificial pretende mimetizar las características más importantes de las neuronas biológicas. Cada neurona i -ésima está caracterizada en cualquier instante por un valor numérico denominado valor o estado de activación asociado a cada unidad, existe una función de salida, que transforma el estado actual de activación en una señal de salida. Dicha señal es enviada a través de los canales de comunicación unidireccionales a otras unidades de la red; estos canales la señal se modifica de acuerdo con la sinápsis asociada a cada uno de ellos según determinada regla. Las señales moduladas que han llegado a la unidad j -ésima se combinan entre ellas, generando así la entrada total.

Una función de activación, F , determina el nuevo estado de activación de la neurona, teniendo en cuenta la entrada total calculada y el anterior estado de activación.



La dinámica que rige la actualización de los estados de las unidades puede ser de dos tipos: asíncrono y modo síncrono. En el primer caso, las neuronas evalúan su estado continuamente según les va llegando información, y lo hacen de forma independiente. En el segundo caso, la información llega de forma continua, pero los cambios se realizan simultáneamente, como si existiera un reloj interno que decidiera cuando cambiar su estado. Los sistemas biológicos quedan probablemente entre ambas posibilidades.

Unidades de proceso: La neurona artificial

Si se tienen N unidades (neuronas), podemos ordenarlas arbitrariamente y designar la j -ésima unidad como U_j . Su trabajo es simple y único, y consiste en recibir las entradas de las células vecinas y calcular un valor de salida, el cual es enviado a todas las células restantes.

En cualquier sistema que se esté modelando, es útil caracterizar tres tipos de unidades: entradas, salidas y ocultas. Las unidades de entrada reciben señales del entorno, éstas pueden ser provenientes de sensores o de otros sectores del sistema. Las unidades de salida envían la señal fuera del sistema; éstas pueden controlar directamente potencias u otros sistemas. Las unidades ocultas son aquellas cuyas entradas y salidas se encuentran dentro del sistema; es decir no tienen contacto con el exterior.

Se conoce como nivel o capa a un conjunto de neuronas cuyas entradas provienen de la misma fuente, y cuyas salidas se dirigen a un mismo destino.

Estado de Activación

El procesamiento que realiza la red se ve como la evolución de un patrón de activación en el conjunto de unidades que lo componen a través del tiempo.

Todas las neuronas que componen la red se hallan en cierto estado. Podemos decir que hay dos posibles estados, reposo y excitado, a los que denominaremos estados de activación y a cada uno de los cuales se le asigna un valor. Los valores de activación pueden ser continuos o discretos. Además pueden ser limitados o ilimitados. Si son discretos, suelen tomar un conjunto pequeño de valores o bien valores binarios. En notación binaria, un estado activo se indicaría por un 1, y se caracteriza por la emisión de un impulso por parte de la neurona (potencial de acción), mientras que un estado pasivo

se indicaría por un 0. En otros modelos se considera un conjunto continuo de estados de activación, en cuyo caso se asigna un valor entre $[0,1]$ o en el intervalo $[-1,1]$, generalmente siguiendo una función sigmoideal.

Los criterios o reglas que siguen las neuronas para alcanzar tales estados dependen de dos factores:

Dado que las propiedades macroscópicas de las redes neuronales no son producto de actuación de elementos individuales, es necesario tener idea del mecanismo de interacción entre las neuronas. El estado de activación estará fuertemente influenciado por tales interacciones ya que el efecto que producirá una neurona sobre otra será proporcional a la fuerza, peso de la conexión entre ambas.

La señal que envía cada una de las neuronas a sus vecinas dependerá de su propio estado de activación.

Función de salida o transferencia

Asociada a cada unidad (neurona) hay una función de salida, que transforma el estado actual de activación en una señal de salida.

En algunos modelos, esta salida es igual al nivel de activación de la unidad.

Existen cuatro funciones de transferencia típicas que determinan distintos tipos de neuronas:

- Función escalón
- Función lineal y mixta
- Sigmoideal
- Función gaussiana

La función escalón únicamente se utiliza cuando las salidas de la red son binarias. La salida de una neurona se activa sólo cuando el estado de activación es mayor o igual a cierto valor umbral. La función lineal o identidad equivale a no aplicar función de salida. Las funciones mixtas y sigmoideales son las más apropiadas cuando queremos como salida información analógica.

Neurona de función escalón

La función escalón se asocia a neuronas binarias en las cuales cuando la suma de las entradas es mayor o igual que el umbral de la neurona, la activación es 1, si es menor, la activación es 0 (ó -1). Las redes formadas por este tipo de neuronas son fáciles de implementar en hardware, pero sus capacidades están limitadas.

Neurona de función lineal o mixta

La función lineal o mixta corresponde a la función $F(x) = x$. En las neuronas con función mixta si la suma de las señales de entrada es menor que un límite inferior, la activación se define como 0 (ó -1). Si dicha suma es mayor o igual que el límite superior, entonces la activación es 1. Si la suma de entrada está comprendida entre ambos límites, la activación se define como una función lineal de suma de las señales de entrada.

Neurona de función continua (sigmoidal)

Cualquier función definida simplemente en un intervalo de posibles valores de entrada, con un incremento monotónico y que tengan ambos límites superiores e inferiores (por ejemplo las funciones sigmoidal y arco tangente), podrá realizar la función de activación o transferencia de forma satisfactoria.

Con la función sigmoidal, para la mayoría de los valores del estímulo de entrada, el valor dado por la función es cercano a uno de los valores asintóticos. Esto hace posible que en la mayoría de los casos, el valor de salida esté comprendido en la zona alta o baja del sigmoide. De hecho cuando la pendiente es elevada, esta función tiende a la función escalón. La importancia de ésta función es que su derivada es siempre positiva y cercana a cero para los valores grandes positivos o negativos; además toma su valor máximo cuando x es cero. Esto hace que se puedan utilizar las reglas de aprendizaje definidas para la función escalón, con la ventaja respecto a esta función, que la derivada está definida para todo el intervalo. La función escalón no podía definir la derivada en el punto de transición y esto no ayuda a los métodos de aprendizaje en los cuales se usan derivadas.

Función de transferencia gaussiana

Los centros y anchura de estas funciones pueden ser adaptados, lo cual las hace más adaptativas que las funciones sigmoidales.

Regla de aprendizaje

Existen muchas definiciones del concepto aprendizaje, una de ellas podría ser: La modificación del comportamiento inducido por la interacción con el entorno y como resultado de experiencias conducente al establecimiento de nuevos modelos de respuesta a estímulos externos.

Biológicamente, se suele aceptar que la información memorizada en el cerebro está mas relacionada con los valores sinápticos de las conexiones entre las neuronas que con ellas mismas. En el caso de las RNA, se puede considerar que el conocimiento se encuentra representado en los pesos de las conexiones entre neuronas. Todo proceso de aprendizaje implica cierto número de cambios en estas conexiones. Puede decirse que se aprende modificando los valores de los pesos de la red.

Estructura de una Red Neuronal Artificial

Niveles o capas de neuronas

La distribución de neuronas dentro de la red se realiza formando niveles o capas de un número determinado cada una. Se pueden distinguir tres tipos de capas:

De Entrada: es la capa que recibe directamente la información proveniente de las fuentes externas de la red.

Ocultas: son internas a la red, no tiene contacto directo con el exterior. El número de niveles ocultos puede ser de cero a un número elevado. Las neuronas de las capas ocultas pueden estar interconectadas de distintas maneras, lo que determina junto a su número, las distintas topologías.

De Salida: transfieren información de la red hacia el exterior.

Se dice que una red es totalmente conectada si todas las salidas desde un nivel llegan a todos y cada uno de los nodos del mismo nivel siguiente.

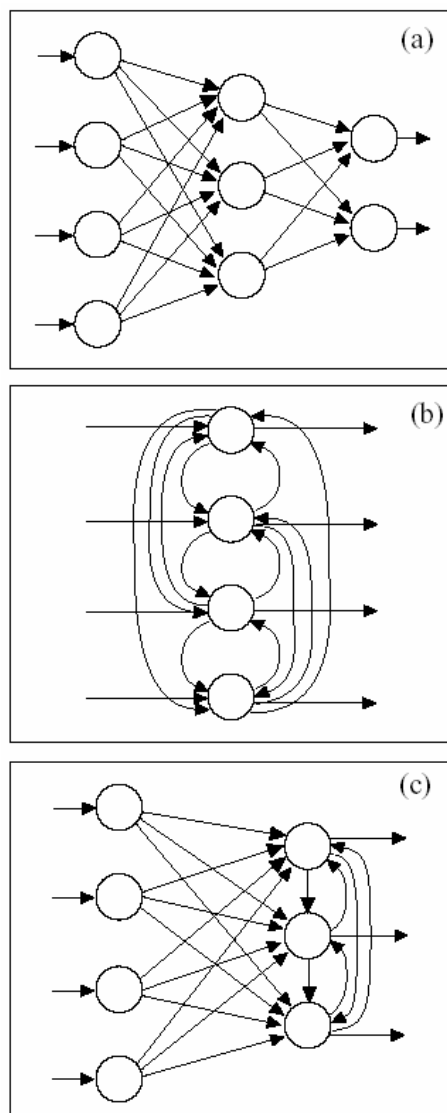
Formas de Conexión entre neuronas

La conectividad entre los nodos de una red neuronal está relacionada con la forma en que las salidas de las neuronas están canalizadas para convertirse en entradas de otras neuronas. La señal de salida de un nodo puede ser una entrada de otro elemento de proceso, o incluso de sí mismo (conexión auto-recurrente).

Cuando ninguna salida de las neuronas es entrada de neuronas del mismo nivel o de niveles precedentes, la red se describe como propagación hacia delante. Cuando las salidas pueden estar conectadas como entradas de neuronas de niveles previos o del mismo nivel, incluyéndose ellas mismas, la red es de propagación hacia atrás. Las redes de propagación hacia atrás que tiene lazos cerrados son sistemas recurrentes.

En la siguiente figura se muestran ejemplos de conexiones.

a) Conexiones hacia delante. b) Conexiones laterales. c) Conexiones hacia atrás (recurrentes).



Topología de las Redes Neuronales

Consiste en la organización de las neuronas en la red formando capas o agrupaciones de neuronas más o menos alejadas de la entrada y salida de la red. Los parámetros fundamentales de la red son: el número de capas, el número de neuronas por capa, el grado de conectividad y el tipo de conexiones entre neuronas.

En términos topológicos podemos clasificar las redes entre: redes de una sola capa y las redes con múltiples capas.

Redes monocapa

En las redes monocapa, como la red HOPFIELD y la red BRAIN-STATE-IN-A-BOX, se establecen conexiones laterales entre las neuronas que pertenecen a la única capa de la red. Además pueden existir conexiones auto recurrente.

Las redes monocapa se utilizan típicamente en tareas relacionadas con lo que se conoce como auto asociación, por ejemplo para regenerar informaciones de entrada que se presentan distorsionadas o incompletas.

Redes Multicapa

Son aquellas que disponen las neuronas agrupadas en varias capas. Una de las formas para distinguir la capa a la que pertenece una neurona, consistiría en fijarse en el origen de las señales que recibe a la entrada y el destino de la señal de salida. Normalmente, todas las neuronas de una capa reciben señales de entrada de otra capa anterior, más cercana a la entrada de la red, y envían su señal de salida a una capa posterior, más cercana a la salida de la red. A estas conexiones se les denominan conexiones hacia delante o feedforward.

Sin embargo en un gran número de estas redes también existe la posibilidad de conectar las salidas de las neuronas de capas posteriores a las entradas de capas anteriores, a estas conexiones se les denomina conexiones hacia atrás o feedback.

Estas dos posibilidades permiten distinguir entre dos tipos de redes: las redes con conexiones hacia adelante (redes feedforward), y las redes que disponen de conexiones tanto hacia delante como hacia atrás (redes feedforward/feedback).

Redes con conexiones hacia delante (feedforward)

Las señales se propagan hacia adelante a través de las capas de la red. No existen conexiones hacia atrás, y normalmente tampoco auto recurrentes, ni laterales, excepto los modelos de red propuestos por Kohonen.

Las redes feedforward más conocidas son: PERCEPTRON, ADALINE, MADALINE, LINEAR ADAPTATIVE MEMORY, DRIVE-REINFORCEMENT, BACKPROPAGATION. Todas ellas son útiles en aplicaciones de reconocimiento o clasificación de patrones.

Redes con conexiones hacia adelante y hacia atrás (feedforward/feedback)

En este tipo de redes circula información tanto hacia delante como hacia atrás durante el funcionamiento de la red. Para que eso sea posible existen conexiones feedforward y feedback entre las neuronas.

En general, suelen ser bicapas, existiendo por lo tanto dos conjuntos de pesos: los correspondientes a las conexiones feedforward de la primera capa (capa de entrada) hacia la segunda (capa de salida) y los de las conexiones feedback de la segunda a la primera. Los valores de los pesos de estos tipos de conexiones no tienen porqué coincidir, siendo diferentes en la mayor parte de los casos.

Este tipo de estructura (bicapa) es particularmente adecuada para realizar una asociación de una información o patrón de entrada (en la primera capa) con otra información o patrón de salida en la segunda capa (lo cual se conoce como heteroasociación), aunque también pueden ser utilizadas para la clasificación de patrones.

Algunas redes tienen un funcionamiento basado en lo que se denomina resonancia, de tal forma que las informaciones en la primera y segundas capas interactúen entre sí hasta que alcanzan un estado estable. Esto permite un mejor acceso a las informaciones almacenadas en la red.

Los dos modelos de red de dos capas más conocidos son la red ART (Adaptive Resonant Theory) y la red BAM (Bidirectional Associative Memory).

También en este grupo de redes existen algunas conexiones laterales entre neuronas de la misma capa.

Mecanismo de Aprendizaje

Es el proceso por el cual una red neuronal modifica sus pesos en respuesta a una información de entrada. Los cambios que se producen durante el proceso de aprendizaje se reducen a destrucción, modificación y creación de conexiones entre las neuronas. La creación de una nueva conexión implica que el peso de la misma pasa a tener un valor distinto de cero; una conexión se destruye cuando su peso pasa a ser cero. El proceso de aprendizaje ha terminado (la red ha aprendido) cuando los valores de los pesos permanecen estables.

Un aspecto importante respecto al aprendizaje es conocer cómo se modifican los valores de los pesos; cuáles son los criterios para cambiar el valor asignado a las conexiones cuando se pretende que la red aprenda una nueva información.

Estos criterios determinan la regla de aprendizaje. Se suelen considerar dos tipos de reglas, las que responden a lo que se conoce como aprendizaje supervisado y aprendizaje no supervisado. Una de las clasificaciones de redes neuronales obedece al tipo de aprendizaje utilizado. Así se pueden distinguir:

- Neuronas con aprendizaje supervisado
- Neuronas con aprendizaje no supervisado

La diferencia fundamental entre ambos tipos es la existencia o no de un agente externo (supervisor) que controle el proceso de aprendizaje.

Otro criterio para diferenciar las reglas de aprendizaje se basan en considerar si la red puede aprender durante su funcionamiento habitual (aprendizaje ON LINE), o si el aprendizaje supone una desconexión de la red; es decir su inhabilitación hasta que el proceso termine (aprendizaje OFF LINE).

Cuando el aprendizaje es OFF LINE se distingue entre una fase de aprendizaje o entrenamiento y una fase de operación o funcionamiento, existiendo un conjunto de datos de entrenamiento y un conjunto de datos de prueba que serán utilizados en la correspondiente fase. En las redes con aprendizaje OFF LINE, los pesos de las

conexiones permanecen fijos después que termina el entrenamiento. Debido a su carácter estático, estos sistemas no presentan problemas de estabilidad en su funcionamiento. En las redes con aprendizaje ON LINE no se distingue entre fase entrenamiento y operación. Los pesos varían siempre que se presenta una nueva información al sistema. Debido a su carácter dinámico, el estudio de la estabilidad es un aspecto fundamental de estudio.

Redes con aprendizaje supervisado

El proceso de aprendizaje se realiza mediante un entrenamiento controlado por un agente externo (supervisor o maestro) que determina la respuesta que debería generar la red a partir de una entrada determinada. El supervisor comprueba la salida de la red y en caso de que ésta no coincida con la deseada, se procederá a modificar los pesos de las conexiones, con el fin de que la salida obtenida se aproxime a la deseada.

Se suelen considerar tres formas de llevar a cabo el aprendizaje:

- Aprendizaje por corrección de error
- Aprendizaje por refuerzo
- Aprendizaje estocástico.

Aprendizaje por corrección de error

Consiste en ajustar los pesos de las conexiones de la red en función de la diferencia entre los valores deseados y los obtenidos en la salida de la red; es decir, en función del error cometido en la salida.

Un algoritmo muy conocido que permite un aprendizaje rápido es propuesto por [Widrow 60], denominado regla delta o regla del mínimo error cuadrado (LMS Error: Least-Mean-Squared Error), que se aplicó en las redes desarrolladas por los mismos, conocidas como ADALINE y MADALINE).

[Widrow 60] definieron una función que permitiría cuantificar el error global cometido en cualquier momento durante el proceso de entrenamiento de la red, lo cual es importante, ya que cuanta más información se tenga sobre el error cometido, más rápido se puede aprender.

Mediante este procedimiento, se llegan a obtener un conjunto de pesos con los que se consigue minimizar el error medio.

Otro algoritmo de aprendizaje por corrección de error lo constituye la denominada regla delta generalizada o algoritmo de retropropagación del error (error backpropagation). Se trata de una generalización de la regla delta para poder aplicarla a redes de conexiones hacia delante (feedforward) con capas o niveles internos ocultos de neuronas que no tienen relación con el exterior.

Estas redes multicapa pueden utilizarse en muchas aplicaciones, pero su proceso de aprendizaje es mucho más lento, debido a que durante el mismo se debe explorar el espacio de posibles formas de utilización de neuronas de las capas ocultas; es decir, establecer cuál es su papel en la red.

Aprendizaje por refuerzo

Es un aprendizaje más lento que el anterior, que se basa en la idea de no disponer de un ejemplo completo del comportamiento deseado; es decir, de no indicar durante el entrenamiento exactamente la salida que se desea que proporcione la red ante una determinada entrada.

En el aprendizaje por refuerzo la función del supervisor se reduce a indicar mediante una señal de refuerzo si la salida obtenida en la red se ajusta a la deseada (éxito = +1 o fracaso = -1), y en función de ello se ajustan los pesos basándose en un mecanismo de probabilidades.

Un ejemplo de algoritmo es el Linear Reward-Penalty o L_{R-P} (algoritmo lineal con recompensa y penalización) presentado por Narendra y Thathacher en 1974. Este algoritmo ha sido ampliado por Barto y Anandan, quienes en 1985 desarrollaron el denominado Associative Reward-Penalty o A_{R-P} (algoritmo asociativo con recompensa y penalización), que se aplica en redes con conexiones hacia delante de dos capas cuyas neuronas de salida presentan una función de activación estocástica.

Otro algoritmo conocido es el Adaptive Heuristic Critic, introducido por Barto, Sutton y Anderson en 1983, que se utiliza en redes feedforward de tres capas especialmente diseñadas para que una parte de la red sea capaz de generar un valor interno de refuerzo que es aplicado a las neuronas de salida de la red.

Aprendizaje estocástico

Este tipo de aprendizaje consiste básicamente en realizar cambios aleatorios en los valores de los pesos de las conexiones de la red y evaluar su efecto a partir del objetivo deseado y de distribuciones de probabilidad.

En el aprendizaje estocástico se suele hacer una analogía en términos termodinámicos, asociando la red neuronal con un sólido físico que tiene cierto estado energético. En el caso de la red, la energía de la misma representaría el grado de estabilidad de la red, de tal forma que el estado de mínima energía correspondería a una situación en la que los pesos de las conexiones consiguen que su funcionamiento sea el que más se ajusta al objetivo deseado.

Según lo anterior, el aprendizaje consistiría en realizar un cambio aleatorio de los valores de los pesos y determinar la energía de la red. Si la energía es menor después del cambio; es decir, si el comportamiento de la red se acerca al deseado, se acepta el cambio; de lo contrario, se aceptaría el cambio en función de una determinada y preestablecida distribución de probabilidades.

Redes con aprendizaje no supervisado

Las redes con dicho aprendizaje no requieren de influencia externa para ajustar los pesos de las conexiones entre sus neuronas. La red no recibe ninguna información por parte del entorno que le indique si la salida generada en respuesta de una entrada es o no correcta. Suele decirse que estas redes son capaces de autoorganizarse.

Estas redes deben encontrar las características, regularidades, correlaciones o categorías que se pueden establecer entre los datos que se presentan en su entrada.

En algunos casos, la salida representa el grado de familiaridad o similitud entre la información que se le está presentando en la entrada y las que se le han mostrado en el pasado. En otro caso podría realizar una clusterización, indicando la red la salida a qué

categoría pertenece la información presentada a la entrada, siendo la propia red quien debe encontrar las categorías apropiadas a partir de correlaciones entre las informaciones presentadas. Una variación de esta categorización es el prototipado. En este caso, la red obtiene prototipos representantes de las clases a las que pertenecen las informaciones de entrada.

También el aprendizaje sin supervisión permite realizar una codificación de los datos de entrada, generando a la salida una versión codificada de la entrada, con menos bits, pero manteniendo la información relevante de los datos.

Algunas redes con aprendizaje no supervisado, realizan un mapeo de características, obteniéndose en las neuronas de salida una disposición geométrica que representa un mapa topográfico de las características de los datos de entrada, de tal forma que si se presentan a la red informaciones similares, siempre sean afectadas neuronas de salidas próximas entre sí, en la misma zona del mapa

Suelen considerarse dos algoritmos de aprendizaje no supervisado:

- Aprendizaje hebbiano
- Aprendizaje competitivo y cooperativo.

Aprendizaje hebbiano

Este tipo de aprendizaje se basa en el postulado formulado por Donald O. Hebb en 1949: “Cuando un axón de una celda A está suficientemente cerca como para conseguir excitar a una celda B y repetida o persistentemente toma parte en su activación, algún proceso de crecimiento o cambio metabólico tiene lugares en una o ambas celdas, de tal forma que la eficiencia de A, cuando la celda a activar es B, aumenta. Por celdas, Hebb entiende un conjunto de neuronas fuertemente conexionadas a través de una estructura compleja. La eficiencia podría identificarse por la intensidad o magnitud de la conexión, es decir, con el peso.

Se puede decir que el aprendizaje consiste en el ajuste de los pesos de las conexiones de acuerdo con la correlación (multiplicación en el caso de valores binarios +1 y -1) de los valores de activación (salidas) de las neuronas conectadas.

Si las dos unidades son activas (positivas), se refuerza la conexión; por el contrario, cuando una es activa y la otra pasiva, se debilita la conexión.

Aprendizaje competitivo y cooperativo

En dicho aprendizaje suele decirse que las neuronas compiten (y cooperan) unas con otras con el fin de llevar a cabo una tarea dada.

La competición entre neuronas se realiza en todas las capas de la red, existiendo en estas neuronas conexiones recurrentes de autoexcitación y conexiones de inhibición por parte de neuronas vecinas. Si el aprendizaje es cooperativo, estas conexiones con las vecinas serán de excitación.

El objetivo de este aprendizaje es clusterizar los datos que se introducen en la red. De esta forma, las informaciones similares son clasificadas formando parte de la misma categoría, y por tanto deben activar la misma neurona de salida. Las categorías deben ser creadas por la misma red, puesto que se trata de aprendizaje no supervisado, a través de las correlaciones entre los datos.

En este tipo de redes, cada neurona tiene asignado un peso total, suma de todos los pesos de las conexiones que tiene a su entrada. El aprendizaje afecta sólo a las neuronas

ganadoras (activas), redistribuyendo este peso total entre todas las conexiones que llegan a la neurona vencedora y repartiendo esta cantidad por igual entre todas las conexiones procedentes de unidades activas.

Existe otro caso particular de aprendizaje competitivo, denominado teoría de la resonancia adaptativa, desarrollado por Carpenter y Grossberg en 1986 y utilizado en la red feedforward /feedback de dos capas conocida como ART. Esta red realiza un prototipado de las informaciones que recibe a la entrada, generando como salida un ejemplar o prototipo que representa a todas las informaciones que podrían considerarse pertenecientes a la misma categoría.

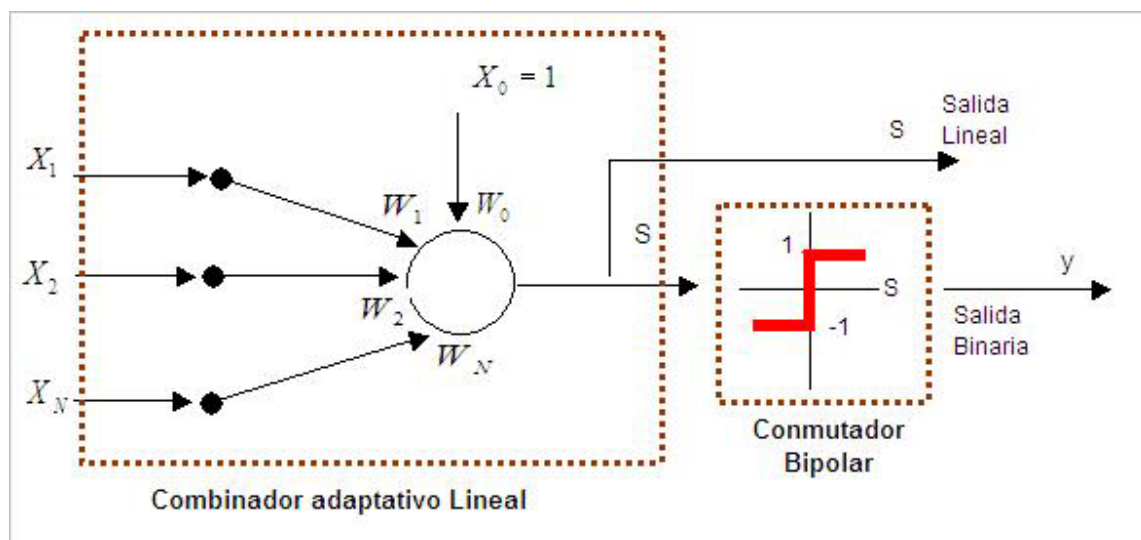
Red ADALINE

Las redes ADALINE (Adaptive Linear Element), fueron desarrolladas por Bernie Widrow en la Universidad de Stanford. Dicha red usa neuronas con función de transferencia escalón, y está limitada a una única neurona de salida.

Utiliza la denominada regla Delta de Widrow-Hoff o regla del mínimo error cuadrado medio (LMS). Estas redes pueden procesar información analógica, tanto de entrada como de salida, utilizando una función de activación lineal o sigmoideal.

En cuanto a su estructura, está formada por un elemento denominado combinador adaptativo lineal (ALC) que obtiene una salida lineal(s) que pueda ser aplicada a otro elemento de conmutación bipolar, de forma que si la salida del ALC es positiva, la salida de la red ADALINE es +1; si la salida es negativa, entonces la salida de la red ADALINE es -1.

En la figura siguiente se muestra la red ADALINE, compuesta por un combinador adaptativo lineal y una función de salida bipolar.



La red ADALINE se puede utilizar para generar una salida analógica utilizando un conmutador sigmoideal, en lugar de binario; en tal caso, la salida se obtendrá aplicando una función tipo sigmoideal, como la tangente hiperbólica ($\tanh(s)$) o la exponencial ($1/(1+e^{-s})$).

Aprendizaje de la red ADALINE

La red ADALINE utiliza un aprendizaje OFF LINE con supervisión denominado LMS (Least Mean Squared) o regla del mínimo cuadrado medio.

Cuando se utiliza una red ADALINE para resolver un problema concreto, es necesario determinar una serie de aspectos prácticos, como el número de vectores de entrenamiento necesarios, hallar la forma de generar la salida deseada para cada vector de entrenamiento, o la dimensión óptima del vector de pesos, o cuales deberían ser los valores iniciales de los pesos, así como si es necesario o no un umbral θ , o cuál debe ser el valor de α , o cuándo se debe finalizar el entrenamiento, etc.

Respecto al número de componentes del vector de pesos, si el número de entradas está bien definido, entonces habrá un peso por cada entrada, con la opción de añadir o no un peso para la entrada del umbral.

La solución es diferente cuando sólo se dispone de una señal de entrada. En estos casos, la aplicación más común es el filtro adaptativo para, por ejemplo, eliminar ruido de la señal de entrada.

Aplicaciones de la red ADALINE

La principal aplicación de las redes ADALINE está en el campo del procesamiento de señales, en concreto, para el diseño y realización de filtros que eliminen el ruido en señales portadoras de información.

Se destaca su uso como filtros de ecuación adaptativos en módems de alta velocidad y canceladores adaptativos del eco para el filtrado de señales en comunicaciones telefónicas de larga distancia y comunicaciones vía satélite.

También se han utilizado para la eliminación del ruido materno de las grabaciones electrocardiográficas (ECG) del latido del corazón del feto humano.

Por otro lado los filtros adaptativos también se suelen utilizar para predecir el valor futuro de una señal a partir de su valor actual, basándose en un aprendizaje en el que se emplea como entrada el valor retardado de la señal actual (la señal en algún momento anterior), y como salida esperada, el valor actual de la señal. El filtro intentará minimizar el error entre su salida y la señal actual, basándose en una entrada que es el valor de la señal en algún instante anterior. Una vez que el filtro predice correctamente la señal actual, basándose en la señal anterior, se puede utilizar directamente la actual como entrada sin retardo. El filtro realizará una predicción del valor futuro de la señal.

La Red Backpropagation

En 1986, Rumelhart, Hinton y Williams, formalizaron un método para que una red neuronal aprendiera la asociación que existe entre los patrones de entrada y las clases correspondientes, utilizando varios niveles de neuronas.

El método backpropagation (propagación del error hacia atrás), basado en la generalización de la regla delta, ha ampliado de forma considerable el rango de aplicaciones de las redes neuronales.

El funcionamiento de la red backpropagation (BPN) consiste en el aprendizaje de un conjunto predefinido de pares de entradas-salidas dados como ejemplo: primero se aplica un patrón de entrada como estímulo para la primera capa de las neuronas de la red, se va propagando a través de todas las capas superiores hasta generar una salida, se compara el resultado en las neuronas de salida con la salida que se desea obtener y se calcula un

valor de error para cada neurona de salida. A continuación, estos errores se transmiten hacia atrás, partiendo de la capa de salida hacia todas las neuronas de la capa intermedia que contribuyan directamente a la salida, recibiendo de error aproximado a la neurona intermedia a la salida original. Este proceso se repite, capa por capa, hasta que todas las neuronas de la red hayan recibido un error que describa su aportación relativa al error total. Basándose en el valor del error recibido, se reajustan los pesos de conexión de cada neurona, de manera que en la siguiente vez que se presente el mismo patrón, la más salida esté cercana a la deseada.

La importancia de la red backpropagation consiste en su capacidad de autoadaptar los pesos de las neuronas de las capas intermedias para aprender la relación que existe entre un conjunto de patrones de entrada y sus salidas correspondientes. Es importante la capacidad de generalización, facilidad de dar salidas satisfactorias a entradas que el sistema no ha visto nunca en su fase de entrenamiento. La red debe encontrar una representación interna que le permita generar las salidas deseadas cuando se le dan entradas de entrenamiento, y que pueda aplicar, además, a entradas no presentadas durante la etapa de aprendizaje para clasificarlas.

Estructura y aprendizaje de la red backpropagation

En una red Backpropagation existe una capa de entrada con n neuronas y una capa de salida con m neuronas y al menos una capa oculta de neuronas internas. Cada neurona de una capa (excepto las de entrada) recibe entradas de todas las neuronas de la capa anterior y envía su salida a todas las neuronas de la capa posterior (excepto las de salida). No hay conexiones hacia atrás feedback ni laterales entre las neuronas de la misma capa.

La aplicación del algoritmo tiene dos fases, una hacia delante y otra hacia atrás. Durante la primera fase el patrón de entrada es presentado a la red y propagado a través de las capas hasta llegar a la capa de salida. Obtenidos los valores de salida de la red, se inicia la segunda fase, comparándose éstos valores con la salida esperada para obtener el error. Se ajustan los pesos de la última capa proporcionalmente al error. Se pasa a la capa anterior con una retropropagación del error, ajustando los pesos y continuando con este proceso hasta llegar a la primera capa. De esta manera se han modificado los pesos de las conexiones de la red para cada patrón de aprendizaje del problema, del que conocíamos su valor de entrada y la salida deseada que debería generar la red ante dicho patrón.

La técnica Backpropagation requiere el uso de neuronas cuya función de activación sea continua, y por lo tanto, diferenciable. Generalmente, la función utilizada será del tipo sigmoideal.

Pasos para aplicar el algoritmo de entrenamiento

Paso 1

Inicializar los pesos de la red con valores pequeños aleatorios.

Paso 2

Presentar un patrón de entrada y especificar la salida deseada que debe generar la red.

Paso 3

Calcular la salida actual de la red. Para ello presentamos las entradas a la red y vamos calculando la salida que presenta cada capa hasta llegar a la capa de salida, ésta será la salida de la red.

Paso 4

Calcular los términos de error para todas las neuronas.

Paso 5

Actualización de los pesos: para ello utilizamos un algoritmo recursivo, comenzando por las neuronas de salida y trabajando hacia atrás hasta llegar a la capa de entrada.

Paso 6

El proceso se repite hasta que el término de error resulta aceptablemente pequeño para cada uno de los patrones aprendidos.

Consideraciones sobre el algoritmo de aprendizaje

El algoritmo encuentra un valor mínimo de error (local o global) mediante una aplicación de pasos (gradiente) descendentes. Cada punto de la superficie de la función corresponde a un conjunto de valores de los pesos de la red. Con el gradiente descendente, siempre que se realiza un cambio en todos los pesos de la red, se asegura el descenso por la superficie del error hasta encontrar el valle más cercano, lo que puede hacer que el proceso de aprendizaje se detenga en un mínimo local de error.

Control de Convergencia

En las técnicas de gradiente decreciente es conveniente avanzar por la superficie del error con incrementos de pesos pequeños. Esto se debe a que tenemos una información local de la superficie y no se sabe lo lejos o cerca que se está del punto mínimo. Con incrementos grandes se corre el riesgo de pasar por encima del punto mínimo sin conseguir estacionarse en él. Con incrementos pequeños, aunque se tarda más en llegar, se evita que ocurra esto.

Dimensionamiento de la red. Número de neuronas ocultas

No se pueden dar reglas concretas para determinar el número de neuronas o número de capas de una red para resolver un problema concreto.

Respecto al número de capas de la red, en general tres capas son suficientes (entrada - oculta-salida). Sin embargo, hay veces que un problema es más fácil de resolver con más de una capa oculta. El tamaño de las capas, tanto de entrada como de salida, suelen venir determinado por la naturaleza de la aplicación. En cambio, decidir cuántas neuronas debe tener una capa oculta no suele ser tan evidente.

El número de neuronas ocultas interviene en la eficiencia de aprendizaje y de generalización de la red. No hay ninguna regla que indique el número óptimo, en cada problema se debe ensayar.

Aplicaciones de la red Backpropagation

Actualmente, este tipo de redes se están aplicando en distintas clases de problema debido a la naturaleza general del proceso de aprendizaje. Algunos de los campos generales de aplicación son:

Codificación de Información.

Traducción de texto a lenguaje hablado.

Reconocimiento de lenguaje hablado.

Reconocimiento óptico de caracteres (OCR).
Software de Redes Neuronales

CARACTERISTICAS DE LAS REDES NEURONALES

Aprendizaje inductivo: No se le indican las reglas para dar una solución, sino que extrae sus propias reglas a partir de los ejemplos de aprendizaje, modifican su comportamiento en función de la experiencia. Esas reglas quedan almacenadas en las conexiones y no representadas explícitamente como en los sistemas basados en conocimiento (simbólico-deductivos)

Generalización: Una vez entrenada, se le pueden presentar a la red datos distintos a los usados durante el aprendizaje. La respuesta obtenida dependerá del parecido de los datos con los ejemplos de entrenamiento.

Abstracción o tolerancia al ruido: Las redes neuronales artificiales son capaces de extraer o abstraer las características esenciales de las entradas aprendidas, de esta manera pueden procesar correctamente datos incompletos o distorsionados.

Procesamiento paralelo: las neuronas reales trabajan en paralelo; en el caso de las redes artificiales es obvio que si usamos un solo procesador no podrá haber proceso paralelo real; sin embargo hay un paralelismo inherente, lo esencial es que la estructura y modo de operación de las redes neuronales las hace especialmente adecuadas para el procesamiento paralelo real mediante multiprocesadores (se están desarrollando máquinas específicas para la computación neuronal).

Memoria distribuida: el conocimiento acumulado por la red se halla distribuido en numerosas conexiones, esto tiene como consecuencia la tolerancia a fallos: una red neuronal es capaz de seguir funcionando adecuadamente a pesar de sufrir lesiones con destrucción de neuronas o sus conexiones, ya que la información se halla distribuida por toda la red, sin embargo en un programa tradicional un pequeño fallo en cualquier punto puede invalidarlo todo y dar un resultado absurdo o no dar ningún resultado.

VENTAJAS DE LAS REDES NEURONALES

Debido a su constitución y a sus fundamentos, las RNA presentan un gran número de características semejantes a las del cerebro. Por ejemplo, son capaces de aprender de la experiencia, de generalizar de casos anteriores a nuevos casos, de abstraer características esenciales a partir de entradas que representan información irrelevante, etc. Esto hace que ofrezcan numerosas ventajas y que este tipo de tecnología se esté aplicando en múltiples áreas.

Estas ventajas incluyen:

Aprendizaje Adaptativo: Es una de las características más atractivas de las redes neuronales, es la capacidad de aprender a realizar tareas basadas en un entrenamiento o una experiencia inicial.

En el proceso de aprendizaje, los enlaces ponderados de las neuronas se ajustan de manera que se obtengan unos resultados específicos. Una RNA no necesita un algoritmo para resolver un problema, ya que ella puede generar su propia distribución de los pesos

de los enlaces mediante el aprendizaje. También existen redes que continúan aprendiendo a lo largo de su vida, después de completado e periodo inicial de entrenamiento.

La función del diseñador es únicamente la obtención de la arquitectura apropiada. No es problema del diseñador el cómo la red aprenderá a discriminar; sin embargo, si es necesario que desarrolle un buen algoritmo de aprendizaje que proporcione la capacidad de discriminar de la red mediante un entrenamiento con patrones.

Autoorganización: Las redes neuronales usan su capacidad de aprendizaje adaptativo para organizar la información que reciben durante el aprendizaje y/o la operación. Una RNA puede crear su propia organización o representación de la información que recibe mediante una etapa de aprendizaje. Esta autoorganización provoca la facultad de las redes neuronales de responder apropiadamente cuando se les presentan datos o situaciones a las que no habían sido expuestas anteriormente.

Tolerancia a Fallos: Comparados con los sistemas computacionales tradicionales, los cuales pierden su funcionalidad en cuanto sufren un pequeño error de memoria, en las redes neuronales, si se produce un fallo en un pequeño número de neuronas, aunque el comportamiento del sistema se ve influenciado, sin embargo no sufre una caída repentina. Hay dos aspectos distintos respecto a la tolerancia a fallos: primero, las redes pueden aprender a reconocer patrones con ruido, distorsionados, o incompleta. Segundo pueden seguir realizando su función (con cierta degradación) aunque se destruya parte de la red. La razón por la que las redes neuronales son tolerantes a fallos es que tienen su información distribuida en las conexiones entre neuronas, existiendo cierto grado de redundancia en ese tipo de almacenamiento, a diferencia de la mayoría de los ordenadores algorítmicos y sistemas de recuperación de datos que almacenan cada pieza de información en un estado único, localizado y direccionable.

Operación en Tiempo Real: Los computadores neuronales pueden ser realizados en paralelo, y se diseñan y fabrican máquinas con hardware especial para obtener esta capacidad.

Fácil inserción dentro de la tecnología existente: Debido a que una red puede ser rápidamente entrenada, comprobada, verificada y trasladada una implementación hardware de bajo costo, es fácil insertar RNA para aplicaciones específicas dentro de sistemas existentes (chips, por ejemplo). De esta manera, las redes neuronales se pueden utilizar para mejorar sistemas de forma incremental, y cada paso puede ser evaluado antes de acometer un desarrollo más amplio.

APLICACIONES DE LAS REDES NEURONALES

Las redes neuronales son una tecnología computacional emergente que puede utilizarse en un gran número y variedad de aplicaciones, tanto como comerciales como militares.

Hay muchos tipos diferentes de redes neuronales, cada uno de los cuales tiene una aplicación particular más apropiada. Separándolas según las distintas disciplinas algunos ejemplos de sus aplicaciones son:

Biología:

- Aprender más acerca del cerebro y otros sistemas.
- Obtención de modelos de la retina.

Empresa

- Reconocimiento de caracteres escritos.
- Identificación de candidatos para posiciones específicas.
- Optimización de plazas y horarios en líneas de vuelo.
- Explotación de bases de datos.
- Evaluación de probabilidad de formaciones geológicas y petrolíferas.
- Síntesis de voz desde texto.

Medio Ambiente:

- Analizar tendencias y patrones.
- Previsión del tiempo.

Finanzas:

- Previsión de la evolución de los precios.
- Valoración del riesgo de los créditos.
- Identificación de falsificaciones.
- Interpretación de firmas.

Manufacturación:

- Robots automatizados y sistemas de control (visión artificial y sensores de presión, temperatura, gas, etc.)
- Control de producción en líneas de proceso.
- Inspección de calidad.
- Filtrado de señales.

Medicina:

- Analizadores del habla para la ayuda de audición de sordos profundos.
- Diagnóstico y tratamiento a partir de síntomas y/o de datos analíticos (encefalograma, etc.).

Monitorización en cirugía:

- Predicción de reacciones adversas a los medicamentos.
- Lectoras de Rayos X.
- Entendimiento de causa de ataques epilépticos.

Militares:

- Clasificación de las señales de radar.
- Creación de armas inteligentes.
- Optimización del uso de recursos escasos.

REDES NEURONALES Y CONTROL

Lo que se hace en control es modelar, según los parámetros aprendidos en sistemas dinámicos, los sistemas para luego controlarlos, sin embargo en ese modelamiento se desprecian muchos datos debido a la linealidad de los mismos, por ejemplo, al modelar un motor se desprecian datos como el desgaste de máquina, esos son valores importantes, pero al tenerlos en cuenta la solución de un sistema se haría imposible, así que se hace necesario despreciar esos términos, sin ellos el modelamiento funciona pero en la vida práctica no es tan preciso. Ese problema se soluciona con redes neuronales, debido a las teorías anteriormente expuestas, si por ejemplo, usted modela un sistema de manera tradicional y luego este sufre variación alguna los planteamientos iniciales ya no funcionan, con las redes neuronales eso ya no sucede, porque el sistema después de haber

recibido unos patrones iniciales comienza a identificar, acepta, aprende y responde ante diferentes señales. Sin importar que estas no sean idénticas a los patrones iniciales.

FUTURO

Es necesario resaltar la significación e importancia que las redes neuronales están adquiriendo en la actualidad como lo evidencia el hecho de formar parte de los estudios centrales de instituciones gubernamentales a nivel mundial.

Así que la intención principal es profundizar en esta nueva tecnología, aprovechando el hecho de que será está una materia en cursos posteriores en el transcurso de mi carrera, para así lograr un alto grado de aprendizaje y con el la implementación practica de alguna red neuronal.

BIBLIOGRAFIA

<http://www.monografias.com/trabajos12/redneuro/redneuro.shtml>
<http://www.gc.ssr.upm.es/inves/neural/ann2/anntutor.htm>
<http://www.monografias.com/trabajos/redesneuro/redesneuro.shtml>
<http://www.iiia.csic.es/~mario/rna/intro-rna.html>

Artificial Neural Networks: A Tutorial

Anil K. Jain

Michigan State University

Jianchang MAO

K.M. Mohiuddin

IBM Almaden Research Center