

Redes Neuronales

Las redes neuronales son más que otra forma de emular ciertas características propias de los humanos, como la capacidad de memorizar y de asociar hechos. Si se examinan con atención aquellos problemas que no pueden expresarse a través de un algoritmo, se observará que todos ellos tienen una característica en común: la experiencia. El hombre es capaz de resolver estas situaciones acudiendo a la experiencia acumulada. Así, parece claro que una forma de aproximarse al problema consista en la construcción de sistemas que sean capaces de reproducir esta característica humana.

En definitiva, las redes neuronales no son más que un modelo artificial y simplificado del cerebro humano, que es el ejemplo más perfecto del que disponemos para un sistema que es capaz de adquirir conocimiento a través de la experiencia. Una red neuronal es **"un nuevo sistema para el tratamiento de la información, cuya unidad básica de procesamiento está inspirada en la célula fundamental del sistema nervioso humano: la neurona"**.

Todos los procesos del cuerpo humano se relacionan en alguna u otra forma con la (in)actividad de estas neuronas. Las mismas son un componente relativamente simple del ser humano, pero cuando millares de ellas se conectan en forma conjunta se hacen muy poderosas.

También, es bien conocido que los humanos son capaces de aprender. Aprendizaje significa que aquellos problemas que inicialmente no pueden resolverse, pueden ser resueltos después de obtener más información acerca del problema.

Por lo tanto, las Redes Neuronales consisten de unidades de procesamiento que intercambian datos o información. Se utilizan para reconocer patrones, incluyendo imágenes, manuscritos y secuencias de tiempo, tendencias financieras.

Existen numerosas formas de definir a las redes neuronales; desde las definiciones cortas y genéricas hasta las que intentan explicar más detalladamente qué son las redes neuronales. Por ejemplo:

1) Un sistema de computación compuesto por un gran número de elementos simples, elementos de procesos muy interconectados, los cuales procesan información por medio de su estado dinámico como respuesta a entradas externas.

2) 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.

Ventajas que ofrecen las redes neuronales

Debido a su constitución y a sus fundamentos, las redes neuronales artificiales 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. Entre las ventajas se incluyen:

Aprendizaje Adaptativo: capacidad de aprender a realizar tareas basadas en un entrenamiento o en una experiencia inicial.

Auto-organización: una red neuronal puede crear su propia organización o representación de la información que recibe mediante una etapa de aprendizaje.

Tolerancia a fallos: la destrucción parcial de una red conduce a una degradación de su estructura sin embargo, algunas capacidades de la red se pueden retener, incluso con un gran daño.

Operación en tiempo real: los cálculos neuronales pueden ser realizados en paralelo; para esto se diseñan y fabrican máquinas con hardware especial para obtener esta capacidad.

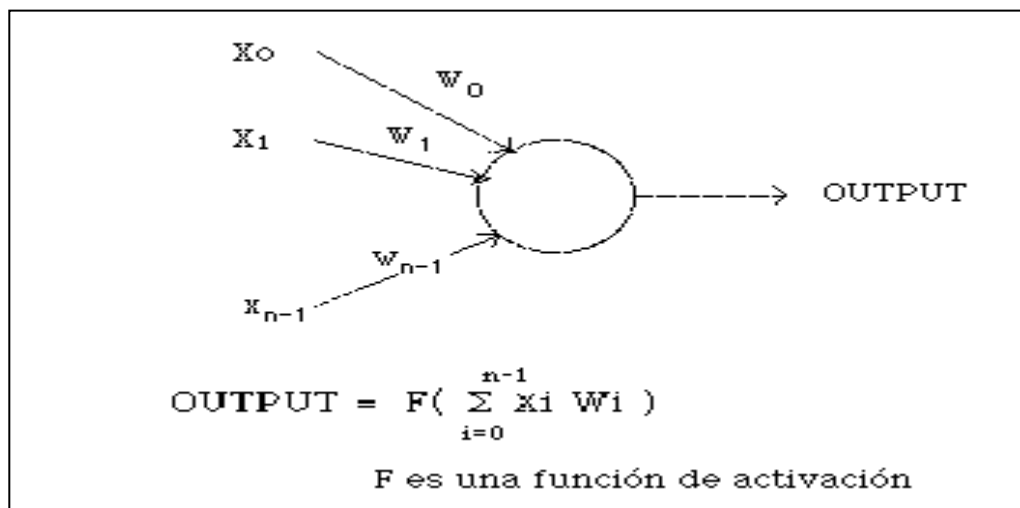
Fácil inserción dentro de la tecnología existente: se pueden obtener chips especializados para redes neuronales que mejoran su capacidad en ciertas tareas. Ello facilitará la integración modular en los sistemas existentes.

Algunas aplicaciones de redes neuronales artificiales.

- Reconocimiento de caracteres impresos
- Reconocimiento de caracteres manuscritos.
- Memorias asociativas.
- Reconocimiento de voz
- Control de robots
- Toma de decisiones
- Reconocimiento de enfermedades cardiacas
- Reconocimiento de señales de radio
- Generación de reglas para sistemas expertos

El Neuron Artificial

Es un elemento procesador al que se aplican un conjunto de entradas, cada una representando la salida de otro neurón. Estas entradas que podemos llamar X , se multiplican por un peso (número real) asociado a ellas, que podemos llamar W



$$\text{NET} = \sum X_i W_i \text{ con } i=0 \text{ a } n-1$$

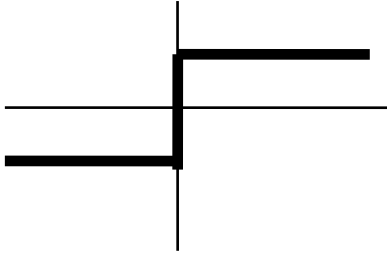
La suma ponderada se "filtra" por medio de una función, llamada función de activación:

$$\text{OUT} = F(\text{NET})$$

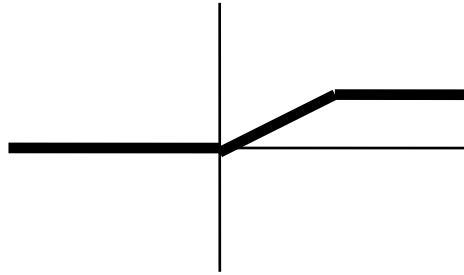
Se dice que un neurón "enciende" si su valor de salida es mayor o igual a un valor umbral, el cual es determinado por los posibles valores de la función de activación.

Las funciones de activación pueden ser las siguientes:

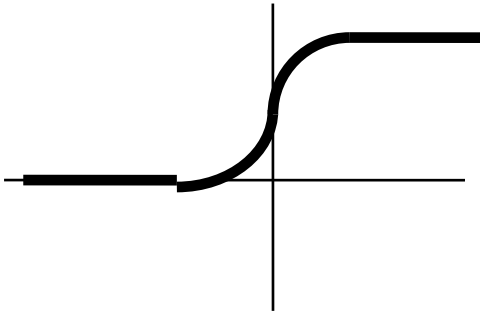
a) escalón



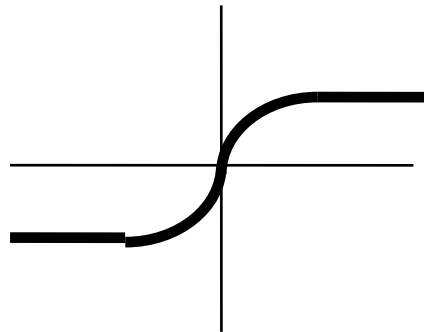
b) lineal con umbral



c) sigmoide



d) tangente hiperbólica



Características de las redes neuronales

Existen cuatro aspectos que caracterizan una red neuronal: su topología, el mecanismo de aprendizaje, tipo de asociación entre la información de entrada y de salida, y la forma de representación de estas informaciones.

Topología : Monocapa y Multicapa

Consiste en la organización y disposición de las neuronas en la red formando capas o agrupaciones de neuronas. Los parámetros fundamentales de la red son: número de capas, número de neuronas por capa, grado de conectividad y tipo de conexión entre neuronas.

Al hacer una clasificación topológica de las RNAs se suelen distinguir:

Redes monocapa : se establecen conexiones laterales entre las neuronas que pertenecen a la única capa que constituye la red. Ejemplos red de HOPFIELD. Las redes monocapa se utilizan típicamente en tareas relacionadas con lo que se conoce como autoasociación; por ejemplo, para regenerar informaciones de entrada que se presenta como incompleta o distorsionada.

Redes multicapa : disponen las neuronas agrupadas en varios niveles. Dado que este tipo de redes disponen de varias capas, las conexiones entre neuronas pueden ser del tipo feedforward (conexión hacia adelante) o del tipo feedback (conexión hacia atrás).

Que es el aprendizaje? :

El 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 la etapa de aprendizaje se reducen a la destrucción (el peso de la conexión toma el valor 0), modificación y creación (el peso de la conexión toma un valor distinto de 0) de conexiones entre las neuronas.

Podemos considerar que el proceso de aprendizaje ha terminado:

- a) Cuando los valores de los pesos permanecen estables
- b) Mediante un número fijo de ciclos
- c) Cuando el error es menor que el error inicialmente establecido

Un aspecto importante es determinar los criterios de la regla de aprendizaje; cómo se van a modificar los pesos.

Mecanismos de aprendizaje Aprendizaje supervisado Aprendizaje NO supervisado

La diferencia entre ambos tipos estriba en la existencia o no de una agente externo que controle todo el proceso.

Otro criterio para diferenciar las reglas de aprendizaje se basa en considerar si la red puede aprender durante su funcionamiento (aprendizaje ON LINE) o requiere de una fase previa de entrenamiento (aprendizaje OFF LINE). En este último debe existir un conjunto de datos de entrenamiento y un conjunto de datos de test o prueba; igualmente los pesos de las conexiones no se modifican después de terminar la etapa de entrenamiento de la red. En la red ON LINE los pesos varían dinámicamente cada vez que se presente una nueva información al sistema.

Redes con aprendizaje supervisado:

Se caracteriza porque el proceso de aprendizaje se realiza mediante un entrenamiento controlado por un agente externo (supervisor, maestro) que determina la respuesta que debería generar la red a partir de una entrada determinada. El supervisor comprueba la salida generada por el sistema y en el caso de que no coincida con la esperada, se procederá a modificar los pesos de las conexiones. En este tipo de aprendizaje se suelen distinguir a su vez tres formas de llevarlo a cabo:

- **Por corrección de error**
- **Por refuerzo**
- **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.

Algoritmos que utilizan este tipo de aprendizaje son:

Regla de aprendizaje del perceptron: utilizada en la red PERCEPTRON

Regla delta generalizada: utilizada en redes multicapa

Aprendizaje por refuerzo :

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

Aprendizaje estocástico :

Consiste básicamente en realizar cambios aleatorios en los valores de los pesos y evaluar su efecto a partir del objetivo deseado y de distribuciones de probabilidad. Un red que utiliza este tipo de aprendizaje es la red Boltzman Machine, ideada por Hinton, Ackley y Sejnowski en 1984 y la red Cauchy Machine desarrollada por Szu en 1986.

Redes con aprendizaje NO supervisado

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 a una determinada entrada es o no correcta; son capaces de autoorganizarse.

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

Pero, ¿qué genera la red en la salida?. Existen varias posibilidades en cuanto a interpretación:

- La salida representa el grado de familiaridad o similitud entre la información de entrada y las informaciones mostradas con anterioridad.
- Clusterización o establecimiento de categorías, indicando la red a la salida a qué categoría pertenece la información de entrada, siendo la propia red la que debe establecer las correlaciones oportunas.
- Codificación de los datos de entrada, generando a la salida una versión codificada con menos bits, pero manteniendo la información relevante de los datos.
- Mapeo de características, obteniéndose una disposición geométrica que representa un mapa topográfico de las características de los datos de entrada.

Tipo de asociación entre las informaciones de entrada y salida

Las redes neuronales son sistemas que almacenan cierta información aprendida; esta se registra de forma distribuida en los pesos asociados a las conexiones entre neuronas. Hay que establecer cierta relación o asociación entre la información presentada a la red y la salida ofrecida por esta. Es lo que se conoce como memoria asociativa.

Existen dos formas primarias de realizar esta asociación entrada/salida y que generan dos tipos de redes:

Redes heteroasociativas :

La red aprende parejas de datos $[(A_1, B_1), (A_2, B_2), \dots, (A_n, B_n)]$, de tal forma que cuando se le presente determinada información de entrada A_i responda con la salida correspondiente B_i . Al asociar informaciones de entrada con diferentes informaciones de salida, precisan al menos de 2 capas, una para captar y retener la información de entrada y otra para mantener la salida con la información asociada. Si esto no fuese así se perdería la información inicial al obtenerse la salida asociada; es necesario mantener la información de entrada puesto que puede ser necesario acceder varias veces a ella, por lo que debe permanecer en la capa de entrada. El aprendizaje de este tipo de redes suele ser **supervisado**.

Redes autoasociativas :

La red aprende ciertas informaciones A_1, A_2, \dots, A_n de forma que cuando se le presenta una información de entrada realizará una autocorrelación, respondiendo con uno de los datos almacenados, el más parecido al de entrada. Este tipo de redes pueden implementarse con una sola capa de neuronas. El tipo de aprendizaje utilizado habitualmente es el **no supervisado** y suelen utilizarse en tareas de filtrado de información para la reconstrucción de datos, eliminando distorsiones o ruido, explorar relaciones entre informaciones similares para facilitar la búsqueda por contenido en bases de datos y para resolver problemas de optimización

Representación de la información de entrada y salida

Redes continuas :

En un gran número de redes, tanto los datos de entrada como de salida son de naturaleza analógica (valores reales continuos y normalmente normalizados, por lo que su valor absoluto será menor que la unidad). En este caso las funciones de activación de las neuronas serán también continuas, del tipo lineal o sigmoideal.

Redes discretas :

Por el contrario, otras redes sólo admiten valores discretos $[0,1]$ a la entrada, generando también en la salida respuestas de tipo binario. La función de activación en este caso es del tipo escalón.

Redes híbridas :

La información de entrada es continua pero a la salida ofrecen información binaria.