

1 Introducción a las Redes Neuronales Artificiales.

Contenido

1	Introducción a las Redes Neuronales Artificiales.....	0
1.1	Introducción.....	1
1.1.1	Inteligencia, Inteligencia Artificial e Inteligencia Computacional.....	1
1.1.2	¿Qué es computación neuronal?.....	5
1.1.3	Breve Historia de las Redes Neuronales Artificiales.....	7
1.2	Fundamentos de Las RNA.....	9
1.2.1	Estructura de la Neurona.....	10
1.2.2	Elementos de la Red Neuronal Artificial.....	12
1.3	Características de una red neuronal artificial.....	15
1.3.1	Topología de las RNA.....	16
1.3.2	Mecanismo de Aprendizaje.....	18
1.3.3	Tipo de Asociación entre la Información de Entrada y de Salida.....	22
1.3.4	Representación de la Información de Entrada y Salida.....	24
1.4	Áreas de Aplicación de las RNA.....	24
1.5	Algunas Preguntas.....	27
1.6	Referencias Bibliográficas.....	29

1.1 Introducción.

Las redes neuronales artificiales (RNA), son una aproximación a la inteligencia artificial fundamentada en el cerebro de los mamíferos, más explícitamente en las neuronas biológicas. En la actualidad son ampliamente usadas en diferentes campos del conocimiento para: análisis de datos, reconocimiento de patrones y problemas de predicción. (Werbos, 2005) Define tres enfoques de las RNA, el más elemental presenta a las RNA como una caja de herramientas para la solución de un tipo particular de problemas; una segunda visión indica que las redes neuronales se podrán dar cuando se disponga de la sexta generación de computadores¹, sistemas a los que se implementaría el aprendizaje. Y la tercera que define a las RNA como sistemas matemáticos bien especificados, diseñados para capturar la clase más alta de inteligencia que se encuentra en el cerebro de los mamíferos, es decir capturar la capacidad funcional del sistema. Por lo tanto la visión general de las redes neuronales es muy amplia y muy ambiciosa, en la actualidad estamos en el nivel de herramientas y en este sentido serán revisadas en el presente curso.

1.1.1 Inteligencia, Inteligencia Artificial e Inteligencia Computacional.

Definir la inteligencia, no es una tarea sencilla, puesto que existen muchas sutilezas considerar. En términos generales, la inteligencia está relacionada a las habilidades mentales para la resolución de problemas. Habilidades como: recordar, usar el conocimiento aprendido para responder a situaciones nuevas en forma correcta, adaptarse a esas situaciones y como resultado manipular el ambiente. Aspectos como la experiencia son importantes, ya que esta podría conducir a respuestas en forma más certera. Una característica que siempre es tomada en cuenta en los test de inteligencia es la velocidad con la que se responde. No hay duda de que un elemento clave de la inteligencia es la habilidad para aprender.

Si “adaptamos” este concepto a las computadoras, podríamos decir de manera general que la inteligencia artificial (IA) es la ciencia de imitar las facultades mentales humanas en un computador.

Esta definición, aunque dice mucho, puede dejar una vaga impresión del término de Inteligencia Artificial, y conducir a falsas interpretaciones. (Hopgood, 2003), explica la definición de IA en el contexto del espectro del comportamiento inteligente desde el nivel

¹ Las PC actuales ya constan de algunos núcleos por procesador. Existen los supercomputadores con varios procesadores y gran poder de cálculo. La idea es que los PC deberían evolucionar a supercomputadores con muchos chips de CPU en una sola caja en el que cada uno ejecuta una secuencia de sentencias. La sexta generación con millones de simples streams por chips u óptica. Las redes neuronales se definirían como diseñadas para la 6ta generación más el aprendizaje.

base, referido a las reacciones instintivas como el apartar la mano de un objeto caliente hasta el nivel más alto como el del conocimiento de los especialistas (Tabla 1-1)².


Nivel de Entendimiento		Experto, Planeación
		Interacción
		Sentido común
		Visión
		Coordinación
		Regulación
		Reacción

Tabla 1-1. Culturalmente se ha aceptado que el nivel reactivo es el que requiere menor grado de inteligencia y la experiencia (especialización) el más alto grado

Si la inteligencia consistiera solamente en almacenar conocimiento, hacer reconocimiento de patrones, resolución de problemas y la habilidad para aprender, entonces no tendríamos problemas en decir que los programas son inteligentes. Sin embargo hay otras cualidades que algunos críticos consideran necesarias para la inteligencia, algunas de ellas son: la intuición, la creatividad, la habilidad de pensar, la habilidad de entender, el tener conciencia y sentimientos.

Cabe anotar, que aunque podemos ver a la IA como un tipo diferente de inteligencia, no obstante, ésta se plantea como objetivo el **desarrollar programas de computación, algoritmos y arquitecturas de computador que lleguen a realizar actividades que sean semejantes a las que realizan las personas y que hagan aquellas cosas que requieren de inteligencia tales como, entendimiento, pensamiento y razonamiento**. La ciencia se plantea que la mente no es más que un producto del cerebro y en este sentido factible de ser modelada.

Existen avances importantes en el campo de la robótica e Inteligencia Artificial, por ejemplo Asimo (de Honda), el robot más famoso puede entre otras actividades: caminar, subir escaleras, esquivar obstáculos, algunos videos pueden verse en [YouTube](https://www.youtube.com/watch?v=U3333333333). Algunos grupos en Europa están trabajando en el robot iCub, semejante a un niño pequeño, aprende

² Aunque el conocimiento del especialista sea visto como el del más alto nivel, en realidad esto no significa que sea el más complejo. Una vez entendido el problema este puede ser definido en forma clara y específico. En realidad esto hace la diferencia entre fácil y difícil. La respuesta a la pregunta ¿Cuánta inteligencia se requiere para ejecutar una tarea? reside en que existen tareas en que los conceptos no pueden ser claramente definidos o especificados. Por lo tanto si la tarea es clara y bien definida, esto no será complejo para el computador, no así cuando las definiciones no lo son, como por ejemplo el caso del sentido común.

en forma progresiva de sus propias habilidades motoras y también como actuar con el mundo. Existen otros modelos en varios laboratorios y universidades.

Finalmente anotaremos que con el término de “singularidad tecnológica” algunos autores plantean la posibilidad de que se dé un punto futuro de progreso tecnológico sin precedentes causado, en parte, por la capacidad de las máquinas de mejorar ellas mismas utilizando inteligencia artificial. Algunos investigadores esperan que para los años 2030 al 2045 la inteligencia no biológica alcance el rango de la sutileza de la inteligencia humana.

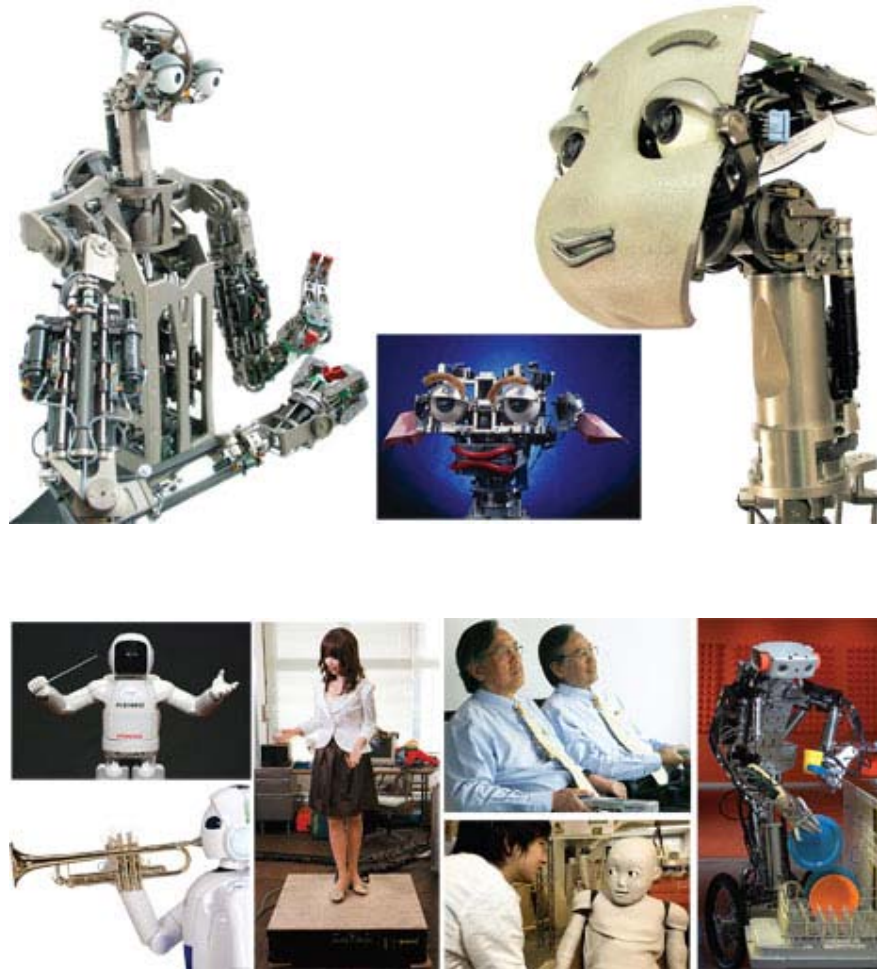


Figura 1-1. Algunos ejemplos de robots que puede consultar en Internet: de arriba hacia abajo y de izquierda a derecha: Máquinas sociables: Domo, Kismet, Merts; Androides: Asimo, Actroid, Zou Ren Ti twin (derecha) Anibots, Toyota droids que interpretan la trompeta y el violín.

Resumiendo, existe un gran objetivo que consiste en encontrar cómo es que trabajan el pensamiento humano y la inteligencia, de tal forma que los mismos métodos o métodos similares puedan llevarse a cabo en un computador. Un segundo objetivo más modesto es producir programas de computación que funcionen de forma parecida a las acciones racionales que realizan las personas, con la finalidad de hacer de la computación algo más

útil para que las computadoras puedan hacer las tareas que realiza la gente, quizás más rápido y en ciertos aspectos, en forma más precisa que el ser humano.

El intento de entender la inteligencia implica construir y probar modelos de la mente y el cerebro, reales y posibles.

Las ciencias de la computación juegan un rol especial en el estudio de la mente y el cerebro ya que forman la base para su desarrollo. “El lenguaje de computación en el estudio de la mente es lo que es el cálculo y las ecuaciones diferenciales en el estudio de la Física Clásica”.

Los dos campos que revisaremos en el presente curso: Redes Neuronales (Neuro Computación) y Computación Evolutiva juntamente con la Lógica Difusa se agrupan en lo que se conoce como Inteligencia Computacional (IC), las dos primeras se conocen como áreas de la IA, la lógica difusa está más cerca al campo de control que al de IA. Las áreas cubiertas por la Inteligencia Computacional son conocidas además por el término de Soft Computing

En general las técnicas típicas de IA tienen un enfoque de arriba hacia abajo (top-to-bottom) en donde por ejemplo la estructura del modelo, soluciones, y otros aspectos son impuestos desde el análisis y la descomposición. Las técnicas de IC son de abajo hacia arriba (bottom-up) donde el orden y la estructura emerge desde un inicio no estructurado.

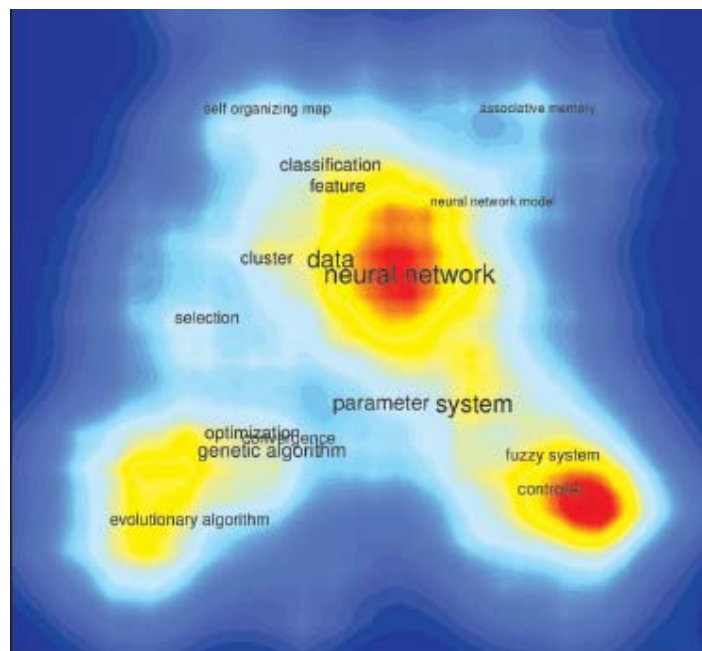


Figura 1-2. Mapa de densidad de conceptos del campo de CI al 2006 según van Eck, et al (2006)

(van Eck Nees Jan, Waltman Ludo, van den Berg Jan, Kaymak Uzay, 2006) Hacen un análisis del campo de la inteligencia computacional basado en los resúmenes presentados

en el congreso mundial de IC tanto del año 2002 como del 2006. De este análisis se concluye, como se muestra en la Figura 1-2, que los conceptos del campo de la CI se agrupan en los tres conjuntos: redes neuronales, computación evolutiva y sistemas difusos. De esta forma se confirma la aceptación generalizada de la conformación del campo de IC.

De la Figura 1-2, se observa que los sub-campos de redes neuronales y lógica difusa están menos separados entre ellos que el de computación evolutiva, en este sentido se puede interpretar que el campo de la CE tiene una posición más independiente. Adicionalmente se observa que el sub-campo de la RN ocupa una amplia parte del espacio en el mapa, lo que sugiere que éste sub-campo está compuesto relativamente de diversos tópicos de investigación. El sub-campo de lógica difusa por otro lado, no toma mucho espacio, el color rojo indica un alto número de conceptos altamente relacionados y todos estos conceptos se reúnen en el tópico de control. Finalmente el sub-campo de CE también cubre algunos tópicos y en este sentido se parece más al sub-campo de RN. El color indica que ningún tópico dentro de este sub-campo domina sobre el otro.

1.1.2 ¿Qué es computación neuronal?

El conjunto de técnicas que hoy se conocen como redes neuronales, emerge y toma su mayor inspiración en el cerebro humano, es así que el campo de la computación neuronal se ha desarrollado con la influencia de dos áreas del conocimiento

- Ingeniería Neuronal la misma que se ve motivada principalmente por los cambios científicos e ingenieriles y se enfoca en el diseño, análisis y aplicaciones de la neurobiología inspirando algoritmos o redes neuronales artificiales para resolver problemas en las ciencias de la computación (Ej. Optimización, predicción) inteligencia artificial (reconocimiento de patrones, representación del conocimiento) e ingeniería (Ej. Robótica), problemas particulares que no han podido ser resueltos usando algoritmos convencionales de las técnicas de las ciencias de la computación.
- Modelamiento Neuronal el mismo que se respalda principalmente por los cambios científicos en las ciencias cognitivas y se enfoca en el modelamiento computacional de la mente y el cerebro para tratar de explicar comportamientos observados (Ej. Percepción, aprendizaje, desarrollo cognitivo). Los modelos computacionales son herramientas invaluable para transformar la Psicología del desarrollo desde una ciencia descriptiva en una ciencia explicativa madura (Mareschal & Thomas in press).

El campo de la computación neuronal comparte con la inteligencia artificial y las ciencias cognitivas, la siguiente hipótesis: “El conocimiento puede ser modelado por la computación.” En el presente curso sobre redes neuronales no enfatizaremos en el estudio de los modelos computacionales de neuronas biológicas sino más bien iremos a las arquitecturas que son herramientas útiles en la solución de cierto tipo de problemas.

Examinaremos sus capacidades de cálculo, con énfasis particular en los algoritmos de diseño, altamente paralelos, adaptativos y tolerables a fallos e inspirados neuralmente.

Una red neuronal artificial no es más que la representación simplificada de ciertas partes del sistema nervioso central de los animales, estas **ayudan en la resolución de ciertos problemas de la informática** y aclaran algunas características del funcionamiento del cerebro. **En el presente curso nos interesa el primer punto.**

Como los distintos tópicos de la inteligencia artificial, en las redes neuronales artificiales, por ejemplo (Nadal, 1988) distingue dos propiedades “inteligentes”.

1. La primera propiedad “inteligente” de las redes de neuronas concierne a la memoria. La memoria humana es diferente a la de un ordenador, ella es asociativa³ (una memoria asociativa es igualmente “direccionable por el conocimiento.”) El físico John Hopfield de la Universidad Técnica de California demostró que las redes neuronales tienen la misma propiedad de asociatividad. Su modelo puede ser considerado como la formulación matemática de las ideas de Donald Hebb de la Universidad de McGill.
2. La segunda propiedad “inteligente” de las R.N.A. es la generalización. El cerebro humano, generaliza sin cesar, los Psicólogos han pretendido evaluar la inteligencia considerando esta capacidad por los test de Q.I. que por ejemplo comprenden muchos ejercicios que demandan encontrar el camino lógico de una secuencia como por ejemplo 1 1 2 3 5 8

³ Reflexionamos un instante en nuestros propios recuerdos. Si pensamos en una persona amiga, nos vendrá a la memoria muchos datos: nombre, edad, color de pelo, estatura, ocupación, formación educativa, familia, casa, experiencias que hemos compartido, de las cuales podemos hasta recordar hábitos de comida y de bebida, aficiones y preferencias y así. Todos esos hechos se articulan y constituyen nuestro recuerdo de esa persona. La forma cómo recuperamos de nuestra memoria este tipo de recuerdos nada tiene que ver con una agenda donde tenemos nombre, dirección y teléfono. En lugar de direcciones de almacenamiento se utilizan fragmentos de la propia información. Imaginemos un espacio de muchas dimensiones cartesianas, cada uno de cuyos ejes estuviese rotulado con uno de los atributos que puede tener una persona. Un eje sería de estaturas; otro el de color de cabello, éste el de peso, aquél el de todas las personas que al comer mariscos en buen estado sienten placer, el de más allá el de todas las Marías, y así sucesivamente. Cada uno de nuestros amigos ocupa un punto particular de ese espacio. Dado que el número de nuestros amigos es muy pequeño en comparación con el conjunto de todos los individuos posibles, si pusiéramos una marca en la posición de cada una de las personas que conocemos, serían muy pocos los puntos marcados en el seno de un espacio muy grande. Cuando alguien nos da información parcial acerca de una persona, por ejemplo sexo, color de pelo y peso, pero no el nombre, está describiendo aproximadamente una porción del espacio de gente posible. La tarea de una memoria asociativa se aplica en este caso a la de hallar, entre nuestros conocidos, aquél que mejor se ajuste a los datos parciales (Hopfield, Tank, 1988)

Las redes neuronales proporcionan soluciones robustas a problemas en un amplio rango de disciplinas, particularmente áreas que involucran: clasificación, predicción, filtrado, optimización, reconocimiento de patrones y aproximación de funciones.

Para tener una visión más clara del campo de las redes neuronales dentro del contexto de la inteligencia artificial, comparemos a estas con los sistemas expertos. Los sistemas expertos reproducen el comportamiento humano en un estrecho ámbito del conocimiento, son programas tan variados como los que diagnostican infecciones en la sangre e indican un tratamiento, los que interpretan datos sísmológicos en exploración geológica y los que configuran complejos equipos de alta tecnología. Al comparar los dos campos podemos anotar las siguientes características:

Sistemas Expertos	Redes Neuronales
Enfoque Descendente	Enfoque Ascendente
Basado en Psicología	Basado en Neurofisiología
Qué hace el cerebro	Como lo Hace el cerebro
Reglas Si / Entonces	Generalización a partir de ejemplos
Sistemas Programados	Sistemas Entrenados
Arquitectura de Von Newmann Separación Hardware / Software	Arquitecturas paralelas, distribuidas, adaptativas. Autoorganización

Tabla 1-2.- Tabla comparativa entre Sistemas expertos y Redes Neuronales

La Tabla 1-2 muestra en forma clara que las Redes Neuronales y los Sistemas Expertos pertenecen a dos categorías muy diferentes, pero que son pensados para la resolución de problemas que de alguna manera podemos decir que muestran “capacidades inteligentes”.

1.1.3 Breve Historia de las Redes Neuronales Artificiales

El inicio de la historia de las RNA data de los años 1940s, la década del primer computador electrónico. Los primeros estudios teóricos importantes en las redes neuronales, se deben a los trabajos de Mc Coulloc (Neurofisiólogo) y Pitts (Matemático) en 1943, quienes presentaron un modelo de neurona biológica y sus componentes conceptuales para circuitos que podrían llevar tareas de cálculo

Sin embargo el primer paso significativo toma lugar en 1958 cuando Rosenblatt introdujo el primer modelo neuronal concreto, el perceptrón. Rosenblatt también tomó parte en la construcción del primer neurocomputador exitoso, el Mark I perceptrón. La Figura 1-3 ilustra el desarrollo de unas de las principales arquitecturas.

El modelo del perceptron original de Rosenblatt contenía solo dos capas (entrada-salida). Luego de éste, un modelo multicapa fue obtenido en 1960. Los primeros usos del perceptrón multicapa (PMC) fueron complicados por la falta de un algoritmo de aprendizaje adecuado. En 1974 Werbos introduce el llamado algoritmo de retropropagación (backpropagation) en su tesis doctoral para una red perceptrón de tres capas (Werbos, 2005). Sin embargo, el área de aplicación del PMC, permanece más bien limitada hasta que en 1986 en que Rummelhart y Mclelland publican en la prestigiosa publicación Nature, el algoritmo general backpropagation para un perceptrón multicapa

Un trabajo importante es el de Hopfield en 1982, quien publicó sus ideas de una red neuronal, diferente a las neuronas en el PMC. La red de Hopfield consiste de una sola capa cuyas neuronas están totalmente conectadas una con otra. Luego de esta primera red, muchas redes de Hopfield fueron desarrolladas. La máquina de Boltzmann tuvo influencias tanto de las redes de Hopfield como del PMC

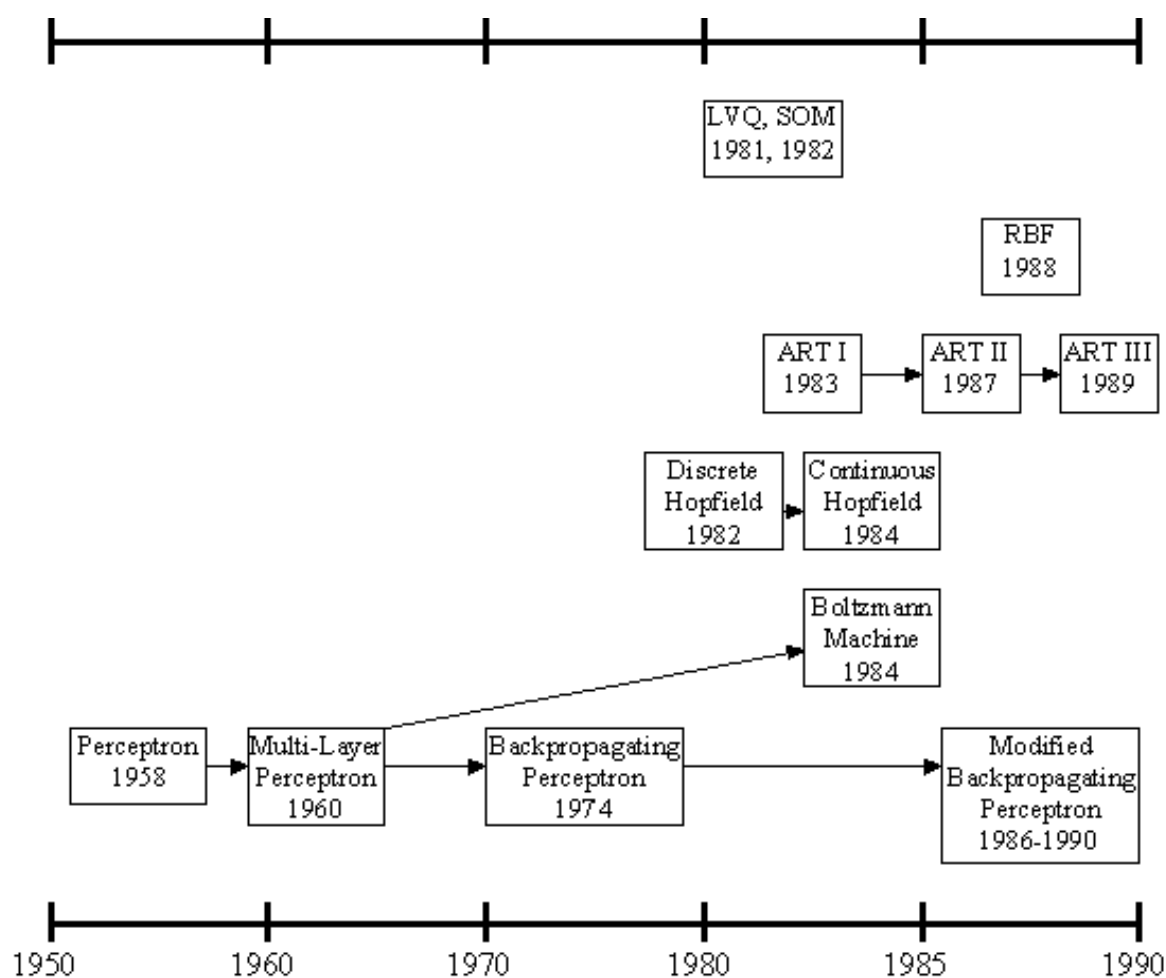


Figura 1-3.- La evolución de las RNA más populares (Modificado de Petri Holdju y Jokoo Halme)

La red ART (Teoría de Resonancia Adaptativa) fue introducida por Carpenter y Grossberg en 1983. El desarrollo de ART continuó, lo que llevó a los modelos de red más avanzados ART II y ART III.

Las Redes de Función de Base Radial (RBF) fueron desarrolladas por Broomhead & Lowe en 1988. Estos trabajos abrieron nuevas fronteras a la comunidad de redes neuronales.

Una clase única de modelo de red son los Mapas Auto organizativos (SOM) presentada por Kohonen en 1982. SOM es una cierta clase de mapa topológico, la misma que se auto organiza, basada sobre los patrones de entrada con los que es entrenada.

1.2 Fundamentos de Las RNA.

Definición: “Una Red Neuronal Artificial es un sistema de computación, hecho por un gran número de elementos simples, elementos de proceso muy interconectados, los cuales procesan información por medio de su estado dinámico como respuesta a entradas externas” (Hech- Nielsen.)

Objetivo: El objetivo general de una RNA es reproducir las funciones realizadas “en forma natural” por los seres vivos tales como: percibir, aprender reconocer, adaptarse a su medio y tomar decisiones.

Modelo: Un modelo es una abstracción simplificada de aspectos seleccionados del fenómeno a ser modelado. Las abstracciones se describen en algún lenguaje seleccionado. Por ejemplo el cálculo y las ecuaciones diferenciales constituyen el lenguaje para modelar la Física Clásica. Para nosotros, los algoritmos constituyen el lenguaje para el modelamiento neuronal. Una vez con el lenguaje, se debe trazar el conjunto total de herramientas disponibles (Ej. Análisis de algoritmos) dentro del paradigma de modelamiento escogido.

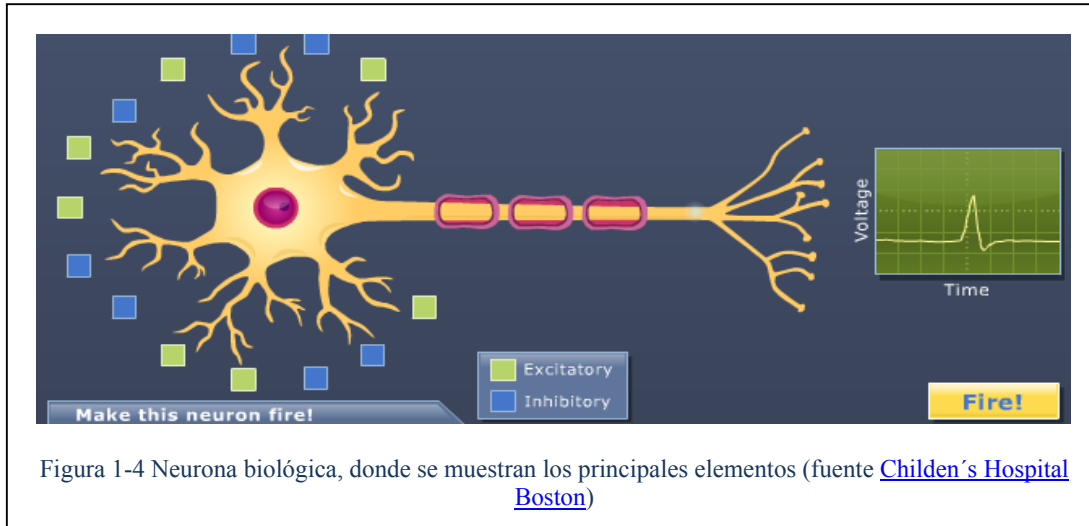
Un modelo (o hipótesis) es tan bueno como las explicaciones que proveen para el comportamiento observado del sistema a ser modelado. Estos modelos están sujetos a revisión, a la luz de la evidencia experimental.

Es importante entender exactamente, que mientras se usa un formalismo de modelización simplificado (ej. modelos de redes neuronales) que aspectos del sistema se están ignorando, y basados sobre las simplificaciones, que predicciones pueden hacerse en los diferentes niveles. Por ejemplo en el caso del algoritmo backpropagation, usado para entrenar una clase de redes neuronales artificiales, puede ser un algoritmo biológicamente irreal, sin embargo, justamente, este es uno de los varios algoritmos que pueden llevar a cabo una de las tareas particulares de computo – gradiente descendente o superficie de error – y por qué no decirlo, esto podría dar pistas sobre los tipos de computo biológico que puede estar dándose en la realidad.

Un modelo de necesidad, abstrae y simplifica (de otra forma sería tan complejo como el sistema que se está tratando de modelar).

1.2.1 Estructura de la Neurona.

Para entender el comportamiento de una red neuronal, empezaremos introduciendo un modelo muy simple de neurona biológica.



Lo que sigue es un cuadro simplificado que describe el trabajo del sistema nervioso. Una neurona típica consiste de tres partes: **las dendritas, el cuerpo celular, y el axón,**

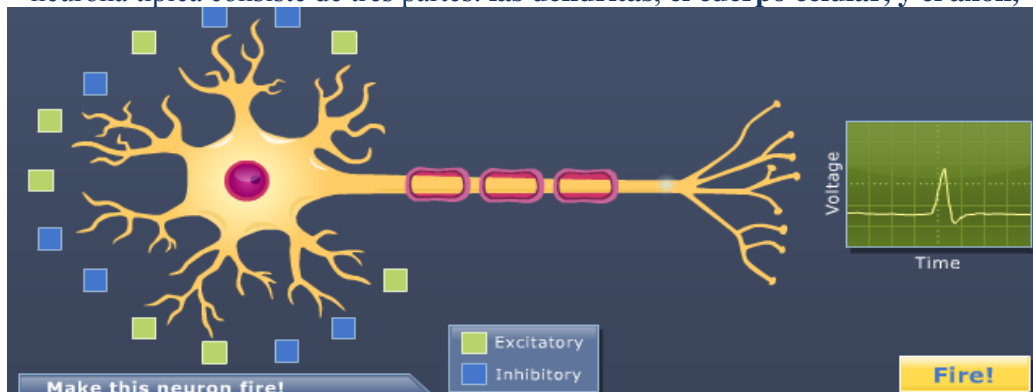


Figura 1-4. Las dendritas conducen las señales nerviosas hacia el cuerpo celular mientras que el axón lo hace desde el cuerpo celular hacia otras neuronas. Los circuitos celulares están formados por grupos de neuronas conformados con los brazos terminales del axón y cierran el circuito con las dendritas de otras neuronas. El punto de contacto entre las componentes de dos neuronas (dendritas) se llama **sinapsis**. Existe un pequeño espacio microscópico entre la sinapsis de las neuronas.

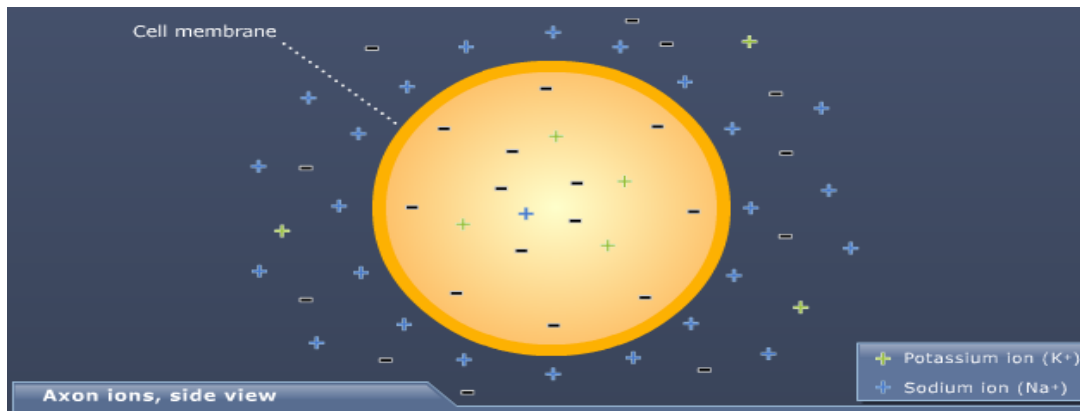


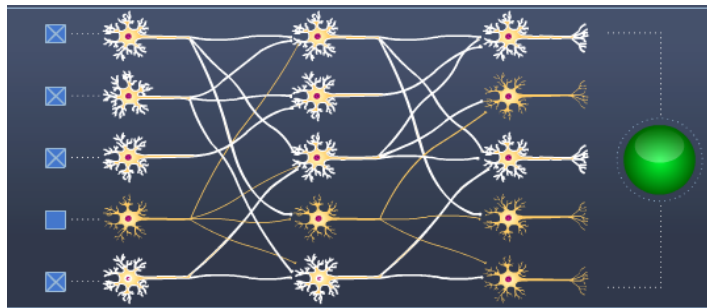
Figura 1-5. Diferencia de cargas a nivel de la membrana celular que genera una diferencia de potencial en la misma, (fuente [Children's Hospital Boston](#))

A nivel de la membrana celular, existe una concentración de cargas negativas y positivas como se observa en la Figura 1-5, que genera una diferencia de potencial, en estado de reposo esta diferencia está en alrededor de 70-80mV, que puede ser medida a través de la membrana plasmática, siendo el interior negativo con respecto al exterior. Este potencial bio-eléctrico es el resultado de la diferencia de concentración de varios aniones y cationes dentro y fuera de la célula, que depende en última instancia de las propiedades de permeabilidad de la membrana externa y del transporte activo de los iones, especialmente sodio y potasio por la membrana. Cuando una neurona recibe la entrada de otras neuronas (cuadrados verdes y azules), el proceso electroquímico involucrado causa que su voltaje se incremente. Cuando el voltaje excede un cierto umbral, resulta una descarga del impulso nervioso en forma rápida que viaja hacia el axón (estimulando a otras neuronas), se dice que la neurona se ha activado. El potencial de reposo puede variar de dos formas distintas, dando lugar a los potenciales graduales o a los potenciales de acción, respectivamente. Los potenciales graduales se dan en las membranas de las dendritas y del soma neuronal, y consiste en fluctuaciones de aumento o disminución del potencial de reposo. Se dice que estas membranas están relativamente despolarizadas o hiperpolarizadas. Por el contrario, el potencial de acción, consiste en un cambio completo, transitorio y reversible de la polaridad a través de la membrana de los axones (y también de algunas dendritas largas que parecen axones en cuanto a estructura).

Las variaciones graduales de potencial pueden ser excitatorias o inhibitorias. Cuando son excitatorias, ocurre un aumento de la permeabilidad de la membrana a los iones de sodio, los cuales disminuyen su gradiente de concentración al entrar en la célula y, consecuentemente, despolarizan progresivamente la membrana. Este súbito cambio es el resultado de la apertura transitoria de canales iónicos especiales, llamados puertas de sodio y de potasio. El tiempo que tarda en producirse el fenómeno y el tamaño alcanzado por la despolarización en un determinado punto de la superficie, de una membrana al aplicar un estímulo activo es proporcional al logaritmo de la energía del estímulo. Es más las zonas de membrana que rodean el área receptora también se despolarizan pero con amplitud decreciente conforme nos alejamos del foco inicial. Cuanto más grande sea la despolarización inicial, tanto más rápidamente se afectan las zonas de membrana que

rodean al foco. Este mecanismo favorece la SUMA de los efectos de los múltiples estímulos excitatorios aplicados sobre numerosos puntos de la superficie de la membrana.

Los estímulos inhibitorios actúan facilitando la penetración de iones de cloro, con carga negativa, dentro de la célula, lo que tiende a aumentar la carga negativa intracelular y, por lo tanto la membrana se hiperpolariza, oponiéndose o disminuyendo el estado excitatorio total de la neurona



Estas neuronas forman circuitos en los que las señales pueden o no pasarse, dependiendo de la actividad conjunta de las neuronas. La ilustración muestra este proceso

Para tener una comprensión más clara de este proceso, se le recomienda revisar una animación de este mecanismo en la página del [Children's Hospital Boston](#).

1.2.2 Elementos de la Red Neuronal Artificial

Las arquitecturas que revisaremos, son en realidad abstracciones muy simplificadas de las redes reales ya que estamos interesados en las herramientas para análisis de datos, no en los procesos en sí mismos. Las arquitecturas que revisaremos en los capítulos siguientes son algoritmos que pueden ser implementados sin mayores complicaciones.

Ahora nuestro objetivo es extraer los elementos principales del modelo anterior para modelarlo. Extraemos los siguientes elementos.

- El conjunto de unidades de procesamiento (neuronas)
- Un estado de activación de cada neurona, la cual equivale a la salida de la unidad (y_j)
- El patrón de conectividad entre las unidades (neuronas). Generalmente cada conexión es definida por un peso w_{ij} , los cuales determinan el efecto que tiene la unidad i sobre la unidad j .
- Una regla de propagación, la cual determina la entrada efectiva, s_j , a una unidad desde sus entradas externas
- Una función de activación F_j , la cual determina el nuevo nivel de activación basado en la entrada efectiva $s_j(t)$ y el nivel actual $y_j(t)$.
- Una entrada externa (bias, umbral) θ_j para cada unidad, que tiene el efecto de rebajar la entrada de la red al rango de la función de activación.

- Un medio ambiente dentro del cual el sistema debe operar, proveyendo señales de entrada, y si es necesario, señales de error.
- Un método para inferir la información: la regla de aprendizaje. Este último no parecería derivarse del modelo de la neurona, sin embargo no hay que olvidar que estamos asumiendo que las neuronas y redes neuronales tienen que ver con el aprendizaje y la memoria.

Las redes neuronales artificiales (RNA) son sistemas paralelos para el procesamiento de la información, inspirados en el modo en el que las redes de neuronas biológicas del cerebro procesan información.

1.2.2.1 Unidad de procesamiento.

Cada unidad lleva a cabo un trabajo simple: recibe entradas de fuentes externas y usa estas para calcular la señal de salida que se propaga a otras unidades. Las entradas para la neurona j (dentritas en el caso de la neurona biológica), las llamaremos $X_j = (x_{1j}, \dots, x_{nj})$ o vector de entrada.

Una segunda tarea está relacionada con el ajuste de los pesos sinápticos (la fuerza del acoplamiento sináptico o conexiones entre neuronas) estos se modelan como parámetros numéricos, $W_j = (w_{1j}, w_{2j}, \dots, w_{nj})$ o vector de pesos

En una red, se distinguen tres tipos de neuronas:

1. De entrada: recibe señales desde el entorno; por ejemplo sensores
2. De salida: envían señales fuera del sistema; salida de la red
3. Ocultas: no tienen contacto exterior; generan la representación interna de la información.

La forma más simple de modelar el efecto de las diferentes neuronas a una neurona específica (j) es a través de una sumatoria. De esta forma, la entrada total a la neurona j (S_j) viene dado por la siguiente fórmula:

$$S_j = \sum_i^n w_{ij} x_i$$

Ecuación 1-1 Suma ponderada de los pesos

1.2.2.2 Conexión entre Neuronas.

En la mayoría de los casos se asume que cada unidad da una contribución aditiva a la unidad a la cual se conecta. La entrada a la unidad j será S_j más el bias o umbral θ_j . En un principio, podemos considerar a este umbral, como el valor a superar para que la neurona se active.

Se puede tener sinapsis excitadoras o inhibitoras. Cuando el valor de activación de la neurona i , x_i , multiplicada por los pesos, w_{ij} , es positivo, entonces la unidad i sirve para

activar la unidad j . De otro modo, cuando el valor de activación de la neurona i , x_i , multiplicada por w_{ij} es negativa, entonces la unidad i sirve para inhibir a la unidad j . El valor de activación de la neurona es dado por una función $f(S_j)$, llamada función de activación o función de transferencia.

$$y_j = \sum_i w_{ij}(t)x_i(t) + \theta_j(t)$$

Ecuación 1-2.- Salida de la neurona

En algunos casos, reglas más complejas para combinar las entradas son usadas en los cuales se hace una distinción entre unidades excitadoras e inhibidoras.

1.2.2.3 Función de activación o de salida:

En una neurona biológica, la señal puede transmitirse o no a otras neuronas, este fenómeno está relacionado con procesos químicos, de tal forma que, por ejemplo, en uno de los casos, si se sobrepasa un cierto potencial, la señal se transmite, caso contrario no se transmite.

La forma más simple de modelar este proceso es mediante una regla, de tal forma que si se sobrepasa el umbral, la neurona se active, caso contrario no, este sería el caso más simple, pero no el mejor.

Podemos pensar en una función, la misma que dé el efecto de la entrada total sobre la activación de la unidad. Es decir, una función f que tome la entrada total S_j y la activación actual y_j y produzca un nuevo valor de activación de la unidad j

$$y_j(t+1) = f_j(y_j(t), S_j(t))$$

Ecuación 1-3.- Salida al tiempo $t+1$

$\varphi(x) = \begin{cases} 1 & \text{si } x \geq 0 \\ -1 & \text{si } x < 0 \end{cases}$	$\varphi(x) = \begin{cases} -1 & \text{si } x < -c \\ 1 & \text{si } x > c \\ a.x & \text{en otro caso} \end{cases}$	$\varphi(x) = \frac{1}{(1 + e^{-ax})}$
Ecuación 1-4.- funciones: escalón; semi lineal; y sigmoidal		

Una función de activación sencilla es tomar simplemente la suma de las entradas a la neurona, S_j , este será el valor de activación de la neurona.

Otra función de activación común es probar si la suma S_j es mayor que un valor umbral (mínimo) entonces la señal se propaga (la neurona se activa, usualmente con un valor de activación igual a 1) de otra forma la neurona permanece en el mismo estado (usualmente con un valor de activación igual a 0, por ejemplo esto se simula con la función escalón).

Otras funciones de transferencia típicas son: Función lineal y función sigmoidal (Figura 1-6)

La función lineal es inadecuada para aproximar a la mayoría de las funciones (problemas), en este caso trabajan mejor las funciones no lineales

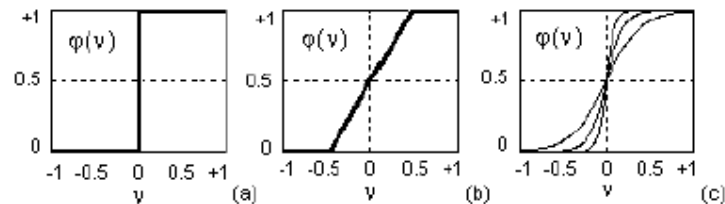


Figura 1-6 Funciones de Activación. (a) Función Umbral. (b) Función lineal por partes. (c) Función Sigmoidea y Tangente

En la Figura 1-6c observamos diferentes funciones sigmoides superpuestas, indicando una dependencia con la pendiente de las ecuaciones 1.4 (tercera ecuación): cuanto mayor sea esta, más abrupta será la caída o pendiente de la función de activación y más se asemejará a la función de activación escalón (Figura 1-6a)

Otra función usada comúnmente es la función tanh, su salida está en el rango de -1 a 1 y puede ser escrita como

$$y = 2 / (1 + \exp(-2 * x)) - 1$$

Ecuación 1-5 Función tangencial

Aunque las funciones de activación tangente hiperbólica y sigmoidal son las más empleadas, las posibilidades son más amplias. Por ejemplo, la posibilidad de descomponer cualquier función en series de Fourier lleva a la conclusión que podría emplearse una función de activación sinusoidal.

1.2.2.4 Regla de aprendizaje.

Biológicamente se acepta que la información memorizada en el cerebro está más relacionada con los valores sinápticos de las conexiones entre neuronas que con ellas mismas. En el caso de las redes neuronales artificiales, se puede considerar que el conocimiento se encuentra representado en los pesos de las conexiones entre neuronas. Todo proceso de aprendizaje implica un cierto número de cambios en estas conexiones (w_{ij}).

La regla de aprendizaje difiere según este sea supervisado o no supervisado.

En el aprendizaje no supervisado, no existe información de salida, mientras que en el aprendizaje supervisado, si se conoce la salida.

La memoria de las neuronas formales reside en los pesos sinápticos.

1.3 Características de una red neuronal artificial.

Las redes neuronales pueden ser clasificadas o caracterizadas desde cuatro aspectos diferentes, siguiendo a (Hilera y Martínez, 1995):

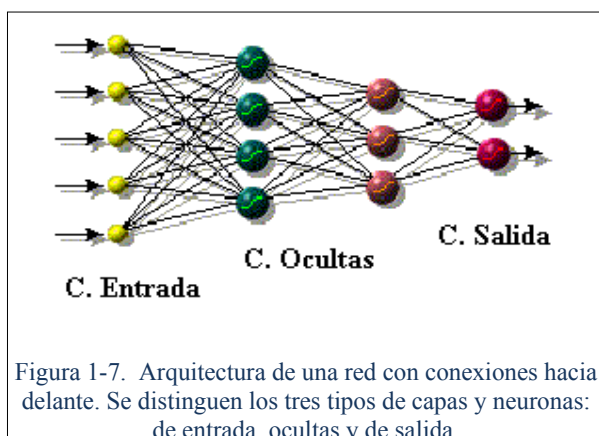
- Topología
- Mecanismo de aprendizaje
- Tipo de Asociación realizada entre la información de entrada y salida
- Forma de representación de la información de entrada y salida

1.3.1 Topología de las RNA

La topología o arquitectura de una red consiste en la organización y disposición de las neuronas en la red formando capas de neuronas, en este sentido, los parámetros fundamentales para esta distinción son:

1. Número de capas,
2. Número de neuronas por capa,
3. Grado de conectividad y
4. Tipo de conexiones entre neuronas.

1.3.1.1 Por el número de capas.



En una red neuronal artificial, La neuronas están conectadas entre sí, formando capas o niveles como se ilustra en la Figura 1-7.

Al igual que tenemos tres tipos de neuronas, lo mismo se da con las capas.

De entrada: Codifica las variables que definen el problema.

Ocultas: Internas a la red, no tienen contacto con el entorno exterior. Su importancia radica en

que generan la representación de la información, en el sentido de representar todas las conexiones posibles entre todas las capas

De salida: Transfieren información de la red hacia el exterior

Por el número de capas se distinguen entre las redes mono capa y multicapa

1.3.1.1.1 Redes mono capa (1 capa):

En la Figura 1-8b se presenta un modelo de red de una capa simple totalmente conectada a todas las neuronas vecinas en la capa. En este tipo de red, todas las neuronas son a la vez de entrada y de salida, también pueden existir conexiones auto recurrentes (salida de la neurona conecta a su propia entrada). La red mejor conocida de este tipo es la red de Hopfield

Generalmente este tipo de redes se utiliza en las tareas denominadas de auto asociación, por ejemplo para regenerar señales que se presentan incompletas o distorsionadas como el caso de señales con ruido.

1.3.1.1.2 Redes multicapa

Son redes con grupos de neuronas agrupadas en más de una capa. Normalmente, todas las neuronas de una capa reciben señales de entrada de otra capa y envían la salida a la capa posterior, más cercana a la salida de la red, a este tipo de conexiones se denominan conexiones hacia adelante o feedforward.

También se pueden conectar las salidas de las neuronas de las capas posteriores a las entradas de las capas anteriores (conexiones hacia atrás).

- Redes con conexiones hacia adelante (feedforward)

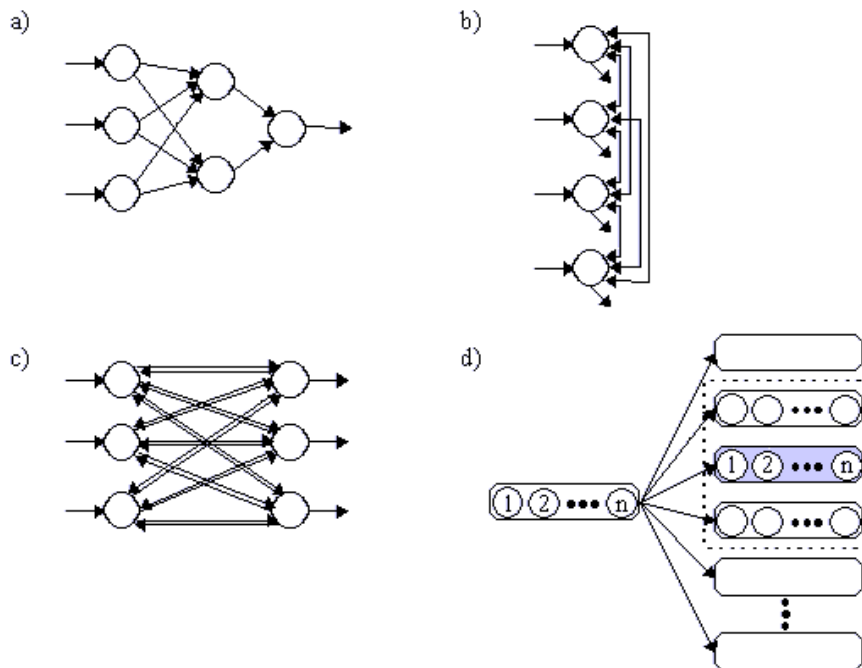


Figura 1-8: Diferentes estructuras de RNA. a) Redes multi-capa con conexión hacia adelante, b) Redes totalmente conectadas de una capa-simple, c) Redes de dos capas feedforward/feedbackward y d) mapa de vector topográficamente organizado. (Fuente, Petri)

La Figura 1-8a) representa la estructura de una red multi-capa con conexiones hacia adelante. En este tipo de red, todas las señales neuronales se propagan hacia delante (de izquierda a derecha) a través de las capas de la red. No existen conexiones hacia atrás, por lo general, tampoco se dan conexiones auto-recurrentes, ni laterales.

Uso: Reconocimiento o clasificación de patrones

Ejemplos de estas redes son: Perceptron, Adaline, Madalena, Backpropagation., Base Radial

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

Figura 1-8c) demuestra las conexiones en una red de dos capas con conexiones hacia adelante/ hacia atrás (feedforward/feedbackward). Las capas en este modelo están conectadas en ambas direcciones, la información circula tanto hacia adelante como hacia atrás, durante el funcionamiento de la red. Cuando un patrón se presenta a esta red, esta “resuena” un cierto número de veces entre las capas antes de recibir una respuesta de la capa de salida. El mejor tipo de red conocido de este tipo es la red de Teoría de Resonancia Adaptativa (ART).

Uso: Este tipo de estructura es adecuado para tareas de hetero-asociación, es decir asociar la información del patrón de entrada con la información del patrón de salida.

La Figura 1-8d) ilustra la idea de un mapa de características organizado. En este modelo, cada neurona en la red contiene un vector de características. Cuando un patrón de entrenamiento se le da a la red, la neurona cuyo vector característico es más parecido al vector de entrada se activa y esta es actualizada para reflejar el vector de entrada que motivó la activación. En el proceso de actualización, también se actualizan las neuronas vecinas. El tipo de red que exhibe este tipo de comportamiento son los mapas auto-organizativos de Kohonen. (SOM, por sus siglas en inglés).

1.3.2 Mecanismo de Aprendizaje

El aprendizaje, no es más que modificar los pesos, en respuesta a la información de entrada. Esto implica:

Destrucción de las conexiones: $W_{ij} = 0$.

Modificación de las conexiones $W_{ij}(t+1) \neq W_{ij}(t)$.

Creación de conexiones: $W_{ij} \neq 0$.

Se dice que el proceso de aprendizaje ha terminado o la Red ha aprendido cuando los pesos permanecen estables: expresando en forma matemática se tiene que la derivada parcial de los pesos con respecto al tiempo tiende a 0:

$$\frac{\partial W_{ij}}{\partial t} \approx 0.$$

En el contexto de redes neuronales se distingue entre la etapa de aprendizaje que se refiere al proceso por el cual la red modifica los pesos ante la información de entrada. Una vez que la red ha aprendido, esta está lista para realizar la tarea para la cual se la entrenó (a esta etapa podemos denominarla de funcionamiento). Normalmente en la mayoría de las redes se distingue entre etapa de aprendizaje y etapa de funcionamiento. En función de la forma de aprendizaje, se puede clasificar a una red como:

1. Redes ON-LINE: La Red “aprende” durante su funcionamiento: W_{ij} varían para adaptarse a la nueva información de entrada todo el tiempo, aunque al inicio la variación de pesos es mayor, en la de funcionamiento las modificaciones son mínimas por obvias razones.
2. Redes OFF-LINE: Se distingue entre una fase de entrenamiento o aprendizaje: (W_{ij} varía) y una de funcionamiento (W_{ij} constante), en la fase de entrenamiento, la red llega a fijar los pesos, esto implica que los mismos permanecerán constantes en todo su proceso de funcionamiento. Si necesitase que la red aprenda un nuevo tipo de patrón, se debe re entrenar nuevamente la red.

1.3.2.1 Reglas de Aprendizaje

Por el proceso de aprendizaje, se distingue entre:

1.3.2.1.1 Aprendizaje supervisado: Implica la existencia de un “supervisor” que controle el proceso de aprendizaje. Esto implica que se debe diseñar un mecanismo que ajuste los pesos en función de cuán parecida es la salida de la red a la salida deseada. La variación de la eficacia sináptica se estabiliza cuando las variaciones de los pesos, dw_i , se vuelve nula en promedio. De esta forma, en una red de este tipo los pesos pueden evolucionar con una ecuación de la forma:

$$dw_j = ux_j(t)(y(t) - y_{d_j}(t))$$

Ecuación 1-6.- Variación de pesos para aprendizaje supervisado

$0 < u \leq 1$ Regula la velocidad de aprendizaje.

Entre las formas de aprendizaje más conocidos dentro de este tipo están

Aprendizaje por corrección de error:

Se basan en la comparación de las salidas obtenidas por la red, los Y , con la salida deseada, Y_d para modificar los pesos W_{ij} . Por lo tanto esta diferencia define el error de la red y el objetivo es minimizar este error.

Por ejemplo, la ecuación de la expresión 1.6 representa una forma de medir un error. Y los pesos se actualizan tomando en cuenta esta variación y el valor actual

de los mismos, la idea es que esta variación es el valor en que se modificara a los pesos.

$$\Delta W_{ij} = u X_i(Y_j - Y_{dj}) = W_{ij}^{(actual)} - W_{ij}^{(anterior)}$$

Ecuación 1-7.- Un ejemplo para modificar los pesos

Este algoritmo fue implementado por la primera red neuronal formal (el perceptrón), pero presenta un fuerte inconveniente ya que no considera el error en su conjunto.

Regla Delta. La regla delta toma en cuenta el error medio cuadrático y lo minimiza para la modificación de los pesos. Esto corrige el problema del algoritmo anterior

$$E_T = \frac{1}{2P} \sum_{k=1}^P \sum_{i=1}^N (Y_i^{(k)} - Y_{di}^{(k)})^2$$

Ecuación 1-8.- Ecuación de error

En donde:

N → Número de neuronas de salida

P → Información que debe aprender la red (número de vectores para el entrenamiento)

Los pesos se ajustan minimizando este error de la siguiente forma.

$$\Delta W_{ij} = u \frac{\partial E_T}{\partial W_{ij}}$$

Aprendizaje Estocástico.

Consiste en realizar cambios aleatorios en los pesos W_{ij} y evaluar su efecto a partir del objetivo deseado y de la distribución de probabilidades.

Por lo general estos métodos parten de principios físicos que se dan en los sólidos (Termodinámica) El objetivo es llegar a un estado de mínima energía, evaluada con alguna función de energía.

Si $E(t+1) < E(t)$ se acepta el cambio de los pesos

Si $E(t+1) \geq E(t)$ se acepta el cambio en función de una función de probabilidades preestablecida.

Redes que usan este tipo de algoritmo de aprendizaje son: Máquina de Boltzman y Máquina de Cauchy

Uso: Reconocimiento de patrones (imágenes de sonar y de radar) y optimización

1.3.2.1.2 Aprendizaje No Supervisado: En este tipo de aprendizaje no se requiere de la influencia externa para ajustar los pesos de las conexiones entre neuronas.

No se tiene información que le indique a la red si la salida generada es o no correcta (en respuesta a la entrada). En este sentido, se habla de capacidad de auto organización. El objetivo es encontrar características, regularidades o correlaciones entre los datos de entrada.

La salida es generada teniendo en cuenta:

1. El grado de familiaridad o similitud entre la información de entrada y las del pasado
2. Agrupando en categorías a partir de correlaciones entre la información de entrada

Por lo tanto, si no se tiene una salida con que comparar aquella de la red, y si la red tiene almacenada la información en los pesos, lo lógico es que compare la información de entrada con la información del sistema, por lo que una regla de aprendizaje podría ser:

$$dw_j = uy(t)(x_j(t) - w_j(t))$$

Ecuación 1-9.- Variación de pesos para aprendizaje no supervisado

Con $u > 0$.

Entre estos tipos de aprendizaje se tienen:

Aprendizaje Hebbiano:

Este tipo de aprendizaje se basa en el postulado formulado por Donald Hebb en 1949: "Cuando un axón de una célula "(pre-sináptica)" está suficientemente próxima (espacialmente) para excitar una célula B (post-sináptica) y de una manera repetida y persistente toma parte en la puesta en marcha de su actividad nerviosa un proceso de crecimiento o un cambio metabólico interviene en una o en ambas células, de tal manera que la eficacia de A, como una de las células que dispara la actividad de B, aumenta⁴. Por célula, Hebb se refiere a un conjunto de neuronas fuertemente conectadas 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.

$$\Delta w_{ij} = y_i * y_j$$

Ecuación 1-10.- la regla de Hebb en su forma más elemental

Esta expresión responde a la idea de Hebb, puesto que 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.

⁴ Fuente: Yves Frénac (1994)

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 competencia entre neuronas se realiza en todas las capas de la red, existiendo en estas neuronas conexiones recurrentes de auto-excitación 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 “agrupar” los datos que se introducen en la red. De esta forma, información similar es clasificada formando parte de una misma categoría, y por tanto deben activar la(s) misma(s) neurona(s) 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, que ha sido producto de la adaptación de la red ante los datos de entrada. El aprendizaje afecta sólo a las neuronas ganadoras (activas), y a las de su vecindad. la variación del peso de una conexión entre una unidad i y otra j será nula si la neurona j no recibe excitación por parte de la neurona i y otra j será nula si la neurona j no recibe excitación por parte de la neurona i , y se modificará si es excitada por dicha neurona i .

1.3.2.1.3 Aprendizaje por refuerzo (Penalización /o Recompensa):

Existe otro tipo de problemas que no pertenece a aprendizaje supervisado porque no se conoce el detalle de la salida deseada, y tampoco pertenece aprendizaje no supervisado ya existe una meta a la que se quiere llegar. Es decir no se disponen de un conocimiento completo del comportamiento o salida deseada en forma exacta.

Este tipo de problemas pertenecen al aprendizaje reforzado. Para alcanzar la meta se ejecutaran acciones que no se conocen, en un principio, pero que llevaran a un resultado. Si el resultado favorece alcanzar la meta propuesta, se recompensa la acción por ejemplo esto implica un éxito (+1) pero si la salida no lleva a la meta o la aleja de ella, se penaliza la acción, esto implica un fracaso (-1). Los pesos se ajustan por probabilidades.

1.3.3 Tipo de Asociación entre la Información de Entrada y de Salida

Una red neuronal como todo sistema recibe unos datos de entrada y entrega unas respuestas o datos de salida. La pregunta que podemos plantearnos es ¿qué relación existe entre esos datos de entrada y los de salida? Ya sabemos que en una red neuronal, la información aprendida se registra en forma distribuida en los pesos W_{ij} y en función de estos y otros parámetros la red da una respuesta. Existen dos formas básicas de realizar la asociación entre la información de entrada/salida.

1. La primera denominada hetero-asociación y se refiere al caso en el que la red aprende parejas de datos $[(X_1, Y_1), (X_2, Y_2)\dots]$, de modo tal que cuando se presenta una información de entrada X_1 deberá responder generando la correspondiente salida asociada Y_1 .
2. La segunda se conoce como auto-asociación y en este caso la red aprende cierto conjunto de datos de entrada X_1, X_2, \dots, X_n , de forma tal que cuando se le presenta

una información de entrada realizará una auto correlación, respondiendo con uno de los datos almacenados, que será el más parecido al de la entrada.

Estos dos mecanismos de asociación dan lugar a dos tipos de redes neuronales: Las redes hetero-asociativas y las redes auto-asociativas respectivamente.

En el primer grupo encontramos aquellas redes que “calculan” ciertas funciones, que no pueden representarse analíticamente, entre un conjunto de entradas y un conjunto de salidas, haciendo corresponder a cada posible entrada una determinada salida. Entre las redes auto asociativas encontramos aquellas cuya misión es reconstruir una determinada información de entrada que se presenta incompleta o distorsionada (le asocia el dato almacenado más parecido).

En realidad, estos dos tipos de redes no son diferentes en principio, ya que una red hetero-asociativa puede ser reducida a una red auto asociativa concatenando una información de entrada y su salida asociada para obtener la información de entrada de la red auto asociativa equivalente. De modo similar, una red auto asociativa puede convertirse en una hetero-asociativa presentando, como entrada parcial de la auto asociativa, la información de entrada para la hetero-asociativa y haciendo que la red complete la información para producir lo que sería la salida de la red hetero-asociativa equivalente.

1.3.3.1 Redes Auto asociativas

Una red auto asociativa asocia una información de entrada con el ejemplar más parecido de los almacenados conocidos por la red. Este tipo de redes puede implementarse con una sola capa de neuronas. Esta capa comenzará reteniendo la información de entrada y terminará representando la información asociada.

En cuanto a la conectividad, en este tipo de redes, existen conexiones laterales entre las neuronas (Hopfield, etc.) y, en algunos casos, conexiones auto concurrentes (Brain-State-In-A-Box, Additive Grossbarg, etc.)

Patrones: X_1, X_2, \dots, X_n . Ante una entrada X_i , se realiza una auto correlación, y se asocia al más parecido.

Uso: Este tipo de redes suelen utilizarse en tareas de filtrado de información para la reconstrucción de datos, eliminando distorsiones o ruidos. También se utilizan para buscar relaciones entre informaciones similares, para resolver problemas de optimización y para búsquedas por contenido en bases de datos.

Aprendizaje: En relación con el tipo de aprendizaje, el más comúnmente utilizado es el no supervisado (Hopfield, Additive Grossbarg, etc.), aunque existen algunas con aprendizaje supervisado (Brain-State-In-A-Box)

1.3.3.2 Redes Hetero asociativas

Estas redes, al asociar ciertas informaciones de entrada con informaciones de salida, precisan al menos dos capas: una para captar y retener la información de entrada y la otra para mantener la salida con la información asociada. Si esto no fuera así, se perdería la

información inicial al obtener el dato asociado, lo cual no debe ocurrir, ya que en el proceso de obtención de la salida se puede necesitar acceder más de una vez a esta información.

En cuanto a su conectividad, existen redes hetero-asociativas con conexiones hacia adelante o feedforward (perceptrón, Backpropagation, etc.), redes con conexiones hacia atrás o feedforward/feedback (Art, Bam, etc.) y redes con conexiones laterales.

Uso: Las redes hetero-asociativas pueden utilizarse para diferentes fines: en algunos casos el objetivo es calcular una función general de los datos de entrada, en otros casos el objetivo será realizar una clasificación, relacionando un gran número de variables de entrada con un pequeño número de variables de salida, que representan los conjuntos en los que se pueden clasificar dichos datos.

Aprendizaje: El aprendizaje de este tipo de redes puede ser con supervisión (Perceptrón, Backpropagation, etc.) o sin supervisión (Art, Bam, Kohonen, etc.).

Calcula cierta función (por lo general no expresable analíticamente):

$(X_1, Y_1), (X_2, Y_2), \dots, (X_n, Y_n)$

Respuesta

Entrada A_i B_i (salida asociada)

1.3.4 Representación de la Información de Entrada y Salida

La representación de la información se refiere al hecho de cómo codificar los datos que alimentaran la red (datos de entrada) y los datos de resultados (salida de la red). Las redes neuronales son sistemas que requieren de datos numéricos para procesar la información. Muchas de las variables, tanto de entrada o de salida vienen expresadas en valores reales o continuos como por ejemplo el valor de la temperatura. Existen otro grupo de variables que representan tipos diferentes de clases como por ejemplo el color. Supongamos que una variable es el color y tenemos solamente cuatro colores: verde, amarillo, rojo y azul, nos topamos con el problema de ¿cómo representar esta variable? Para este tipo de casos la representación recomendada es discreta. La recomendación es usar tantas neuronas para la variable (color) como clases (verde, amarillo, rojo y azul) se tenga:

Clase (color)	representación
verde	1000
amarillo	0100
rojo	0010
azul	0001

Tabla 1-3.- Representación o codificación de clases

Por lo tanto, tenemos que los datos pueden tomar valores continuos (elementos de los reales) y valores discretos en el caso de representación de clases o datos específicos (normalmente 0 y 1). En este sentido se tienen los siguientes tipos de redes.

Tipo de Datos - Red -	Entrada	Salida	Función de Activación
Análoga E \mathfrak{R}	D. Continuos	D. Continuos	Lineal, sigmoides
Discreta	D. Discretos o binarios	D. Discretos o binarios	Escalón
Híbrida	D. Continuos	D. Discretos	

Tabla 1-4 Tipos de redes por la forma en que se representa la información

1.4 Áreas de Aplicación de las RNA

No existen reglas definidas para definir el dominio exacto de aplicación de una u otra arquitectura, las áreas de aplicación generales de las RNA son: reconocimiento de patrones, filtrado, segmentación de datos, compresión de datos, optimización, modelamiento de funciones complejas entre otras.

Tratando de agrupar el tipo de aplicaciones de las RNA, podemos clasificarlas en los siguientes grandes grupos:

- **Clasificación:** La clasificación se refiere a la discriminación entre dos o más cosas basadas en semejanzas, tales como por ejemplo la clasificación que hacemos entre los animales y plantas y las diferentes sub clasificaciones entre estos. Otro ejemplo diferente es la separación entre aplicaciones de riesgo buenas y malas o la distinción una transacción legal de una fraudulenta. Pueden también ser utilizados para discriminar entre las actividades criminales y legales.
- **Agrupamiento (Clustering):** Se refiere a la organización de las observaciones en grupos con características similares, como por ejemplo grupos de los clientes que compran el mismo tipo de productos, también referido como análisis afinidad de cesta - mercado; podríamos pensar en agrupar a los jugadores de fútbol, o a los alumnos y profesores. Un tipo de redes usadas con este propósito son los mapas auto-organizativos o redes de Kohonen
- **Generalización:** la generalización referida a la capacidad de ir a casos nuevos, por ejemplo lo que hace un niño cuando está aprendiendo a leer y escribir, o nuestro aprendizaje, generalizamos a partir de un conjunto base de ejemplos.

- **Predicción⁵:** La predicción trata de adelantarse a los acontecimientos futuros, a partir de la información actual se predice cuál es lo más probable que suceda. La predicción es una forma de clasificación en el futuro. Las RNA se pueden entrenar en observaciones para predecir (con una cierta probabilidad), por ejemplo, quien es más probable que sea un buen jugador.
- **Optimización:** La optimización intenta dar respuesta a un tipo general de problemas de la forma: $\max(\min f(x))$ Donde $x = (x_1, \dots, x_n)$ es un vector y representa variables de decisión, $f(x)$ es llamada función objetivo y representa o mide la calidad de las decisiones (usualmente números enteros o reales) y Ω es el conjunto de decisiones factibles o restricciones del problema.

Algunas veces es posible expresar el conjunto de restricciones Ω como solución de un sistema de igualdades o desigualdades.

$$g(x_1, \dots, x_n) \leq 0, \quad h(x_1, \dots, x_n) = 0$$

Un problema de optimización trata entonces de tomar una decisión óptima para maximizar (ganancias, velocidad, eficiencia, etc.) o minimizar (costos, tiempo, riesgo, error, etc.) un criterio determinado. Las restricciones significan que no cualquier decisión es posible.

Por ejemplo la siguiente tabla ilustra el uso de algunas redes neuronales conocidas

Aplicación	Modelo de Red			
	Back Propagation	Hopfield	Maquina de Boltzman	Kohonen SOM
Clasificación	*	*	*	*
Procesamiento de Imágenes	*			*

⁵ Al parecer esta es la función principal del sistema nervioso, según (Llinás, 2003) el sistema nervioso aparece el momento que hay movimiento, es decir el momento en que es necesario la predicción ya que si hay movimiento el ambiente cambia. Tomemos el ejemplo de levantar un peso cualquiera, digamos una caja de leche. Nuestro cerebro realiza todos los cálculos para que nuestro brazo aplique la fuerza correcta y poder levantar la caja de leche. Qué tal si esta caja, no contuviese leche sino una sustancia más densa. Si estamos convencidos que es leche, aplicaremos la fuerza para este peso y no podremos levantarla. Si queremos poder levantarla necesitamos aplicar una fuerza mayor. Al realizar esta actividad nuestro cerebro realiza una predicción del peso.

Decisión de mercado	*		*	*
Optimización		*	*	*

Tabla 1-5 Tipos de aplicaciones de las redes neuronales

A continuación se presenta algunas aplicaciones de las RNA, tomado de Neural Network Toolbox para usos con Matlab

Biología y Medicina

- Aprender más acerca del cerebro y otros sistemas
- Obtención Modelos de la retina
- Análisis de ECG y EEG
- Diseño de prótesis
- Optimización de los tiempos de trasplante

Empresa

- Evaluación de formaciones geológicas y petrolíferos
- Explotación de Bases de Datos
- Optimización de plazas y horarios en líneas de vuelo
- Reconocimiento de caracteres escritos
- Análisis de diseño de productos químicos
- Identificación de partículas en tiempo real
- Sistemas de guías automática de automóviles

Finanzas

- Predicción de la evolución de precios
- Análisis de uso de líneas de créditos
- Identificación de falsificaciones
- Interpretación de firmas

Entretenimiento

- Animación
- Efectos especiales

- Predicción de mercado

1.5 Algunas Preguntas

- ¿A qué se refiere la Inteligencia Artificial, qué opina del término inteligencia computacional, aclara mejor el concepto o es más restrictivo?
- ¿Cree posible que algún día las máquinas lleguen a imitar las capacidades de la inteligencia humana, en qué se diferenciaran, cuáles son las razones?
- ¿Qué se plantea la computación neuronal y qué implicación tiene esto en el entendimiento del cerebro humano?
- ¿Cuál es el proceso básico de funcionamiento de la neurona?
- ¿Cuáles son los elementos de una red neuronal?
- ¿Qué simula la función de activación?
- Explique por qué es necesario el umbral.Cuál es su efecto y cuáles serán las consecuencias de no tener un umbral
- Explique cuál es el efecto del cambio de los pesos sobre el hiperplano de separación
- Explique cuál es el efecto de de cambiar el umbral sobre el hiperplano que forma las fronteras de decisión
- ¿Explique el aprendizaje supervisado, no supervisado y reforzado mediante un ejemplo de la vida real, como se adapta este problema, u otro, a un computador?
- ¿Qué caracterizan a los algoritmos por corrección de error?
- ¿Qué implica que en el aprendizaje no supervisado se compare las entradas con los pesos de la red, mientras que en el supervisado las salida del sistema con la salida conocida?
- En función de esta forma de ajustar los pesos de las redes de aprendizaje supervisado y no supervisado, ¿qué tipo de aplicaciones o problemas esperaría resolver con cada sistema?

1.6 Referencias Bibliográficas

- Butazzo Giorgio (2001) Artificial Consciosness: Utopia or real Possibility, Computer, July 2001
- Donald Tvetter (1998) The Pattern recognition Basis of Artificial Intelligence, IEEE Computer Society
- Frénac Yves (1994), Las mil y una vidas de la sinapsis de Hebb, Mundo Científico, N° 150
- Hilera y Martínez (1995), Redes Neuronales Artificiales, Addison-Wesley
- Hopgood A (2003). Artificial Intelligence: Hype or Reality, IEEE Computer Society, 03/2003
- Llinás, R. R. (2003). El cerebro y el mito del yo. Bogota: Editorial Norma.
- Krose Ben and van der Smagt (1996) An Introduction to Neural Networks,
- Maeschal and Michael S.C. Thomas (2006), How Computational Models help Explain the Origins of reasoning, IEEE Computational Intelligence
- Nadal Jean Pierre (1988) L'intelligence des réseaux de neurones, Pour La Science N° 254
- Vasant Honavar, (1996) - Iowa State University <http://www.cs.iastate.edu>
- Matthew J. Simoneau and Jane Price, Powerful new algorithms to explore, classify, and identify patterns in data, NEWS & NOTES MATLAB
- Werbos Paul, Neural Networks that actually work in diagnostics, prediction & control: Common Misconceptions Vs real-world success Tutorial IEEE Computational Intelligence Society.
- van Eck Nees Jan, Waltman Ludo, van den Berg Jan, Kaymak Uzay. (2006). Visualizing the Computational Intelligence Field. IEEE Computattional Intelligence, 6-10.