

OPENCOURSEWARE
REDES DE NEURONAS ARTIFICIALES
Inés M. Galván – José M. Valls

Tema 1

Introducción a las Redes de Neuronas

Introducción a las Redes de Neuronas

- Introducción
- Fundamentos biológicos
- Modelo Computacional
- Aprendizaje y Generalización
- Historia de las Redes Neuronales

Introducción

- Objeto de la IA: construcción de sistemas inteligentes
- Sistema Inteligente: dispositivo físico o lógico capaz de realizar tareas como: toma de decisiones, resolución de problemas, aprendizaje...
- Áreas
 - Simbólica – Top-down
 - Subsimbólica – Botton-up
- Las Redes de Neuronas se encuadran en la IA subsimbólica. Intentan emular los sistemas neuronales biológicos

Introducción

- Durante varias décadas los científicos han perseguido la construcción de algoritmos capaces de procesar información al igual que el cerebro humano
- En la actualidad existe un gran número de arquitecturas diferentes de redes de neuronas artificiales.
- Se caracterizan porque gozan de propiedades como la capacidad de aprendizaje a partir de ejemplos
- Se aplican en numerosas áreas como:
 - reconocimiento de patrones (imagen, voz, caracteres)
 - compresión y análisis de datos
 - robótica
 - medicina
 - predicción de series temporales
 -

Introducción

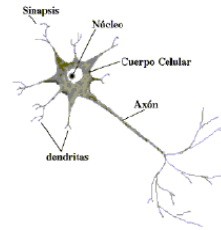
- Las redes de neuronas artificiales se enmarcan dentro de las técnicas de aprendizaje automático
- Las redes de neuronas se utilizan para abordar problemas de:
 - Aproximación, Predicción, Clasificación, Agrupación
- Otras técnicas de aprendizaje automático que pueden también utilizarse para estas familias de problemas son:
 - Máquinas de Vectores de Soporte (Vapnik, 1995)
 - Gradient Boosting (Friedman, 2002)
 - Random Forest (Breiman, 2001)
 - K-medias (agrupación) (MacQueen , 1967)
 -

Introducción

- De manera general:
 - Ventajas de las RNA
 - Capacidad de aprendizaje a partir de ejemplos
 - Tolerancia al ruido en los ejemplos
 - Modelos fáciles de utilizar
 - Diversidad de las áreas de aplicación
 - Problemas específicos de las RNA
 - Tiempo de aprendizaje indeterminado
 - Diseño *en ocasiones* poco sistemático: prueba y error

Fundamentos biológicos

- Elemento fundamental: la neurona
- Reciben los estímulos por las dendritas
- La señal de activación se emiten a través del axón
- Funciones de las neuronas
 - Recogida de información que llega en forma de impulsos
 - de otras neuronas
 - de receptores
 - Transmisión codificada en forma de frecuencia de impulsos
 - Transmisión
 - Neuronas subsiguientes
 - Células efectoras



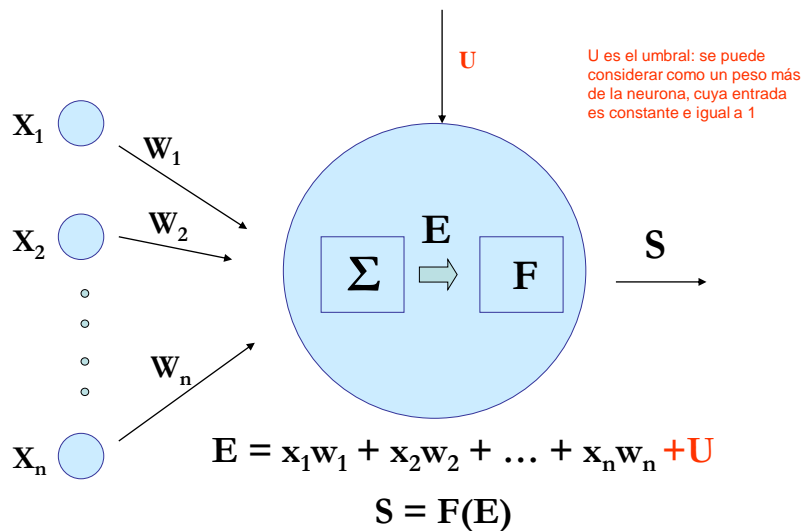
Fundamentos biológicos

- Las neuronas no son elementos lineales
- Si se supera un umbral, se produce una señal de activación
- El sistema nervioso funciona como una enorme malla que propaga señales electroquímicas de unas células a otras, modificando sucesivamente la concentración de iones de las sinapsis.
- Robustez y tolerancia a fallos
- Adaptación y plasticidad
- Manejar información: inconsistente, difusa y con ruido
- Alta dimensión: aprox. 10^{11} neuronas, cada una conectada con entre 1.000 y 10.000 células

Modelo computacional. La neurona artificial

- Cada neurona está caracterizada por un estado interno denominado **nivel de activación**
 - Conjunto de estados posibles de la neurona. Ejemplos:
 - $S=\{0,1\}$, 0: desactivada, 1: activada
 - $S=[0,1]$, intervalo continuo de valores
- **Función de activación**: permite cambiar el nivel de activación a partir de las señales de entrada
- Las señales de entrada se combinan entre sí, generando la entrada total

Modelo computacional. La neurona artificial



Modelo computacional. La neurona artificial

$$E = x_1 w_1 + x_2 w_2 + \dots + x_n w_n + U$$

- Se puede definir en forma vectorial como sigue:

$$E = (x_1, x_2, \dots, x_n, 1)^T (w_1, w_2, \dots, w_n, U)$$

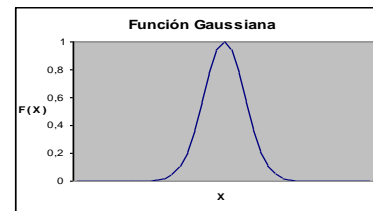
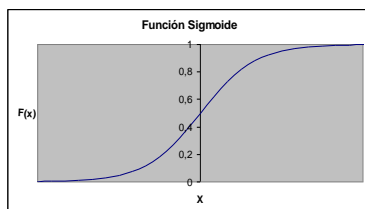
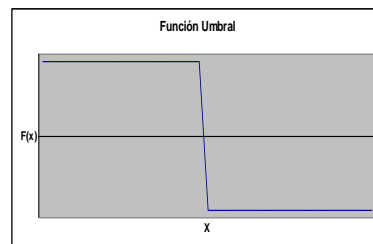
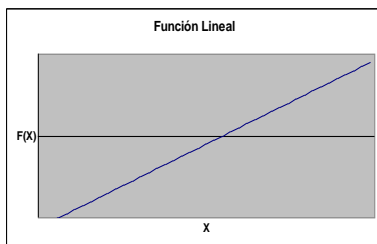
- Función de salida:

- Lineal.- $S = K \cdot E$ con K constante.
- Umbral.- $S = 1$ si $E \geq \theta$, $S = 0$ si $E < \theta$. Siendo θ el umbral constante.
- Cualquier función.- $S = F(E)$. Siendo F una función cualquiera.

- Función Sigmoidal $F(x) = 1/(1+e^{-x})$
- Función Gaussiana $F(x) = e^{-x^2/2}$

Modelo computacional. La neurona artificial

Funciones de Activación



Modelo computacional. RNA

Una red de neuronas artificiales viene determinada por

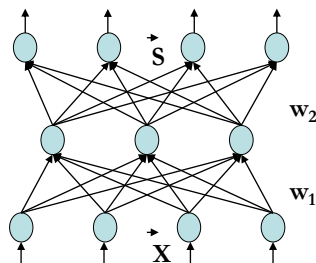
- Conjunto de unidades de proceso o neuronas artificiales
- Conexiones entre las unidades
- Regla de propagación, para propagar las entradas hacia la salida de la red
- Regla de aprendizaje para modificar los pesos de las conexiones

Se necesita un conjunto de ejemplos (muestras, datos, patrones) que representen el problema, a partir del cual se realiza el aprendizaje

Modelo computacional. RNA

- Arquitectura de la red: Distribución de las neuronas y conexión entre ellas
- La estructura básica: Red multicapa
 - Capa de entrada, Capas ocultas y Capa de salida
 - Propagación de la señal hacia delante
 - Los pesos se ajustan en la fase de aprendizaje

Ejemplo



Aprendizaje

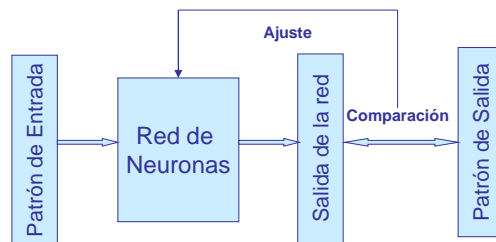
- Aprendizaje de una red de neuronas: determinación de los pesos y umbrales para la resolución eficiente de un problema
- Se determinan a partir de un conjunto de ejemplos o muestras, llamado conjunto de entrenamiento
- La capacidad de aprendizaje está ligada a los ejemplos del conjunto de entrenamiento
 - Significativo (cantidad)
 - Representativo (variedad)

Aprendizaje

- El proceso de aprendizaje consiste en
 - Introducir paulatinamente todos los ejemplos de aprendizaje
 - Modificar los pesos siguiendo una ley o ecuación de aprendizaje
 - Cuando se han introducido todos, se comprueba si se cumple cierto criterio de convergencia. Si no se cumple, se repite el proceso.
- La modificación de los pesos se puede hacer
 - después de introducir cada patrón (más utilizado)
 - después de introducir todos los patrones
- La finalización del aprendizaje se hace en función de un criterio de convergencia determinado por
 - Número fijo de ciclos de aprendizaje o iteraciones
 - Error menor que un valor dado
 - Los pesos dejan de sufrir modificaciones

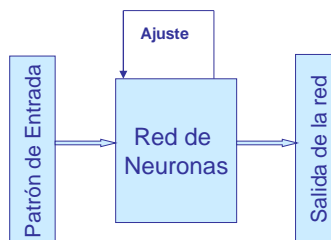
Aprendizaje

- Aprendizaje supervisado
 - Los ejemplos de aprendizaje contienen información relativa a la solución del problema
 - Utiliza profesor externo
 - Maneja información del error cometido por la red
 - La utiliza de forma realimentada para guiar el aprendizaje de la red



Aprendizaje

- Aprendizaje no supervisado
 - No utiliza información externa
 - Reajuste automático de los pesos
 - Auto-organiza la información
 - La red trata de determinar características de los datos: redundancias, regularidades, rasgos significativos



Aprendizaje

- Aprendizaje por refuerzo (Kaelbling et al, 1996)
 - Utiliza información externa de control
 - No existe medida de error, solo si lo ha habido

Generalización

- **Generalización es la capacidad de la red de responder adecuadamente ante entradas no conocidas (no utilizadas para el aprendizaje)**
- El conjunto de datos disponibles se divide en subconjuntos de
 - Entrenamiento
 - Usado para ajustar el valor de los pesos
 - Validación (no siempre se utiliza este conjunto)
 - Usado para determinar los parámetros de la red y del aprendizaje
 - Test
 - Usado para medir la capacidad de generalización de la red
- Los conjunto de validación y test deben ser:
 - » Independientes del aprendizaje
 - » Significativos y representativos
- Problema de sobreajuste (overfitting): Cuando la red se ha sobreajustado a los ejemplos de entrenamiento y no generaliza ante casos nuevos

Tipos de Redes de Neuronas

- Redes feedforward y supervisadas
Conexiones hacia adelante y aprendizaje supervisado
Perceptron simple, Adaline, Perceptron Multicapa, Redes de Base Radial
- Redes no supervisadas
Kohonen, ART
- Redes recurrentes
Conexiones en todas las direcciones
Pueden ser no supervisadas (red de Hopfield) y supervisadas
Las redes parcialmente recurrentes poseen solo unas pocas conexiones recurrentes. Son supervisadas: Red de Jordan, Red de Elman

Historia de las redes neuronales

- Los primeros estudios fueron realizados por McCulloch y Pitts en 1949. Mostraron la habilidad de un grupo de neuronas conectadas para llevar a cabo la implementación de ciertas funciones lógicas. No se introdujo aún el concepto de aprendizaje
- Donald Hebb.- Realiza desarrollos matemáticos de aprendizaje. **Aprendizaje hebbiano** (1949)
- Durante la década de los 50 hubo un considerable crecimiento en este campo.
 - Rosenblatt en 1958 introdujo la primera arquitectura de red de neuronas artificial con capacidad de aprendizaje: **El perceptrón simple**.
 - Bernard Widrow en 1959 introduce el: **ADALINE**, parecida al perceptron pero introduce en el modelo el ajuste de los pesos en base el error cometido por la red.
- **Stephen Grossberg**
 - Relaciones entre fenómenos psicológicos y biológicos
 - Análisis matemático y elaboración de nuevos paradigmas de acceso directo, tiempo real, aprendizaje continuo

Historia de las redes neuronales

- Teuvo Kohonen (70's)
 - RN con conexiones aleatorias, memorias asociativas y matrices de correlación
 - **Mapas autoasociativos (SOM). Learning Vector Quantization (LVQ)**
- Rumelhart propuso el **Perceptron Multicapa** (Algoritmo de BackPropagation) en 1986. Solventa los inconvenientes del perceptron simple en lo que respecta a problemas no lineales
- Las **Redes de Base Radial** fueron introducidas por Powell en 1985 para resolver problemas de interpolación en espacios multidimensionales. Se basan en la idea de aprendizaje local, diferente al concepto de aprendizaje global del perceptron multicapa.
- Los primeros estudios sobre redes recurrentes los realizó John Hopfield (80's). Las conexiones pueden aparecer en todas las direcciones, de la capa de entrada a la capa de salida y viceversa.
 - **Red autoasociativa**. Estudios sobre complejidad y convergencia
 - **Red de Elman** (1988) y **Red de Jordan** (1986). Redes parcialmente recurrentes

Referencias

(Vapnik, 1995): Vladimir N. Vapnik. The Nature of Statistical Learning Theory. Springer-VerlagNew York, Inc., New York, NY, USA, 1995.

(Friedman, 2002): Jerome H. Friedman. Stochastic gradient boosting. Computational Statistics Data Analysis, 38(4):367- 378, 2002.

(Breiman , 2001): Breiman, L. Random forests. Machine learning, 45(1), 5-32, 2001.

(MacQueen , 1967): MacQueen, J. Some methods for classification and analysis of multivariate observations. In *Proceedings of the fifth Berkeley symposium on mathematical statistics and probability*, Vol. 1, No. 14, pp. 281-297, 1967

(Kaelbling et al, 1996): Kaelbling, L. P., Littman, M. L., & Moore, A. W. Reinforcement learning: A survey. *Journal of artificial intelligence research*, 4, 237-285, 1996