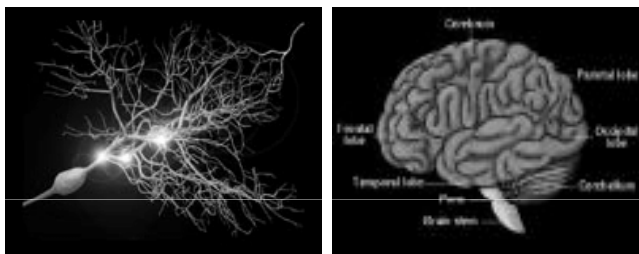


Introducción



Héctor Allende

hallende@inf.utfsm.cl

www.inf.utfsm.cl/~hallende

Bibliografía

Libros

- Arbib M., “*Handbook of Brain Theory and Artificial Neural Network*”, The MIT Press, 1998.
- Richard Golden, “*Mathematical Methods for Neural Network Analysis and Design*”, MIT Press, 1996.
- Bishop C., “*Pattern Recognition and Machine Learning*”, Springer 2006.
- Ripley B. D., “*Pattern recognition and Neural Network*”. Cambridge University Press. 1996.
- **Revistas**
- Journal of Neuro-computing
- IEEE trans. Neural Networks
- Journal Neural Networks
- Journal of Neuro computational

2009

Héctor Allende O

Neurona Unidad Constituyente del Cerebro

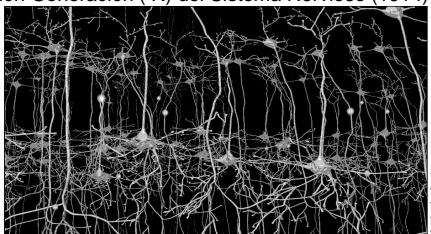
• Santiago Ramón Cajál (1852- 1934)

Descubre la neurona (1888)

Publicación Textura del sistema nervioso del hombre (1902)

Premio Nobel en Medicina (1906)

Publicación Generación (R) del Sistema Nervioso (1914)



Héctor Allende O

2009

Computational Intelligence

- Combinación:
 - Ciencias Cognitivas y lógica
 - Ciencias de la Computación
 - Neurociencia (Fisiología neuronal)
 - Evolución (Comportamiento)
 - Estadística

Creación de Máquinas que puedan aprender de los datos

2009

Héctor Allende O

Introducción

Redes Neuronales Artificiales:

- Modelo Algoritmico inspirado **inicialmente** en el comportamiento de las neuronas biológicas y el cerebro humano

Computational neuroscience:

- Simulación Computacional, como herramienta de modelado para ayudar a la comprensión del funcionamiento del cerebro humano

2009

Héctor Allende O

Corteza Cerebral Humana

- Aproximadamente 10^{11} neuronas
- 1000 a 10.000 Synapsis por neurona
- Comunicación tren de impulsos electro-químicos (mensaje modulado)
- Proceso Cognitivo Humano
 - tiempo (milisegundos)
 - Operación (Masivamente Paralela)

2009

Héctor Allende O

Breve Historia

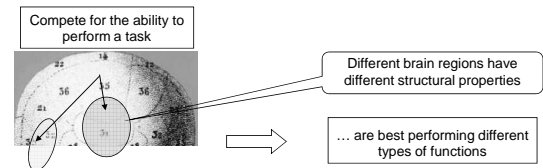
- 1943 W.McCulloch, W. Pitts: Modelo ANN (El Perceptrón)
- 1949 D. Hebb: Neuronas Asociativa
- 1959 Bernard Widrow (SU): Filtro Adaline (Adaptive Linear Neuron) y el Madaline (Multiple Adaline) Basado en (ANN)
- 1962 F. Rosenblatt : El perceptron
- 1969 Minsky y Papert: El Perceptrón (limitaciones).
- 1982 J. Hopfield: Memoria Asociativa "Redes de Hopfield".
- 1986 Rumulhart, Hinton y Williams : redescubren el BPL algoritmo de "back-propagation learning" (Paul Werbor, 1974)
- 1982 Teuvo Kohonen . Redes auto-organizativas (SOM)
- 1989 K. Hornik, M. Stinchcombe, H. White: Multi-FANN y Aproximación Universal
- 1990.....

2009

Héctor Allende O

Ciencias Cognitivas (cerebro)

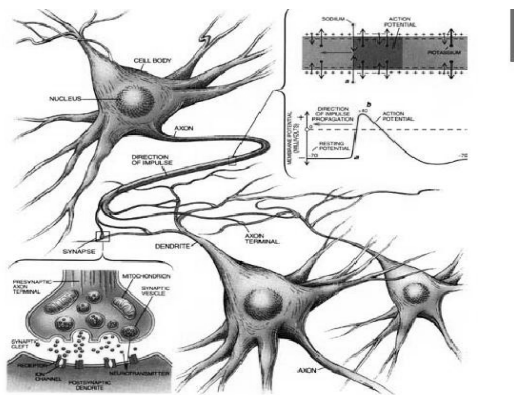
- El cerebro humano es una colección de módulos donde cada uno se especializa en el desempeño de una función (no forma exclusiva)
- Los trabajos de investigación se sustentan en la hipótesis :



2009

Héctor Allende O

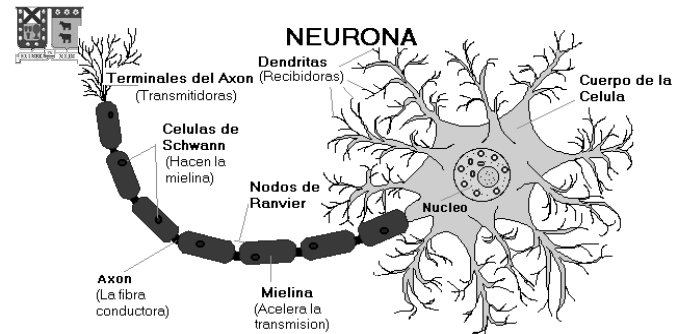
Modelo de Neurona



Arbib and Ito: CS564 - Brain Theory and Artificial Intelligence, USC, Fall 2000. Lecture 2. Networks of Neurons

2009

Héctor Allende O



Soma:	Información + Plasma + Generación Señales
Dendritas:	Recepción Señales → Impulsos
Axón:	Transmisión de Señales
Sinapsis:	Interfaz Neuronal (Inhibitoria, Excitatoria)

2009

Héctor Allende O

Red Neuronal Artificial (ANN)

ANN: Es un sistema dinámico compuesto por redes paralelas y distribuidas de procesadores elementales, con una baja potencia de computo, con la capacidad de **aprender y almacenar "conocimiento"**.

- Arquitectura
- Interacción
- Función de activación

ANN: - Máquina de Aprendizaje

- Modelo Estadístico Semiparamétrico
- Modelo de Caja Negra ; E/S

2009

Héctor Allende O

Teorías de la cognición

- I Etapa (1943-1953) : Cibernética (ciencia de la mente) Trabajo seminal McCulloch and Pitts " Un cálculo inmanente de la actividad nerviosa" Arquitectura de John von Neumann
- II Etapa (1956-) : Cognición se podía entender como la computación de representaciones simbólicas H. Simon ; N. Chomsky; M. Minsky
- III Etapa (1980-) : Conexionismo, emergencia, autoorganización, asociación, dinámica de red
- IV Etapa (1990): la enacción (hacer emerger) una alternativa ante la representación.

2009

Héctor Allende O

Teorías de la cognición

- Qué es la Cognición? Acción efectiva del acoplamiento estructural que enactúa (hace emerger) un mundo.
- Como funciona ? A través de una red de elementos interconectados capaces de cambios estructurales durante un tiempo ininterrumpido.
- Como saber si un sistema cognitivo funciona? Cuando se transforma en parte de un mundo de significación pre-existente o configura un mundo nuevo.

(Francisco Varela 1998)

2009

Héctor Allende O

Procesador Elemental

Unidad básica de procesamiento que posee múltiples entradas y una salida.

- Cada entrada x_i es ponderada por un factor de ponderación (peso) w_i y se calcula la suma ponderada de las entradas:

$$\sum_i w_i x_i = A(\mathbf{x})$$

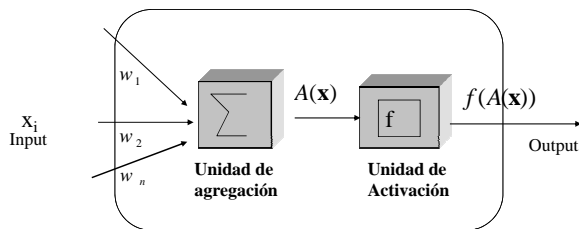
- Luego es aplicada una transformación mediante la función de activación :

$$Salida = f(A(\mathbf{x}))$$

2009

Héctor Allende O

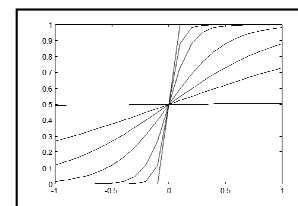
Procesador elemental.



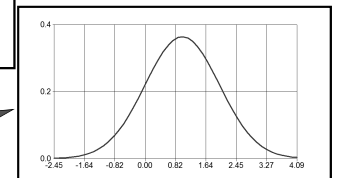
2009

Héctor Allende O

Tipos Funciones de Activación



Funciones Tipo Sigmoidal

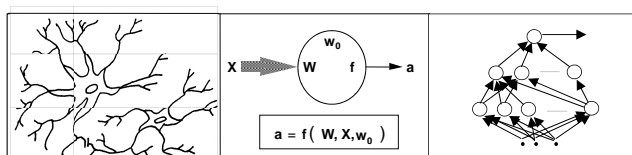


Funciones Base Radial

2009

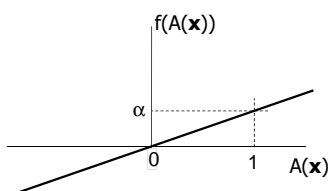
Héctor Allende O

Biology - Engineering



$$A(\mathbf{x}) = \sum_{i=1}^n x_i w_i - w_0$$

$$f(A(\mathbf{x})) = \alpha A(\mathbf{x})$$

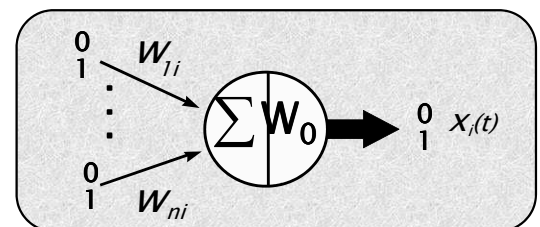


2009

Héctor Allende O

© cm

Mc Culloch & Pitts (1943)



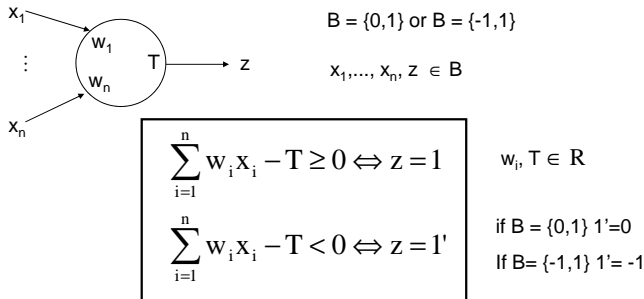
$$x_i(t) = \mathbb{1} \left[\sum_{j=1}^n w_{ij} x_j(t-1) - w_0 \right]$$

2009

Héctor Allende O

El perceptron

El "perceptron" procesador desarrollado F. Rosenblatt (1958),



Note $\sum_{i=1}^n w_i x_i - T = \sum_{i=0}^n w_i x_i$ with $x_0=1$ and $w_0 = -T$

2009

Héctor Allende O

© cm

Rosenblatt's Theorem

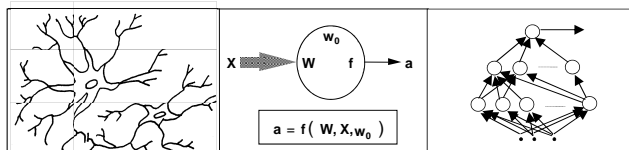
"A perceptron can learn what it can do"

2009

Héctor Allende O

© cm

Biology - Engineering

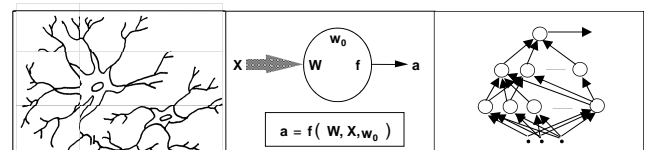


$A(\mathbf{x}) = \sum_{i=1}^n w_i x_i - w_0$
 $f(A(\mathbf{x})) = \text{step}(A(\mathbf{x}))$

2009

Héctor Allende O

Biology - Engineering

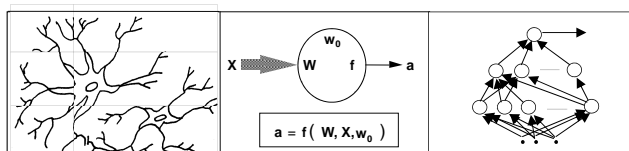


$A(\mathbf{x}) = \sum_{i=1}^n w_i x_i - w_0$
 $f(A(\mathbf{x})) = \text{sign}(A(\mathbf{x}))$

2009

Héctor Allende O

Biology - Engineering



$A(\mathbf{x}) = \sum_{i=1}^n x_i w_i - w_0$
 $f(A(\mathbf{x})) = \frac{1}{1 + \exp(-A(\mathbf{x}))}$

2009

Héctor Allende O

© cm

Funciones de activación

• Función de activación logística:

$$f(a) = \frac{1}{1 + e^{-ca}}$$

- Es monótonamente creciente para $c > 0$

$$f' \equiv \frac{df}{da} = cf(1-f) > 0$$

2009

Héctor Allende O

Neuronas como funciones

- Las neuronas transforman una entrada no acotada $A(t)$ en el tiempo t en una señal de salida acotada $f(A(t))$.
- La función de activación o función de señal: f

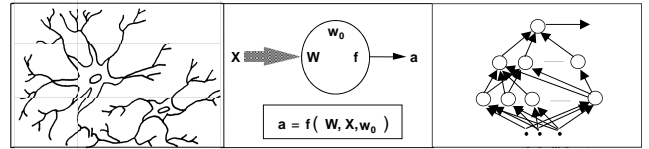
- Velocidad de la señal: $\frac{dA}{dt}$

$$\dot{f} = \frac{df}{dA} * \frac{dA}{dt}$$

2009

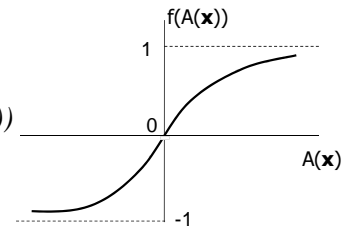
Héctor Allende O

Biology - Engineering



$$A(\mathbf{x}) = \sum_{i=1}^n x_i w_i - w_0$$

$$f(A(\mathbf{x})) = \tanh(A(\mathbf{x}))$$



2009

Héctor Allende O

Funciones de activación

- Tangente hiperbólica:

$$f(a) = \tanh(ca) = \frac{e^{ca} - e^{-ca}}{e^{ca} + e^{-ca}}$$

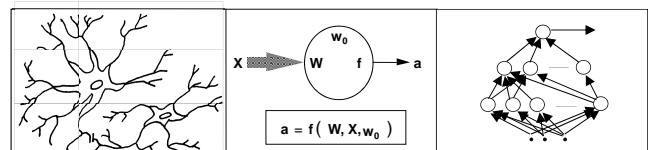
donde $c > 0$.

$$f' = c(1 - f^2) > 0$$

2009

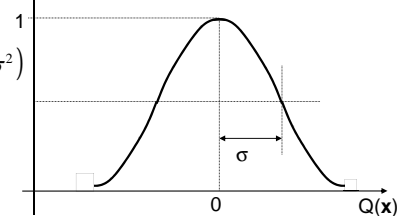
Héctor Allende O

Biology - Engineering



$$Q(\mathbf{x}) = \sum_{i=1}^n ((x_i w_i - w_0)^2 / 2\sigma^2)$$

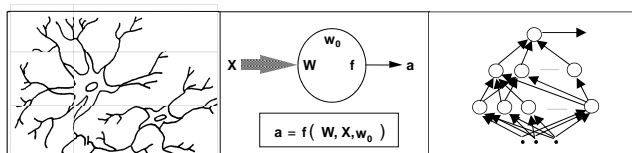
$$f(Q(\mathbf{x})) = \exp(-Q(\mathbf{x}))$$



2009

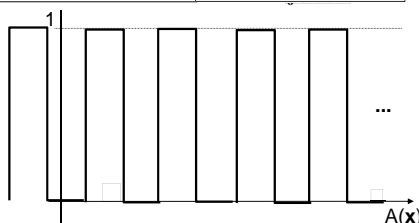
Héctor Allende O

Biology - Engineering



$$A(\mathbf{x}) = \sum_{i=1}^n x_i w_i - w_0$$

$$f(A(\mathbf{x})) = \dots$$

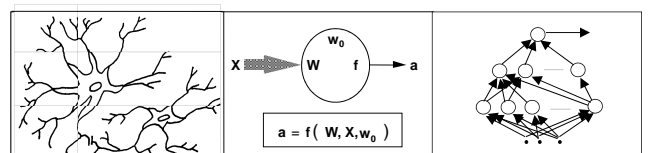


2009

Héctor Allende O

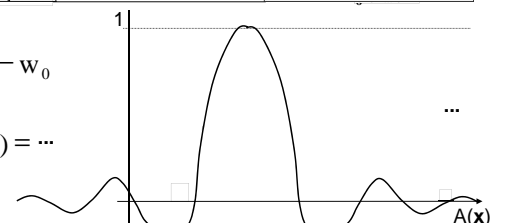
© cm

Biology - Engineering



$$A(\mathbf{x}) = \sum_{i=1}^n x_i w_i - w_0$$

$$f(A(\mathbf{x})) = \dots$$



2009

Héctor Allende O

© cm

Clasificación de las Redes Neuronales

Data flow:	Aprendizaje:
Feedforward Neural Networks	ANN with supervised learning
Recurrent Neural Networks	ANN with competitive learning
	ANN with unsupervised learning
Topología:	Estructura:
Regular	1 Layer
Irregular	Multilayer
Función de Transición	Algoritmo
Deterministas	PPL
Probabilistas	BPL

2009

Héctor Allende O

© cm

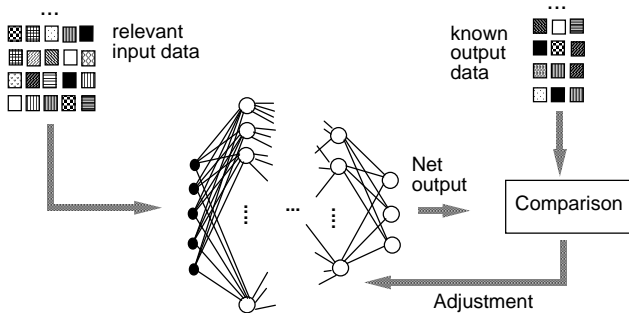
FANN (Feedforward)

- **ANN Feedforward:** Se construye colocando las neuronas en capas y conectando las salidas de una capa con las entradas de las neuronas de la capa siguiente.
- **Capas de una red:**
 - Capa de entrada Zona sensorial (S)
 - Capa de salida Zona de Respuesta (R)
 - Capas ocultas Zona de asociación (A)

2009

Héctor Allende O

Training a neural network

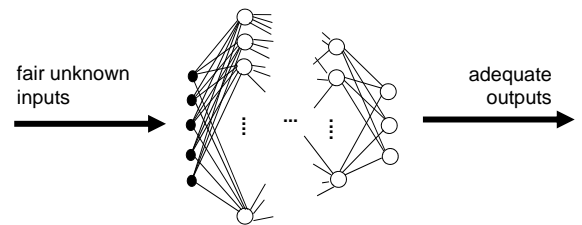


2009

Héctor Allende O

© cm

Generalization

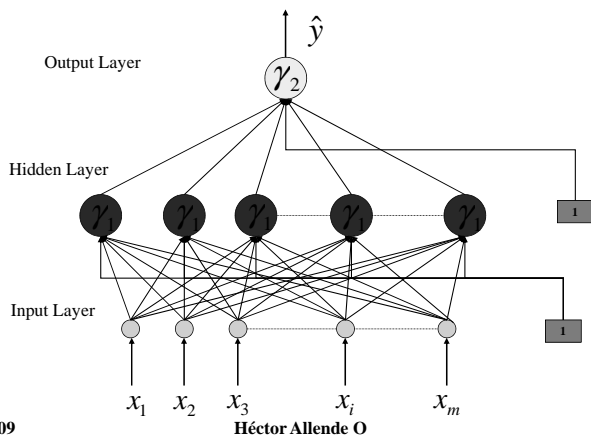


2009

Héctor Allende O

© cm

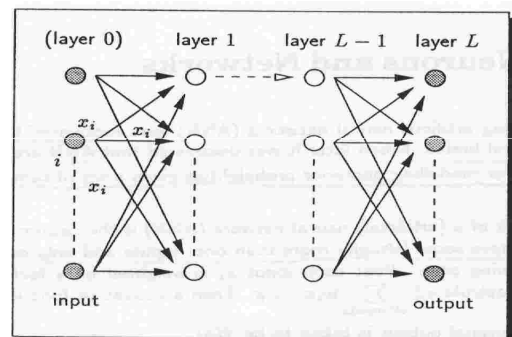
Feedforward Artificial Neural Network



2009

Héctor Allende O

Feedforward Neural Network



2009

Héctor Allende O

Feedforward Artificial Neural Network

•FANN:

$$g_{\lambda}(\underline{x}, \underline{w}) = \gamma_2 \left(\sum_{j=1}^{\lambda} w_j^{[2]} \gamma_1 \left(\sum_{i=1}^m w_{ij}^{[1]} x_i + w_{m+1,j}^{[1]} \right) + w_{\lambda+1}^{[2]} \right)$$

– Number of free parameters

$$d = (m+2)\lambda + 1$$

2009

Héctor Allende O

Aprendizaje Supervisado

- **Estimación de Parámetros:** Los pesos de la red $\{w_i\}$.
- **Aprendizaje o entrenamiento:** Es un algoritmo mediante el cual los pesos libres $\{w_i\}$ de una red son ajustados.
- **Conjunto de entrenamiento:** Conjunto de ejemplos (maestros) conocidos, es decir corresponden a determinados vectores de entrada asociados con vectores de salida: $\{(x_i, y_i)\}_{i=1,n}$.
- **Generalización:** Es la Capacidad de generalización de una red de aprender una distribución desconocida a partir de los datos y de algoritmos apropiados.

2009

Héctor Allende O

Redes Neuronales Feedforward

•Estimación de Parámetros:

$$\hat{w}_n^{LS} = \arg \min \{L_n(w) : w \in W \subseteq \mathbb{R}^d\}$$

$$L_n = \frac{1}{2n} \sum_{i=1}^n (y_i - g(\underline{x}^i, \underline{w}))^2$$

Estimador de mínimos cuadrados

2009

Héctor Allende O

Proceso de Aprendizaje

- Los pesos son ajustados mediante la (Regla Δ):

$$\begin{aligned} w_{lji}(t+1) &= w_{lji}(t) - \mu \frac{\partial E(W)}{\partial w_{lji}} \bigg|_{W(t)} \\ &= w_{lji}(t) - \mu \sum_{p=1}^P \frac{\partial E_p(W)}{\partial w_{lji}} \bigg|_{W(t)} \end{aligned}$$

donde $\mu > 0$ es la constante de aprendizaje.

En notación matricial:

$$W(t+1) = W(t) - \mu \nabla E$$

2009

Héctor Allende O

Estimador- M

- Un M-Estimador T_n^M es la estadística que minimiza la función de costo:

$$\sum_{i=1}^n \rho(X_i, T_n^M)$$

o de manera equivalente, T_n^M es la solución de la ecuación de estimación:

$$\sum_{i=1}^n \psi(X_i, T_n^M) = 0 \quad (\text{Huber 1964})$$

2009

Héctor Allende O

Modelo

- Sistema $y = \varphi(\underline{x})$
- Conjunto de entrenamiento $\mathcal{X} = \{\underline{x}^t, y_t\}_{t=1..n}$

- Aproximación FANN $y_t = g_{\lambda}(\underline{x}^t, \underline{w}) + r_t$

- Supuesto $r \sim \Phi(r/\sigma_r) \quad \underline{x} \sim K(\underline{x})$
 $\Rightarrow f(\underline{x}, y) = \frac{1}{\sigma_r} \phi\left(\frac{y - g_{\lambda}(\underline{x}, \hat{w}_n)}{\sigma_r}\right) k(\underline{x})$

2009

Héctor Allende O

M-Estimador FANN

- El M-estimador \hat{w}_n^M del parámetro w es

$$\hat{w}_n^M = \arg \min \{ RL_n(w) : w \in W \subseteq \mathbb{R}^d \}$$

donde $RL_n(w)$ es una función robusta de costo dada por

$$RL_n(w) = \frac{1}{n} \sum_{i=1}^n \rho \left(\frac{y_i - g_\lambda(x^i, w)}{\sigma_r} \right)$$

siendo

$$r_i = y_i - g_\lambda(x^i, w)$$

2009

Héctor Allende O

M-Estimador FANN

- El M-estimador \hat{w}^M puede ser definido explícitamente por la solución de la ecuación

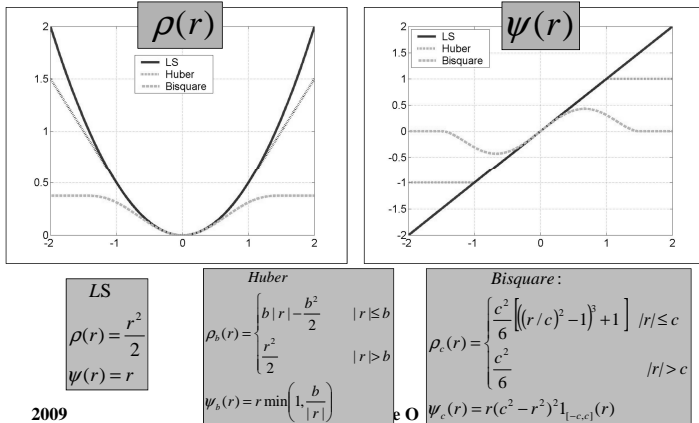
$$\text{donde } \sum_{i=1}^n \psi \left(\frac{y_i - g_\lambda(x^i, w)}{\sigma_r} \right) Dg_\lambda(x^i, w) = 0$$

$$\psi(r) = \frac{\partial \rho(r)}{\partial r}$$

2009

Héctor Allende O

Función robustificadora



2009

Neuronas y Redes Simples

- ANN Recurrente:** La salida de una neurona es la entrada de neuronas de capas anteriores (feedback): Redes de Jordan, Redes de Edelman
- Feedback lateral:** La salida de una neurona es la entrada de otra neurona en la misma capa: Redes autorganizativas (SOM)

2009

Héctor Allende O

Modelo de Turing

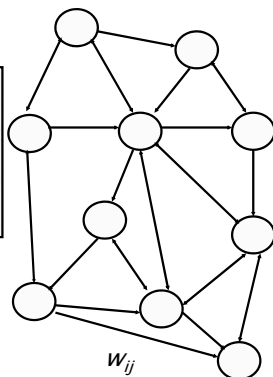
$$X_i(0) \in \{0,1\}$$

$$X_i(t) = \mathbb{1} \left[\sum_{j=1}^n w_{ij} X_j(t-1) - w_o \right]$$

$$\forall i=1, \dots, n$$

W (conectividad de la NN)

W_o = vector de umbrales



2009

Héctor Allende O

Aplicaciones de las ANN

- Telecomunicaciones
- Informática
- Minería
- Energía
- Finanzas
- Transporte
- Salud

2009

Héctor Allende O

Aplicaciones de las ANN

- Clasificación, Clustering
- Pre-procesamiento de datos
- Reconocimiento de patrones
- Aproximadores Universales
- Pronóstico de Series de Tiempo
- Optimización Combinatorial
- Control

2009

Héctor Allende O

Preguntas Abiertas

- Tamaño de las muestras
- Cuántas Neuronas
- Cuantas Capas
- Tipo de Arquitectura (Selección del Modelo ANN)
- Tipo de Aprendizaje
- Algoritmos de Aprendizaje
- ¿Cuándo usar ANN

Modelador

2009

Héctor Allende O

SNNS: Stuttgart neural networks simulator

Contiene:

- Diferentes tipos de redes neuronales
- Diferentes tipos de algoritmos de entrenamientos
- Diferentes representaciones de pesos y salidas
- Varios Ejemplos Demostrativos
- Buena documentación (in English)

SNNS is public domain

<http://www-ra.informatik.uni-tuebingen.de/SNNS/>

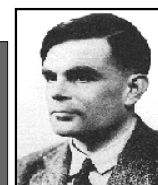
2009

Héctor Allende O

© cm

Test de Turing

“Un computador merece ser llamado inteligente si puede hacer creer a un ser humano que es otro ser humano”



Alan Turing



2009

Héctor Allende O