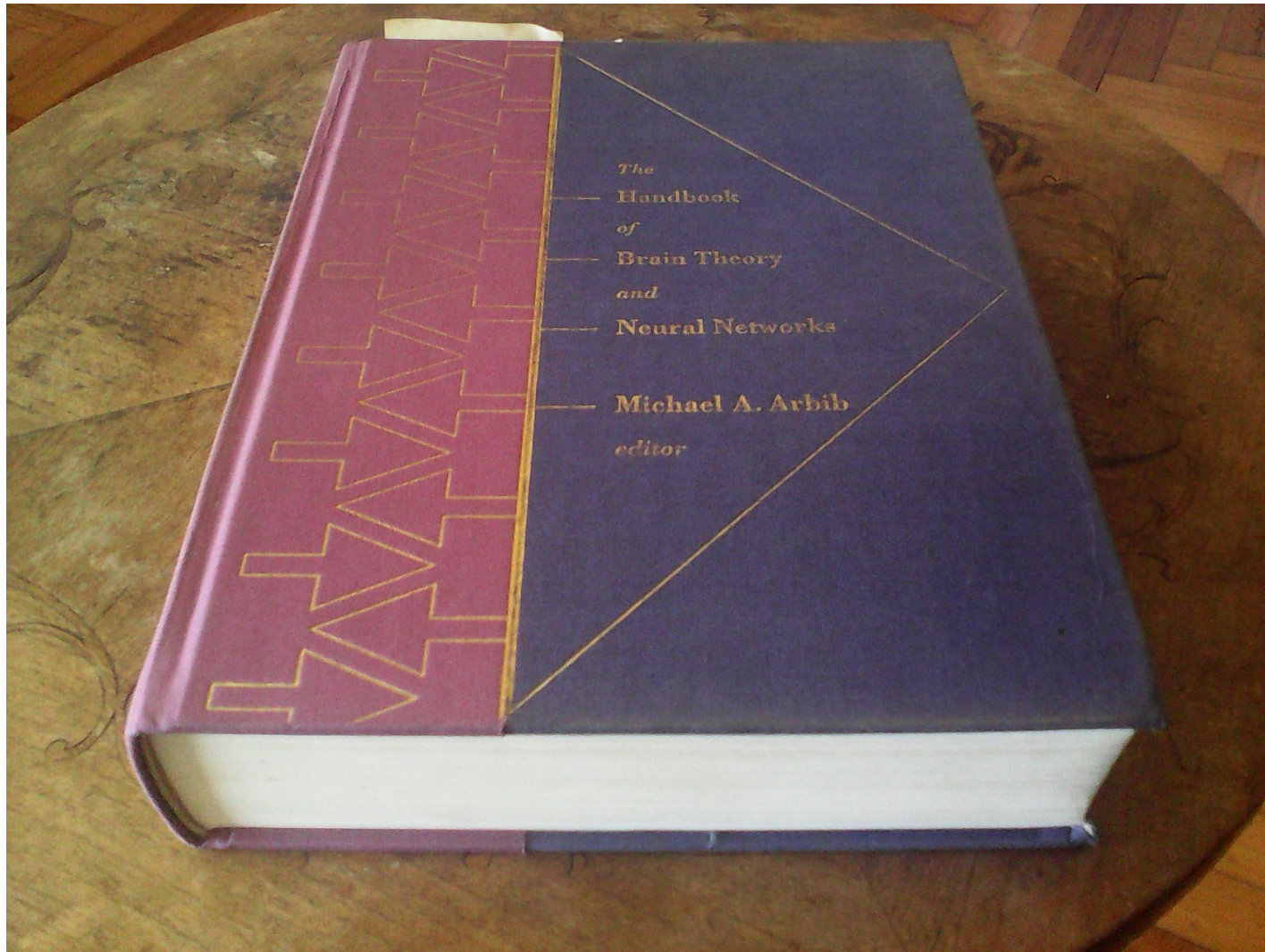


# Introducción a las Redes Neuronales Artificiales

# Preámbulo

- 1) Detesto usar aplicaciones para hacer presentaciones.
- 2) Detesto la información presentada en listas numeradas.
- 3) Detesto las pantallas de presentación llenas de texto.
- 4) Detesto tener que hablar sobre una pantalla de presentación llena de texto porque así los obligo a dividir su atención entre leer lo que dice en la pantalla y escuchar lo que estoy diciendo.
- 5) Y si... también detesto esta tipografía. >:-(

# The Big Book



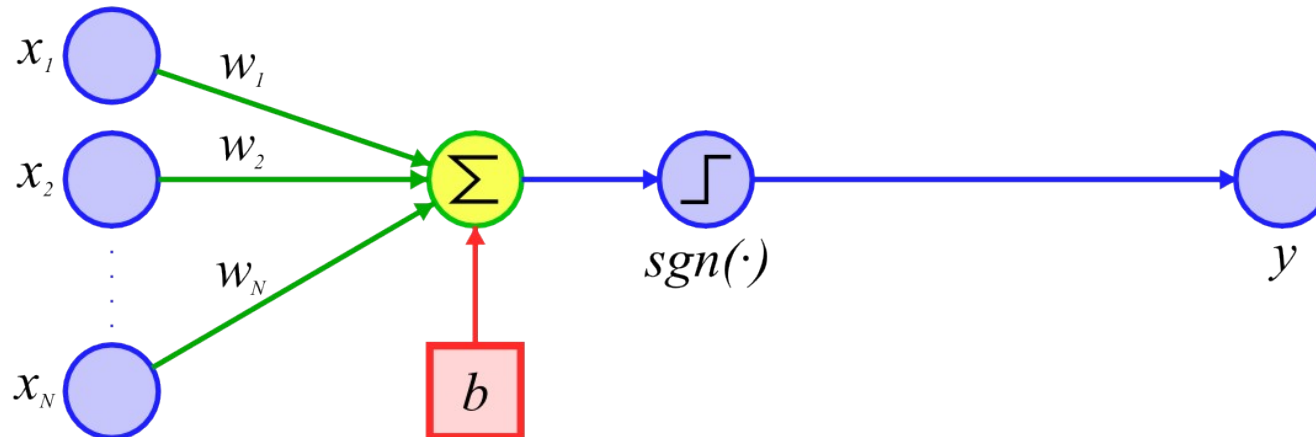
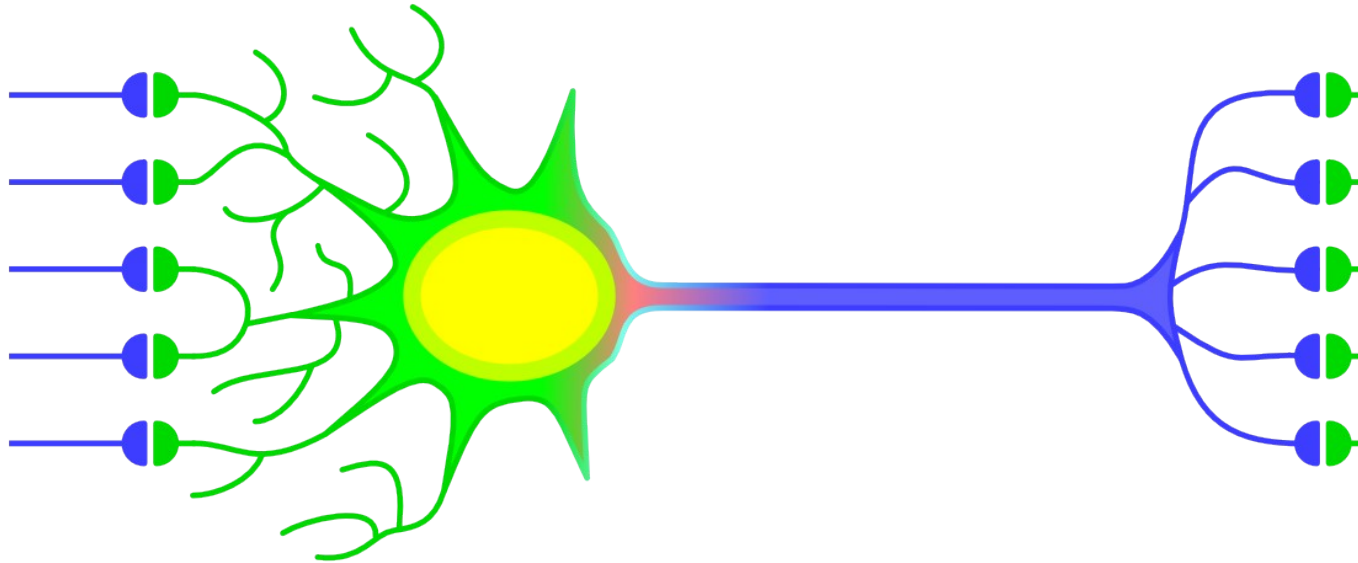
# Las Redes Neuronales son:

Un modelo matemático inspirado en un sistema biológico.

Unidades de procesamiento sencillas que muestran un comportamiento emergente con solo información local.

Un paradigma conexionista que puede operar en paralelo, es robusto ante fallas y capaz de aprendizaje adaptativo.

# La Neurona

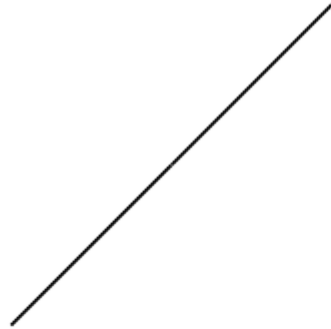


$$y = \text{sgn}\left(\sum_i x_i \cdot w_i - b\right)$$

# La Función de Transferencia

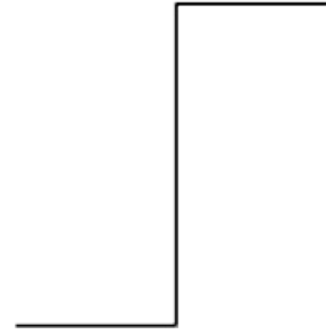
Lineal

$$s(x) = x$$



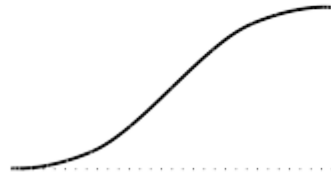
Escalón

$$s(x) = \text{sgn}(x)$$



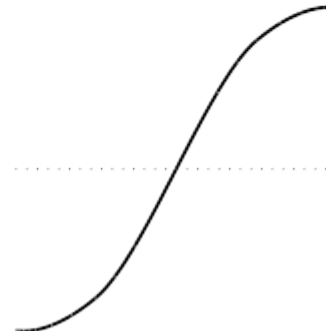
Sigmoidea  
Binaria

$$s(x) = \frac{1}{1 + e^{-x}}$$



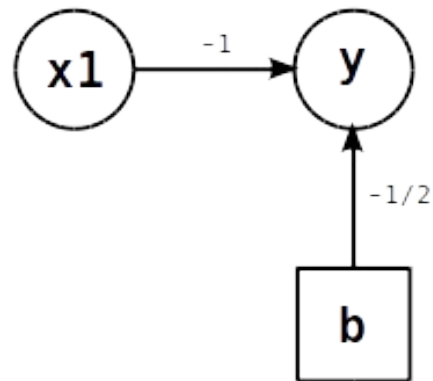
Sigmoidea  
Bipolar

$$s(x) = \tanh(x)$$

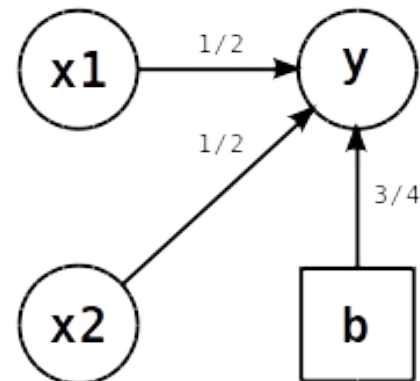


# Las Compuertas Lógicas

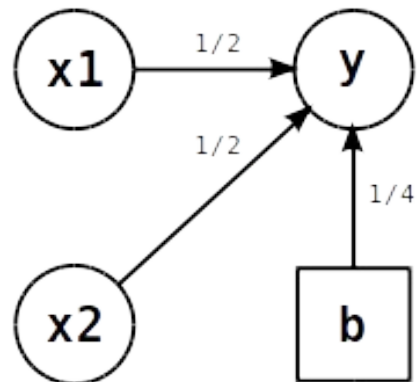
NOT



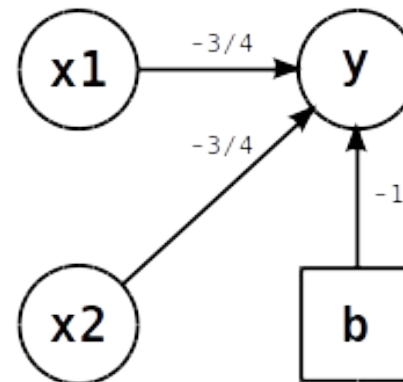
AND



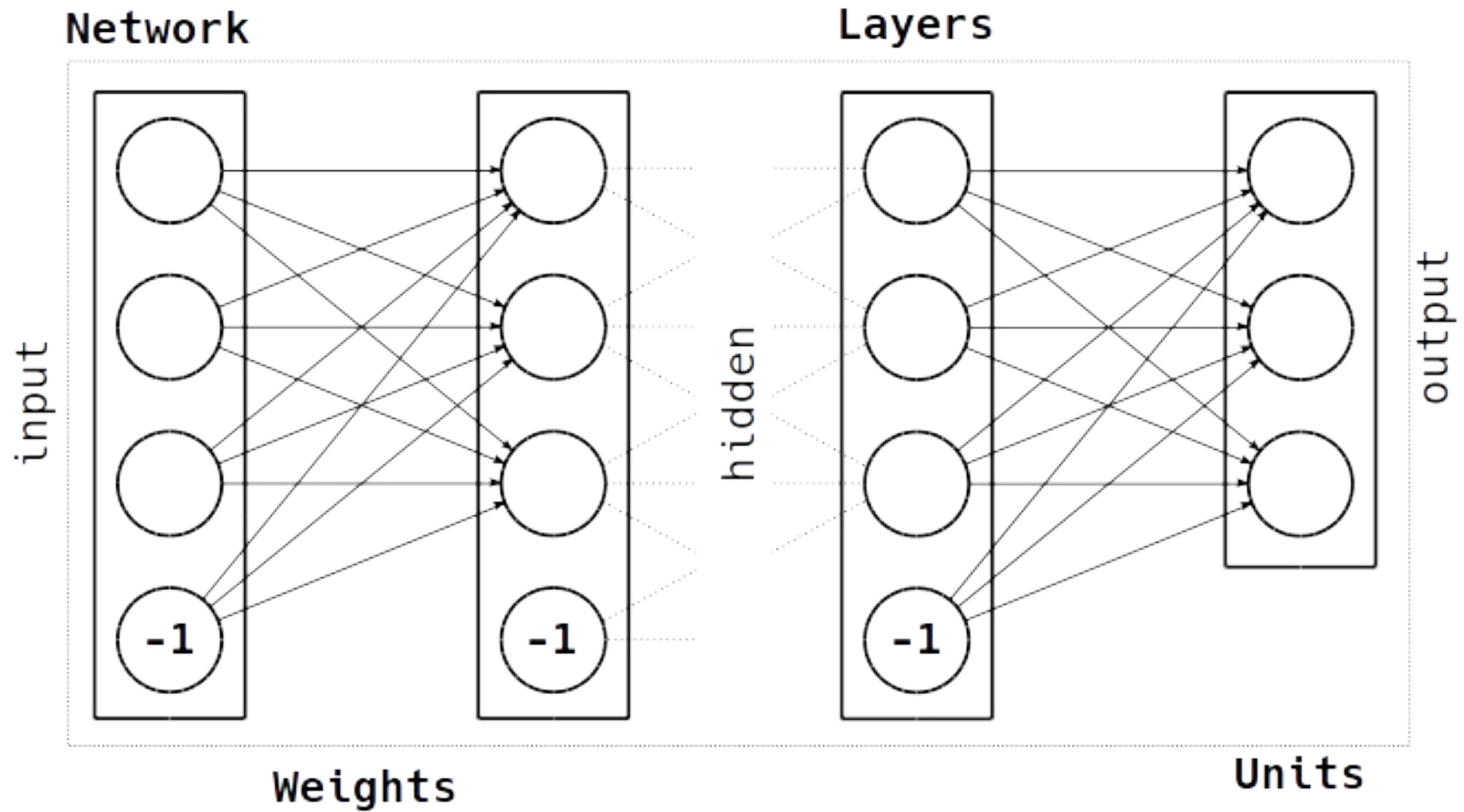
OR



NAND

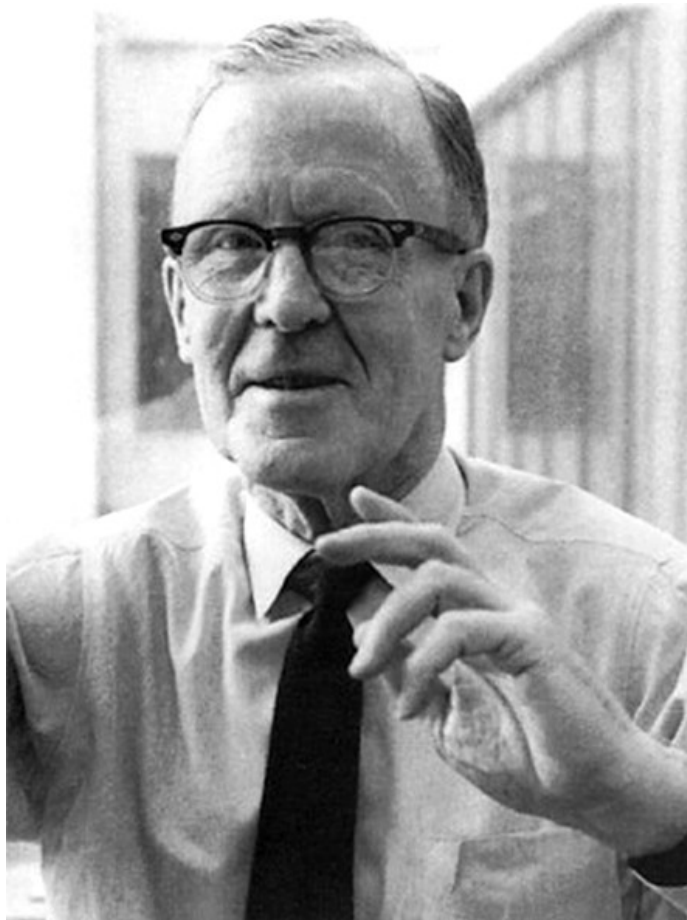


# La Red Neuronal





# El postulado de Hebb

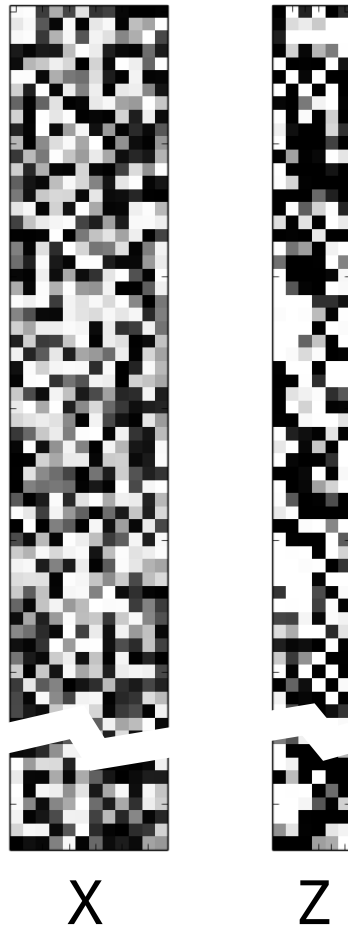


“Cuando el axón de una célula A está lo suficientemente cerca como para excitar a una célula B y repetidamente toma parte en la activación, ocurren procesos de crecimiento en una o ambas células de manera que tanto la eficiencia de la célula A, como la capacidad de excitación de la célula B son aumentadas.”

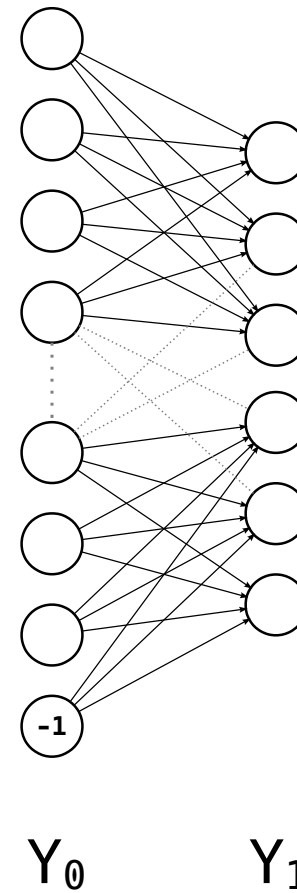
The Organization of Behavior  
Donald Hebb. (1949)

# El Aprendizaje Supervisado

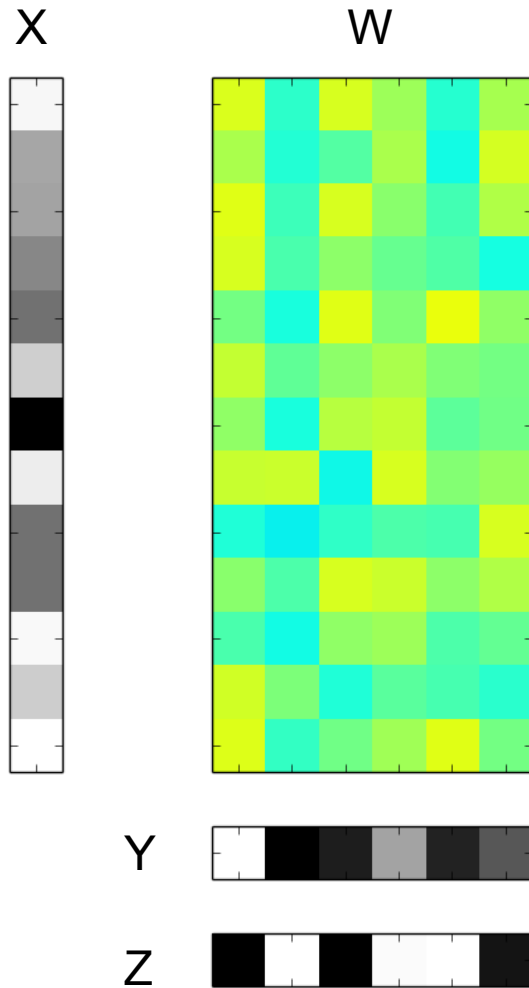
dataset



network



# La Función de Costo

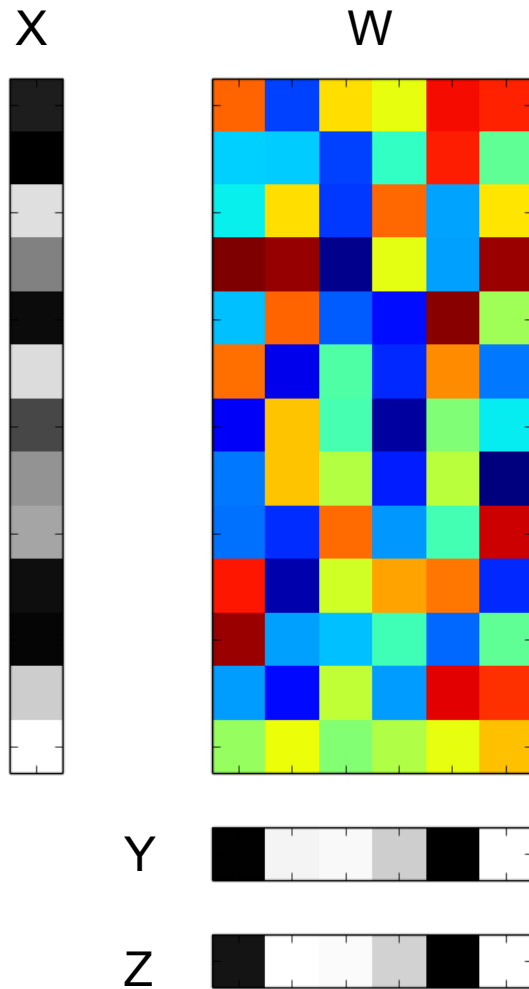


$$Y = s(X \cdot W)$$

$$E = \sum \|Z - Y\|^2$$

$$E = \sum \|Z - s(X \cdot W)\|^2$$

# La Regla Delta



$$Y = s(X \cdot W)$$

$$E = \sum \|Z - Y\|^2$$

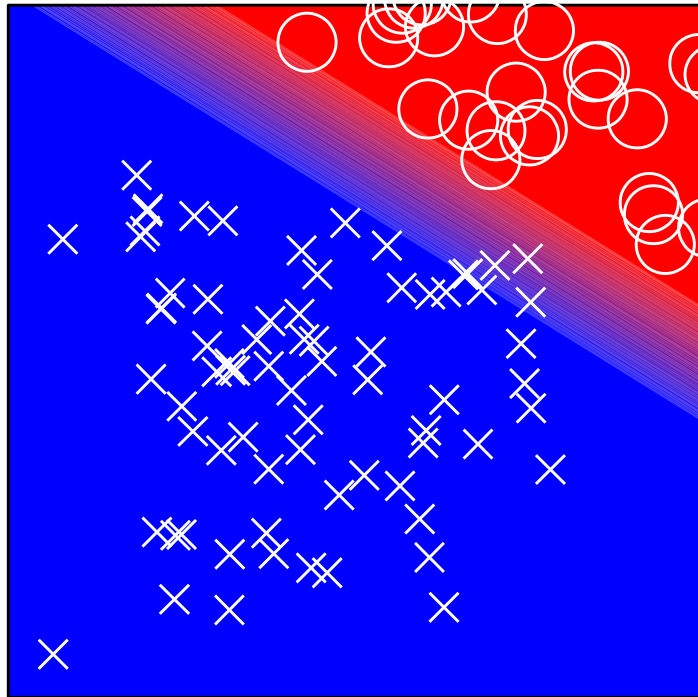
$$E = \sum \|Z - s(X \cdot W)\|^2$$

$$W^{[t+1]} = W^{[t]} + \Delta W^{[t]}$$

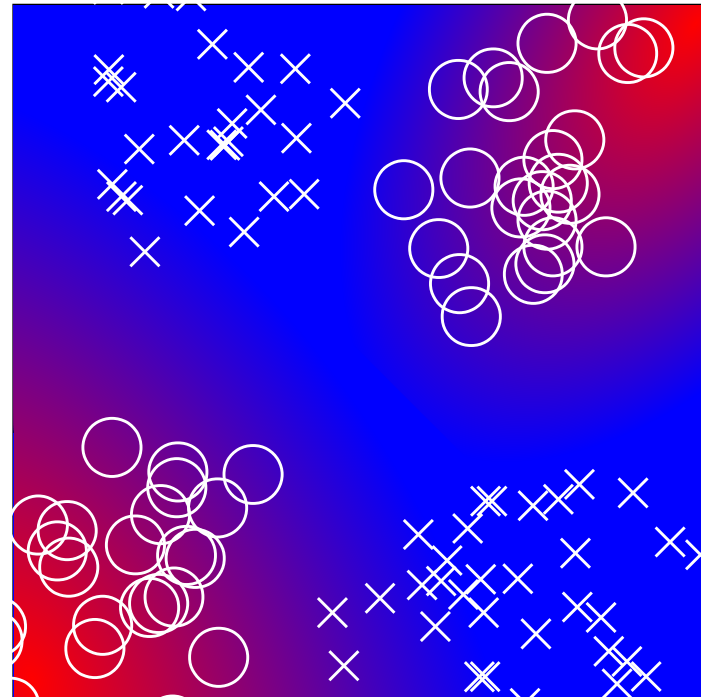
$$\Delta w_{ij} = \eta \cdot x_i \cdot (z_j - y_j)$$

# La Separación

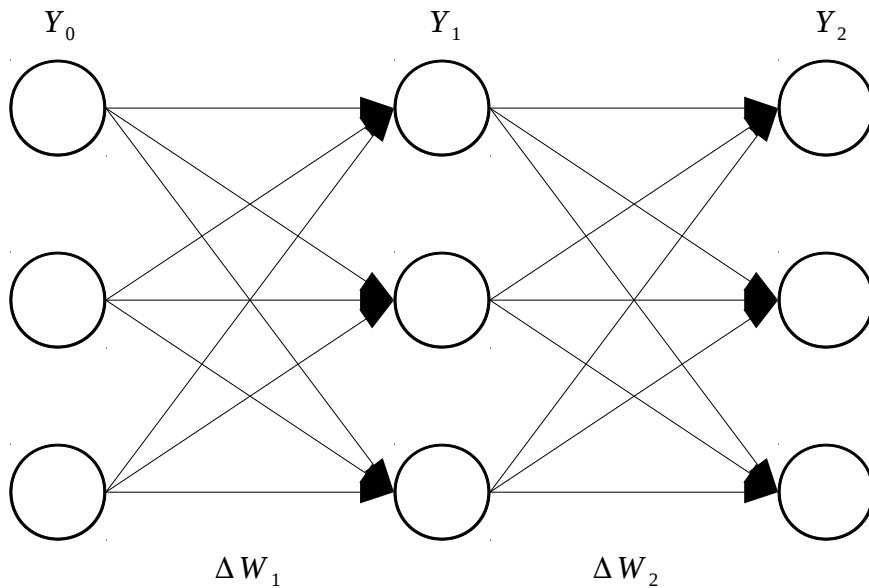
Lineal



No Lineal

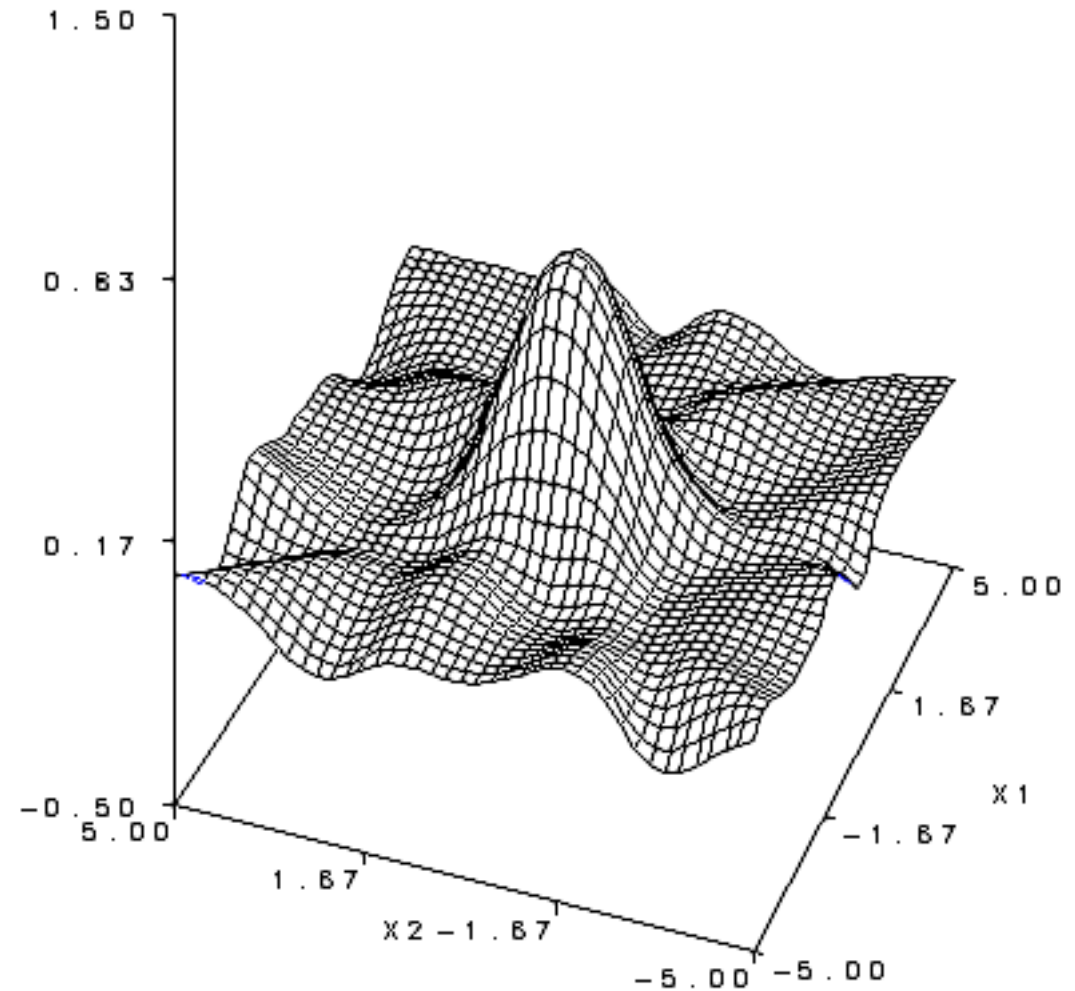


# El Perceptrón Multicapa

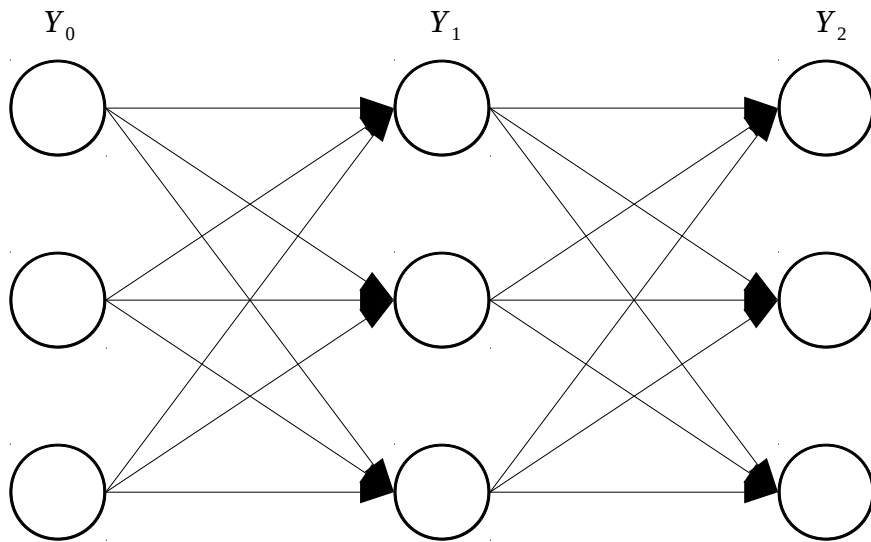


$$\Delta W_2 = \eta \cdot Y_1^T \cdot (Z - Y_2)$$

$$\Delta W_1 = \eta \cdot Y_0^T \cdot (? - Y_1)$$

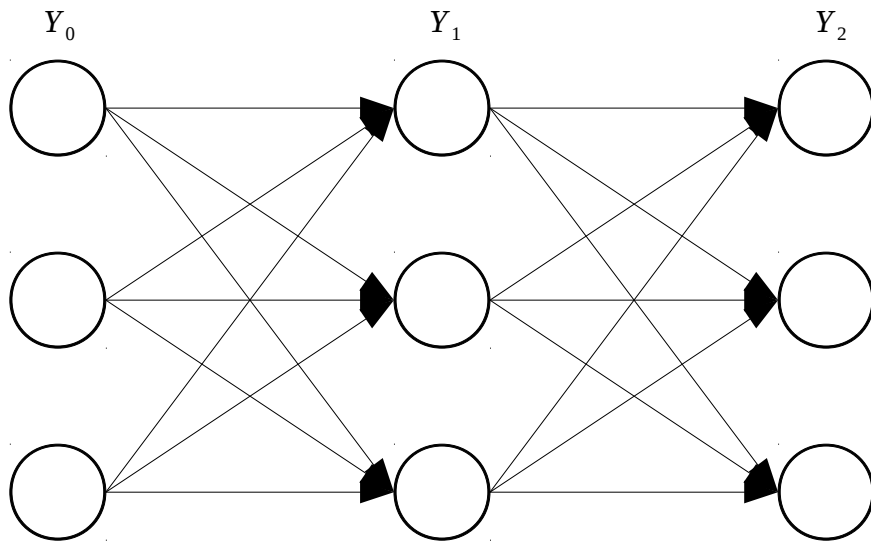


# La Regla Delta Generalizada



$$\Delta w_{ij} = -\eta \cdot \frac{\partial E}{\partial w_{ij}}$$

# La Retropropagación del Error



$$Y_0 = X$$

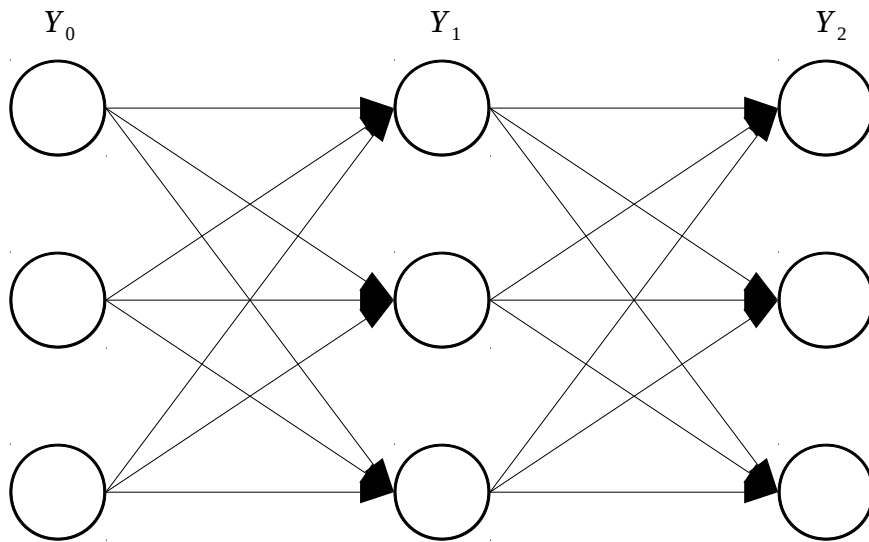
$$Y_1 = s_1(Y_0 \cdot W_1)$$

$$Y_2 = s_2(Y_1 \cdot W_2)$$

$$\Delta w_{ij} = -\eta \cdot \frac{\partial E}{\partial w_{ij}}$$



# La Retropropagación del Error



$$Y_0 = X$$

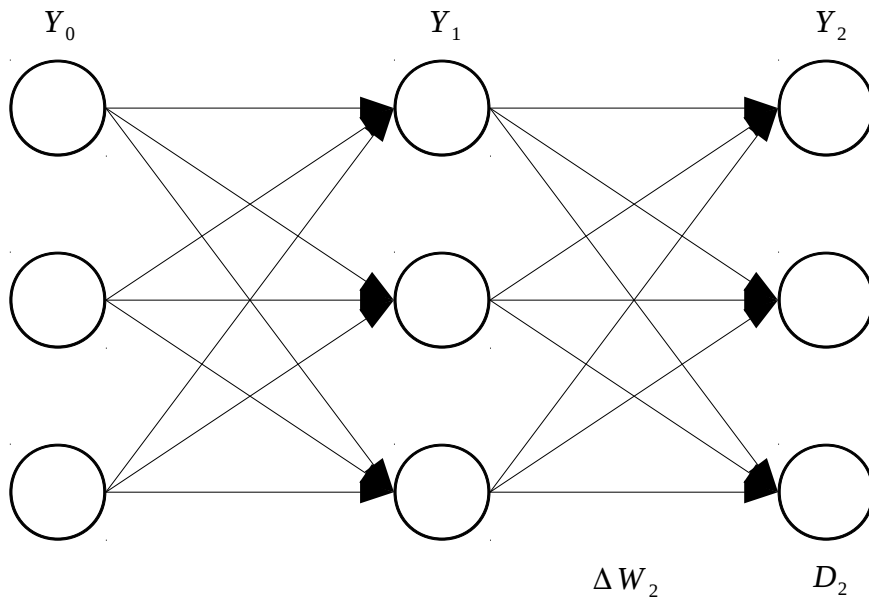
$$Y_1 = s_1(Y_0 \cdot W_1)$$

$$Y_2 = s_2(Y_1 \cdot W_2)$$

$$\Delta w_{ij} = -\eta \cdot \frac{\partial E}{\partial w_{ij}}$$

$$Y_l' = s'(Y_{l-1} \cdot W_l)$$

# La Retropropagación del Error



$$Y_0 = X$$

$$Y_1 = s_1(Y_0 \cdot W_1)$$

$$Y_2 = s_2(Y_1 \cdot W_2)$$

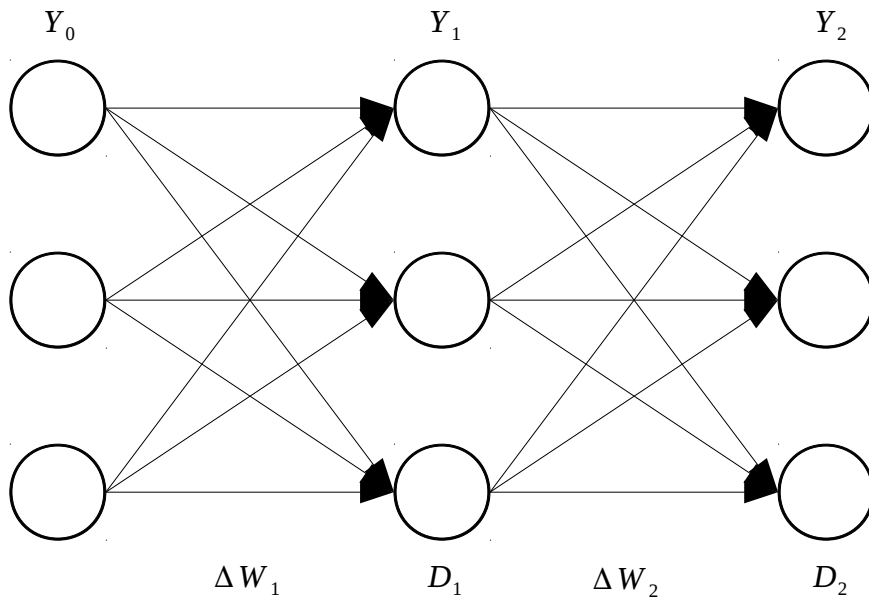
$$D_2 = (Z - Y_2) \cdot Y_2'$$

$$\Delta W_2 = \eta \cdot (Y_1^T \cdot D_2)$$

$$\Delta w_{ij} = -\eta \cdot \frac{\partial E}{\partial w_{ij}}$$

$$Y_l' = s'(Y_{l-1} \cdot W_l)$$

# La Retropropagación del Error



$$\Delta w_{ij} = -\eta \cdot \frac{\partial E}{\partial w_{ij}}$$

$$Y_l' = s'(Y_{l-1} \cdot W_l)$$

$$Y_0 = X$$

$$Y_1 = s_1(Y_0 \cdot W_1)$$

$$Y_2 = s_2(Y_1 \cdot W_2)$$

$$D_2 = (Z - Y_2) \cdot Y_2'$$

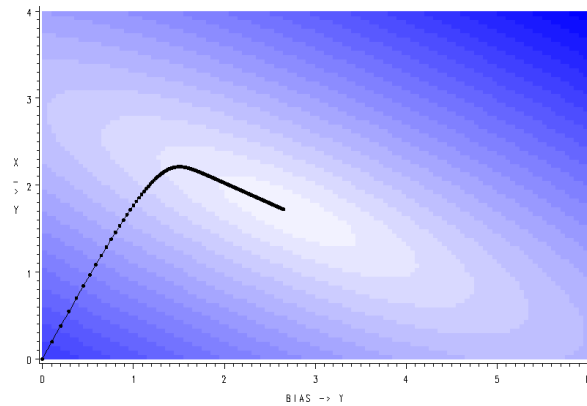
$$\Delta W_2 = \eta \cdot (Y_1^T \cdot D_2)$$

$$D_1 = (D_2 \cdot W_2^T) \cdot Y_1'$$

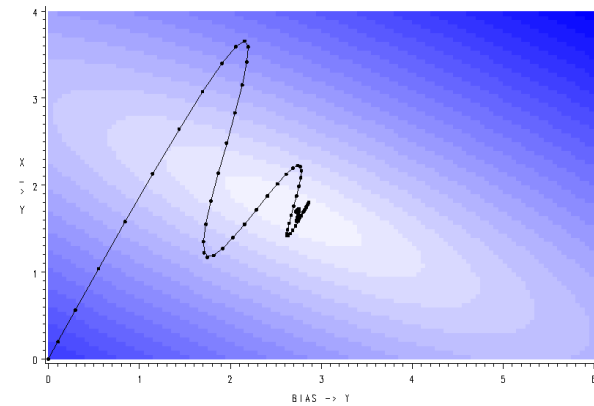
$$\Delta W_1 = \eta \cdot (Y_0^T \cdot D_1)$$

# Mejoras a BackProp

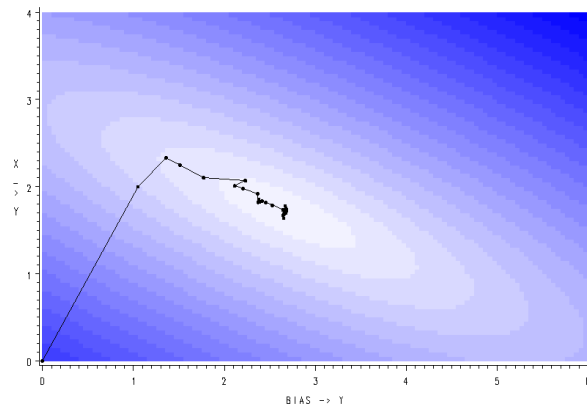
Fair Condition: Backprop Learn=.01 Momentum=0, 593 Iterations



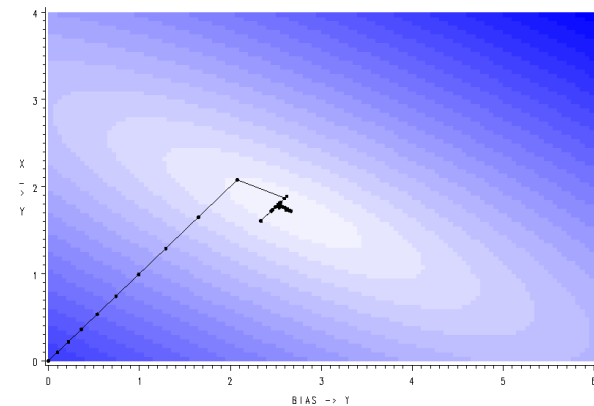
Fair Condition: Backprop Learn=.01 Momentum=.9, 80 Iterations



Fair Condition: Quickprop, 26 Iterations



Fair Condition: RProp, 29 Iterations





*That's all Folks!*