REDES AUTOORGANIZATIVAS

1. Introducción a la Autoorganización.

2. Aprendizaje Competitivo.
    3.1. Carácterísticas.
    3.2. Ventajas y Limitaciones.
3. Modelo de Kohonen. (Clustering)

4. Mapas Topológicos de Características Autoorganizados de Kohonen.
    4.1. Funcionalidad.
    4.2. Algoritmo de entrenamiento.
    4.3. Propiedades de los mapas topológicos.
    4.4. LVQ.
    4.5. Aplicaciones de los Mapas Autoorganizados.

INTRODUCCIÓN

- Consiste en la modificación repetida de conexiones en respuesta a patrones de activación y siguiendo unas reglas preestablecidas, hasta el desarrollo final de la estructura o sistema.

- La organización en las redes neuronales tiene lugar en 2 niveles diferentes, los cuales están íntimamente interconectados en forma de lazo:

  Actividad: Ciertos patrones de actividad son producidos por la estructura en respuesta a señales de entrada.

  Conectividad: Los Pesos de las diferentes interconexiones, son modificados en respuesta a señales producidas por los PE. (Concepto de ***Plasticidad***).

PRINCIPIOS:

Principio1: Modificaciones en los Pesos tienden a ***auto-amplificarse***.

El proceso de auto-amplificación está relacionado con el hecho de que las modificaciones de los Pesos están basadas en señales variables localmente.

Principio 2: Limitaciones debido a la ***competitividad*** establecen la selección de grupos de interconexiones fuertes, a expensas de otras.

Principio 3: Modificaciones en los Pesos tienden a una cooperación entre todos ellos.

Contexto de las Redes Autoorganizativas.

- Existen problemas donde el conjunto de entrenamiento está formado por un conjunto de patrones de entrada y donde las salidas deseadas no están disponibles.

- En estos casos la información relevante debe de ser localizada en los propios patrones de entrada (redundancia).

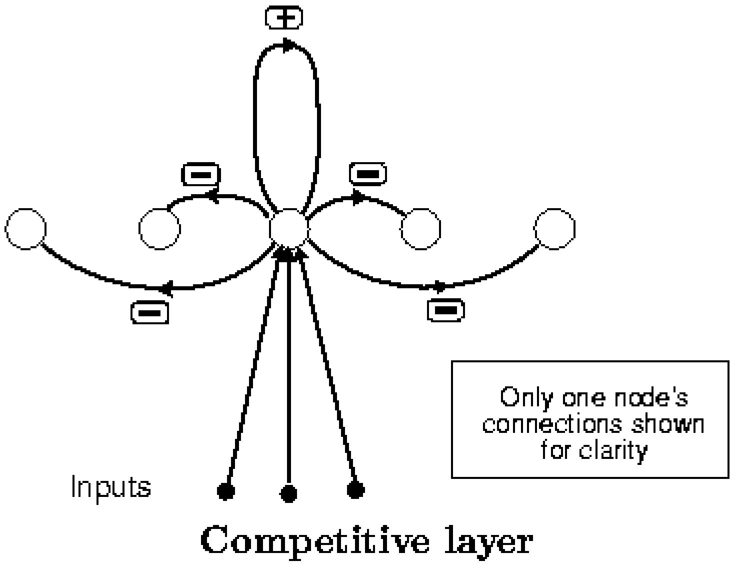
Problemas tales como:

- ***Clustering***: Los datos de entrada pueden ser agrupados en "clusters" y el sistema al procesar los datos debe de encontrar los centros de esos clusters.
- ***Cuantización de Vectores***: Este problema ocurre cuando en un espacio continuo tiene que ser discretizado. La entrada al sistema son vectores n-dimensionales y la salida es una representación discreta del espacio de entradas.
- ***Reducción de Dimensionalidad***: Los datos de entrada deben de ser agrupados en un subespacio con una dimensionalidad más baja que la dimensionalidad de los datos.
- ***Extracción de Características***: El sistema tiene que extraer características de los datos de entrada (supone casi siempre una reducción de la dimensionalidad).

APRENDIZAJE COMPETITIVO

- Esta idea se basa en la existencia de una cierta ***competitividad*** entre los PEs de una cierta capa por la oportunidad de entrenarse (aprender).

    - Esto, se refiere a que, el PE que produce la salida mayor se le considera Ganador, y tiene la capacidad de inhibir a los otros PEs (no presentan activación: salida nula). Todo ello conlleva que solamente los pesos del PE ganador podrán ser ajustados.



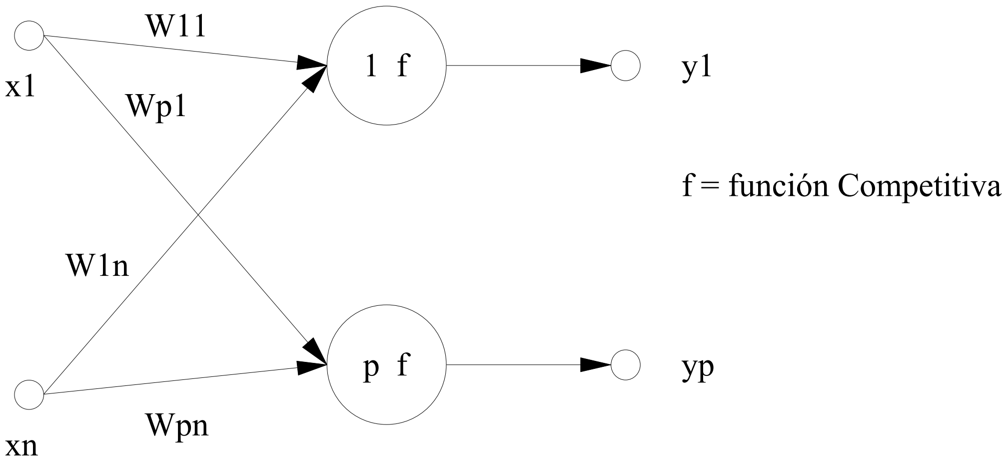
Cada nodo recibe las mismas entradas de una capa previa y entre los nodos de la capa existen conexiones laterales: excitadora (sobre si mismo), inhibidoras (sobre todos los demás nodos de la capa).

MODELO DE KOHONEN. (Clustering)

Matlab, Neural Neurocomputing capt. 5

- Teuvo Kohonen, profesor de la Facultad de Ciencias de la Información (Universidad de Helsinki), trabajó extensivamente, en lo que se denomina ***Memorias Asociativas*** y en modelos para actividad neurobiológica.

- En líneas generales, las redes entrenadas mediante esta regla se caracterizan por diferentes factores:



Estructura Competitiva

- Asocian vectores de entrada a patrones de salida.

- El aprendizaje es Sin Supervisar.

- Las estructuras de las redes las forman solamente 2 capas.

Se determina la similitud entre los

PE y el vector de entrada mediante la formula:

$$d_j = \sum_{i=1}^{n} (x_i(t) - w_{ji}(t))^2$$

Una vez que se determina cual es el PE ganador (supongamos "k") se aplica la función de transferencia de tipo competitiva:

$$y_j = 0 \quad \forall j \neq k$$

$$y_j = 1 \quad j = k$$

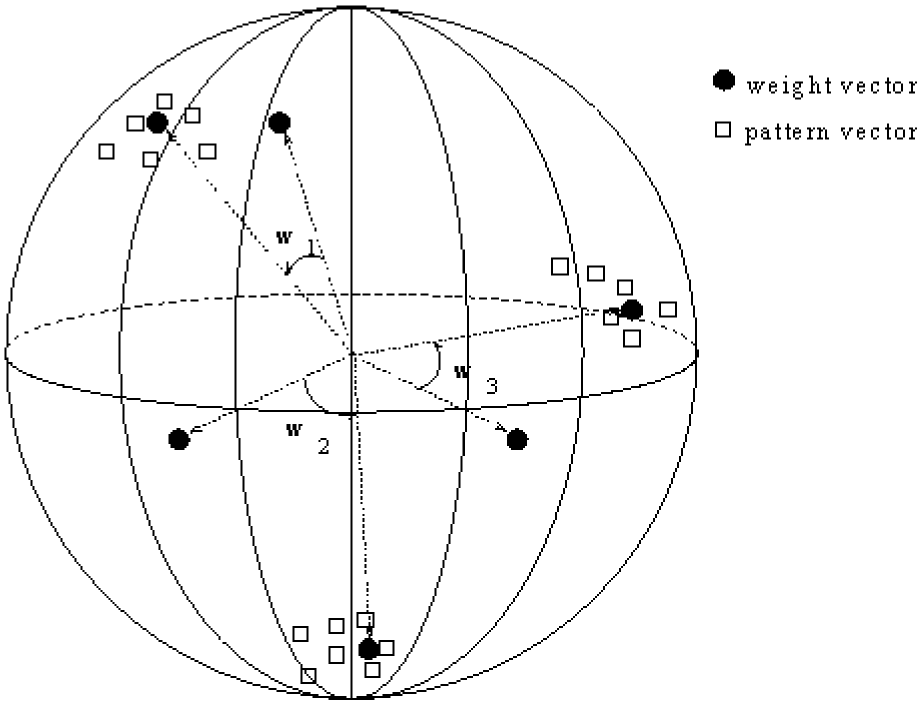
Una vez seleccionado el PE ganador se modifican sus pesos mediante la siguiente regla:

$$w_{ki}(t+1) = w_{ki}(t) + \mu(t)(x_i(t) - w_{ki}(t))$$

- En algunos casos, la velocidad de aprendizaje, suele ***ser dependiente del tiempo y normalmente decrece en cada paso***.

CARACTERÍSTICAS

- La velocidad de aprendizaje suele disminuir con el tiempo, hasta que toma un valor próximo a 0 en cuyo caso el aprendizaje finaliza.



Se trata de asociar cada PE de la capa de salida a un grupo de vectores de entrada, con una cierta similitud, generando de tal manera ***clases o clusters***. Así, los pesos de los Pes, son interpretados como los centros de los clusters.

Si existen más PE en la capa de salida que clases de patrones de entrada, pueden suceder dos cosas:

- quedan PE inactivos, sin asociar a ninguna clase
- una clase tiene asociados más de un PE.

• Normalmente los patrones de entrada se ***Normalizan***. Razón:

El PE ganador se determina calculando una medida de similaridad entre entrada y peso. Dicha similaridad se calcula empleando, normalmente, ***la Distancia Euclídea*** y ésta compara magnitudes y orientación espacial.

Patrones y Pesos Normalizados.

• Este hecho supone incrementar la velocidad de aprendizaje ya que, existe menos variabilidad en el espacio de pesos.

Por ejemplo:(1,1,1,1);(4,4,4,4) son idénticos---- (1/2,1/2,1/2,1/2).

Una ***Limitación de las redes competitivas*** es que algunas neuronas pueden no ser entrenadas. En otras palabras, pueden existir vectores de pesos muy distanciados de las entradas, con lo cual, nunca ganarán. Importantísimo en estos sistemas es la inicialización de pesos.

Demostración de que el aprendizaje competitivo sigue la idea de la Regla Delta.

La función error para el patrón p es: $$E^p = \frac{1}{2}\sum_j \left(w_{kj} - x_j^p\right)^2$$

$$\Delta_p w_{ij} = -\gamma \frac{\partial E^p}{\partial w_{ij}}$$

$$\frac{\partial E^p}{\partial w_{ij}} = \left\{\begin{array}{ll} w_{ij} - x_j^p & \text{i ganadora} \\ 0 & \text{en otro caso} \end{array}\right\}$$

$$\Delta_p w_{ij} = -\gamma\left(w_{ij} - x_j^p\right) = \gamma\left(x_j^p - w_{ij}\right)$$

MAPAS AUTOORGANIZATIVOS (TOPOLÓGICOS)

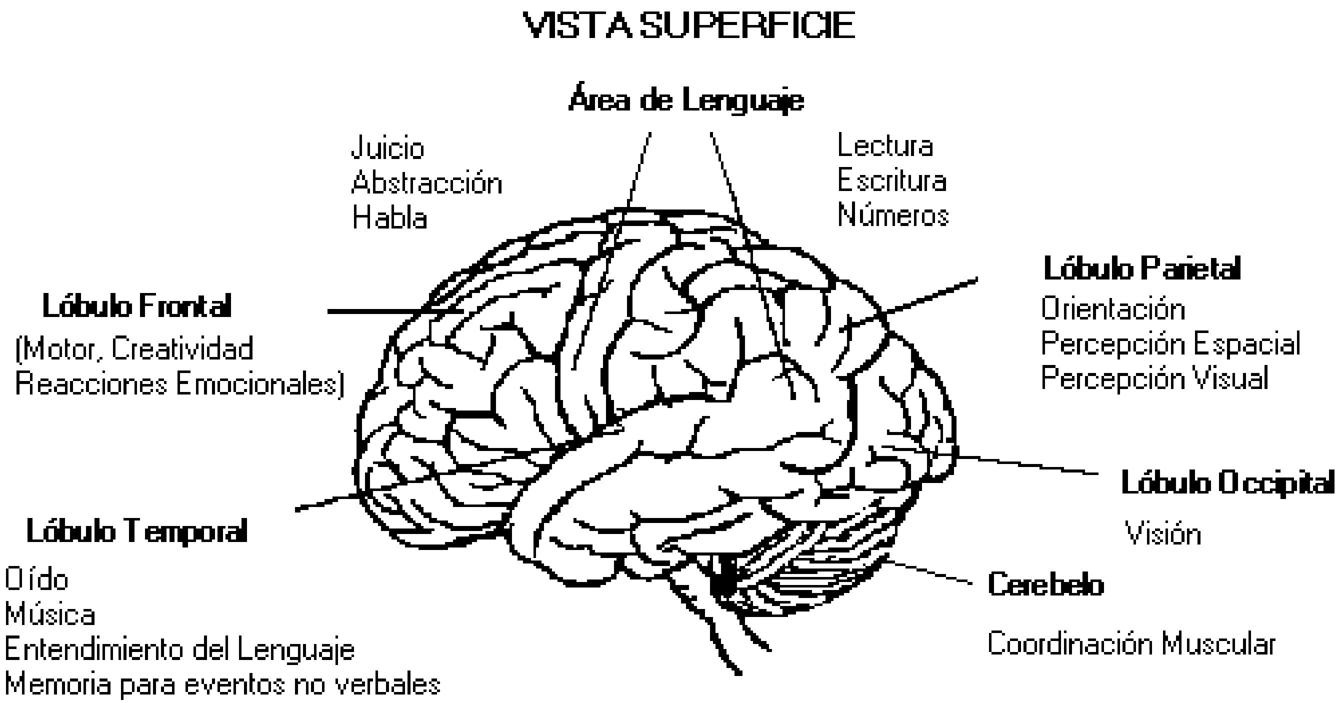
MAPAS CEREBRALES

La representación económica de la información es uno de los problemas centrales de las ciencias de la información.

El cerebro humano posee la capacidad y la habilidad para operar con grandes conjuntos de datos. Es capaz de obtener representaciones reducidas de hechos relevantes sin que ello implique pérdida de información en los datos y sus relaciones.

"*El supuesto del procesado inteligente de información puede ser visto como la creación de imágenes simplificadas del mundo real con diferentes niveles de abstracción, en relación a un subconjunto particular de datos observables*" [kohonen].

Es conocido que varias áreas del cerebro, especialmente de la corteza cerebral, están organizadas de acuerdo a diferentes modalidades sensoriales.



Experimentos más recientes han revelado una estructura-fina en muchas áreas. Por ejemplo: las señales de respuesta en áreas como la visual, son obtenidas en el mismo orden topográfico sobre la corteza con que éstas fueron recibidas en los órganos sensoriales.

La posibilidad de que la representación del conocimiento en una categoría particular pueda asumir la forma de un mapa de características y en cierto sentido organizado geométricamente sobre la corteza del cerebro, ha motivado una serie de investigaciones que han dado lugar a nuevos modelos de redes neuronales artificiales.

Algunas investigaciones han revelado que ciertas capas de neuronas tienen la habilidad de cambiar sus respuestas de tal modo que la localización de la neurona en la red donde la respuesta es obtenida especifica una cierta característica del conjunto de patrones de entrada. Esta especificación ocurre en el mismo orden topológico que describe la relación de similaridad entre los patrones de entrada.

Las neuronas o unidades no se mueven, es el conjunto de sus parámetros internos quienes definen su especificidad.

Las redes son normalmente bidimensionales: Existe un mapeado que es capaz de preservar las relaciones topológicas mientras se consigue una reducción de dimensionalidad del espacio de entradas.

MAPAS AUTO-ORGANIZATIVOS

Estructura de Neuronas Artificiales

Mapas cerebrales

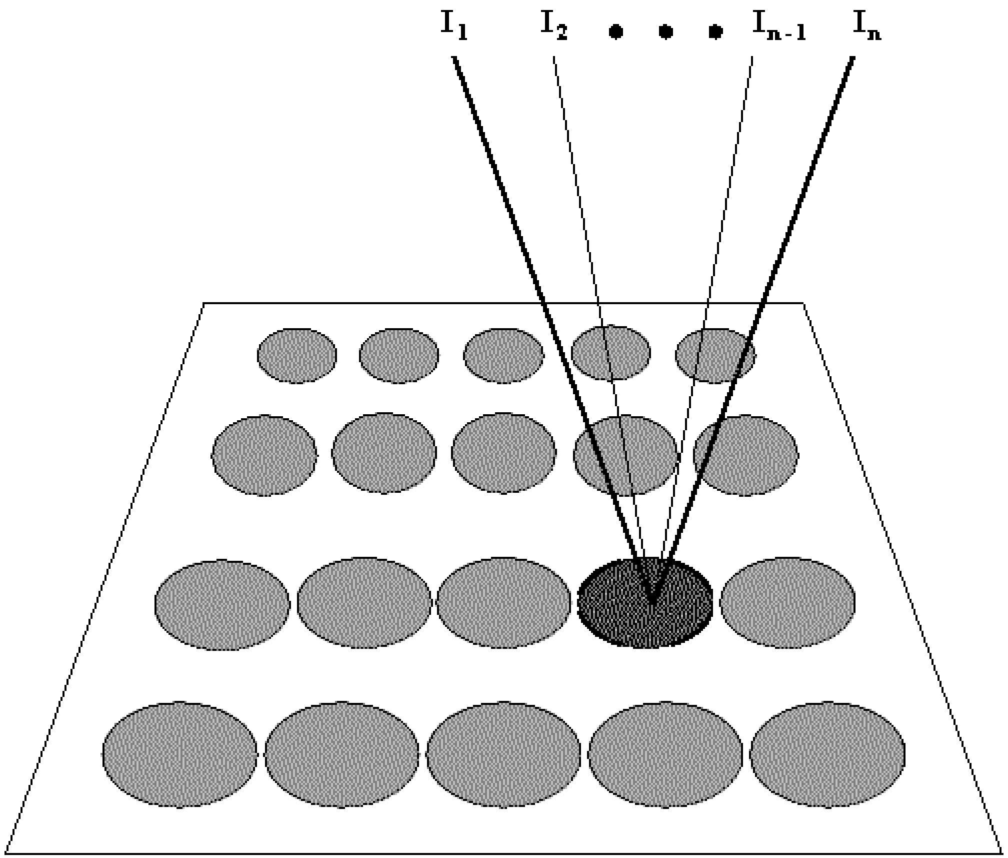
Fundamentos:
- Aprendizaje Competitivo.
- Relación Topológica entre EP.

Modelado de Datos mediante un mapeado espacial

Mapas Autoorganizativos Topológicos (SOM)

T. Kohonen

- Preprocesado de patrones para su reconocimiento
- Proyectar y visualizar espacios de entrada multi-dimensionales en espacios de salida bidimensional. (Mapas Bidimensionales)



- Estructura formada por dos capas.
- Las neuronas en la capa de salida se ordenan en espacios n-dimensionales.
- Los pesos de las unidades de salida se adaptan de forma que el orden presente en el espacio de entradas persiste en la salida.
- La idea del algoritmo de aprendizaje consiste ***en organizar los PEs en el espacio de salida, en regiones (locales)*** que actuarán como clasificadores de los datos de entrada.
- El mapa topográfico es organizado de manera autónoma mediante un proceso cíclico de comparación de patrones de entrada (entrenamiento) y vectores de pesos.

El SOM (Caso Espacio Euclídeo)

Estructura Bidimensional

Inicialización de la estructura:
- Valores de Pesos
- Parámetros de aprendizaje

Presentación Patrones

Selección elemento ganador

- Utilización de una métrica.
  (Distancia Euclídea)

Proceso Aprendizaje

Modificar Pesos:
Se modifican los pesos del EP ganador y de todos aquellos que se encuentren topológicamente cercanos al ganador.

h(i,k)-función dependiente de distancia entre unidades.
S- Región de Interés considerada.

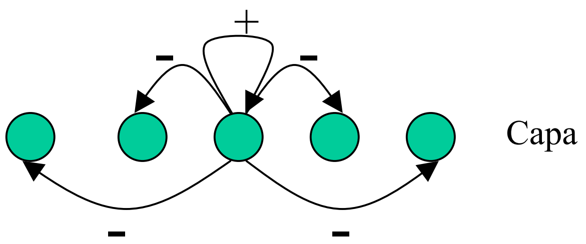
$$w_i(t+1) = w_i(t) + \gamma h(i,k)\big(x(t) - w_i(t)\big)\ \forall i \in S$$

Mecanismos Laterales que favorecen la Autoorganización

CONTROL DE ACTIVIDAD

Aprendizaje Competitivo: se basa en la existencia de una cierta ***competitividad*** entre los EPs de una cierta capa por la oportunidad de entrenarse (aprender).

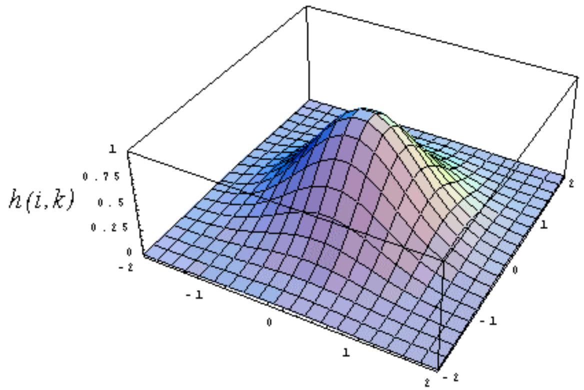
El EP que produce la salida mejor se le considera *Ganador*, y tiene la capacidad de inhibir a los otros EPs. Solamente los pesos del EP ganador podrán ser ajustados.



CONTROL DE PLASTICIDAD

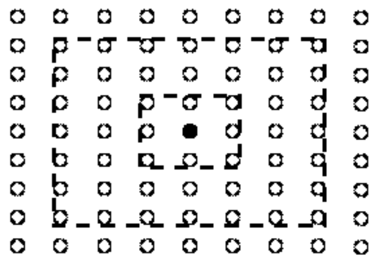
Dependencia topológica en el ajuste de los pesos: Mejora adaptiva en la vecindad de la neurona ganadora.

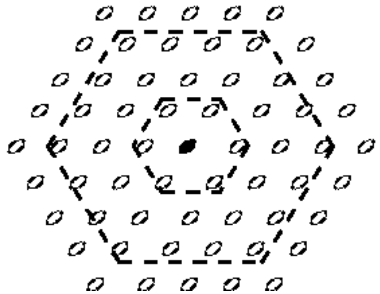
El control de plasticidad precisa como la actividad local determina el parámetro de aprendizaje en su vecindad.



Función que controla la velocidad de aprendizaje en torno a una neurona ganadora

Consideraciones (I)





1- Diferentes Regiones de Interés.

2- La regla de modificación de pesos sigue la idea de la Regla Delta.

$$E^p = \frac{1}{2}\sum_j \left(w_{kj} - x_j^p\right)^2$$

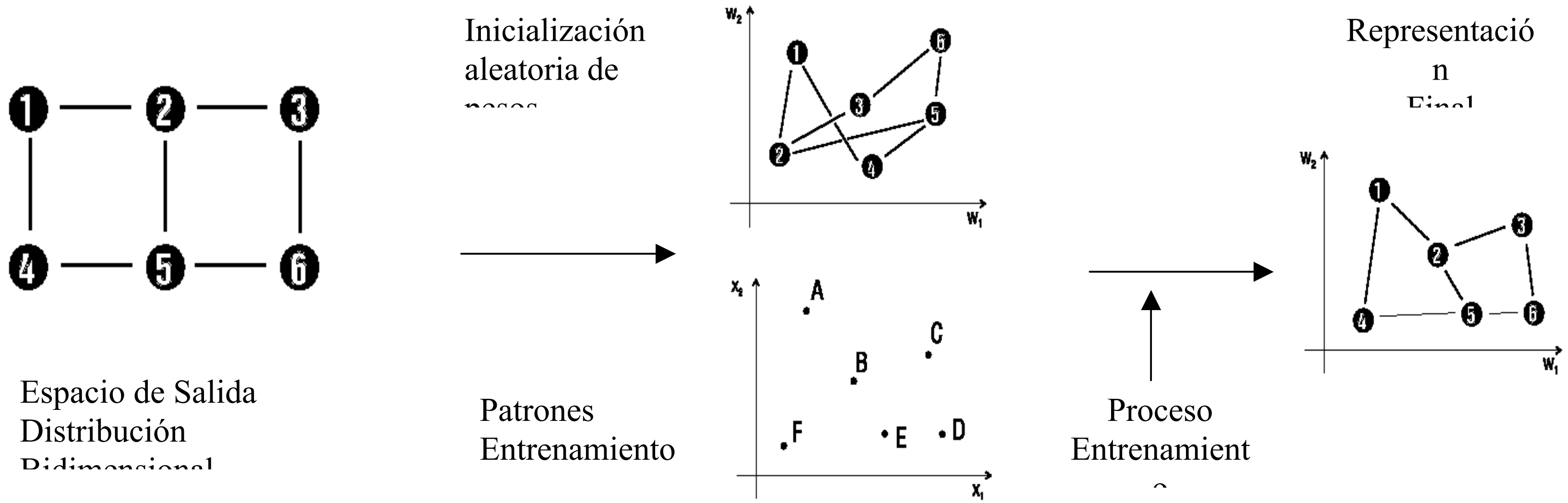
$$\Delta_p w_{ij} = -\gamma \frac{\partial E^p}{\partial w_{ij}}$$

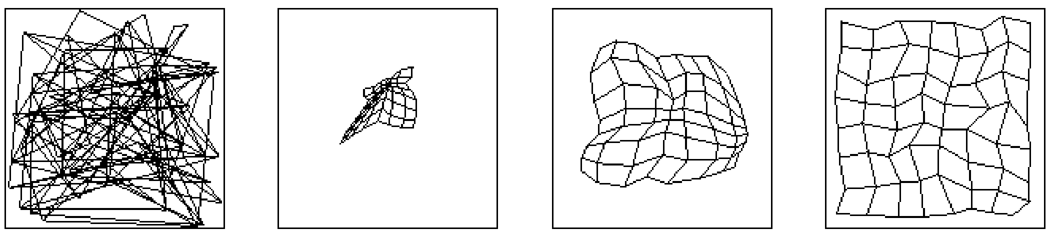
$$\frac{\partial E^p}{\partial w_{ij}} = \begin{cases} w_{ij} - x_j^p & \text{i ganadora} \\ 0 & \text{en otro caso} \end{cases}$$

$$\Delta_p w_{ij} = -\gamma\left(w_{ij} - x_j^p\right) = \gamma\left(x_j^p - w_{ij}\right)$$

Consideraciones (II)

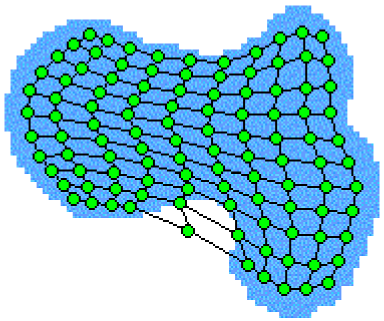
3- El proceso por el que los nodos que se encuentran topológicamente cercanos al ganador modifiquen sus pesos supone un *efecto de suavizado* o de *relajación* sobre los pesos de las neuronas pertenecientes a la región de vecindad que conducen hacia una *ordenación global*.





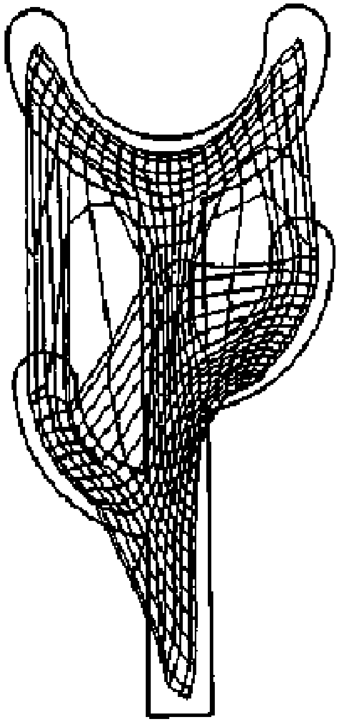
Iteration 0 Iteration 200 Iteration 600 Iteration 1900

PROPIEDADES DE LOS MAPAS AUTOORGANIZATIVOS



Propiedad 1: ***Aproximación del Espacio de Entrada.***
Los Mapas de características, representados por el conjunto de vectores de pesos, nos dan una buena aproximación al espacio de entrada.

Propiedad 2: ***Ordenación Topológica.***
Los Mapas son ordenados topológicamente en el sentido de que la localización espacial de una neurona corresponde a un dominio particular o conjunto de patrones de entrada.



Propiedad 3: ***Densidad de Muestreo.***
Los Mapas autoorganizativos reflejan variaciones en las estadísticas de la distribución de entrada: regiones en el espacio de entrada, cuyos vectores tienen altas probabilidades de ocurrencia, son mapeados en largos dominios del espacio de salida y entonces con mejor resolución que regiones de las cuales sus patrones ejemplo tienen pocas probabilidades de ocurrencia.

Propiedad 4: ***Selección Automática de la dimensionalidad de las Características***

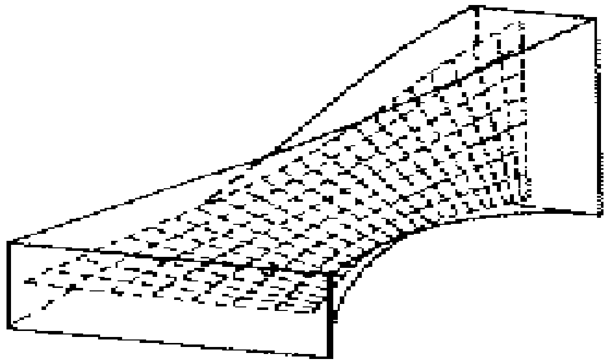
Existen dos tendencias opuestas en el proceso autoorganizativo:

- El conjunto de vectores de pesos tiende a describir la función de densidad de los vectores de entradas.
- Interacciones locales entre unidades de procesado tienden a preservar la continuidad de secuencias de vectores de pesos.

Como resultado de dichas fuerzas es que: la distribución de los vectores de pesos tienden a aproximar una superficie suavizada.

Representación de la forma y orientación óptima en el espacio de patrones que mejor imita la estructura global de la densidad de los patrones de entrada.

La distribución de los vectores de referencia tienden a encontrar aquellas dimensiones del espacio de patrones donde los vectores de entrada presentan una gran variación, representándola en el mapa de salidas.



CARACTERÍSTICAS:

- Con respecto a pesos: lo mismo que para el aprendizaje competitivo.

- Los nodos pertenecientes a la región de interés del ganador son también modificados. La razón de este hecho es que lo que se trata de conseguir es que la red cree regiones que respondan a valores muy próximos al del vector de entrenamiento. Los PE que se encuentren próximos al ganador tendrán un alineamiento similar y sobre el ciclo de entrenamiento esto será una ventaja en la representación de la clase para dicha entrada. Como consecuencia de que son similares espacialmente a vectores de entrenamiento, serán correctamente clasificados incluso aunque éstos no formen parte del conjunto de entrenamiento. Esto demuestra la ***Generalidad*** de esta estructura.

- La región de interés normalmente, se reduce de manera lineal dependiendo del nº de ciclos del entrenamiento:

- Las dos ideas centrales en las que se basa Kohonen para desarrollar sus estructuras usando aprendizaje ***autoorganizativo*** y competitivo son: ***" El proceso de adaptación de pesos y el concepto de geometría topológica de PEs***".

Consideraciones Prácticas para creación de Mapas (I)

Aprendizaje con un número pequeño de patrones de entrenamiento.
• Obtener una representación correcta del espacio de entradas.
• Diferentes métodos en la presentación de patrones.

Representación de manera adecuada los casos más significativos.
• Favorecer los casos importantes.

Escalamiento en las componentes de los patrones.
• Favorecer una correcta ordenación de los vectores representativos.

Forzar representaciones en ciertos lugares del mapa de entrada.
• Forzar la inicialización de pesos.
• Mantener sobre ciertas localizaciones velocidades de aprendizaje bajas.

Visualizar la calidad del aprendizaje.

Consideraciones Prácticas para creación de Mapas (II)

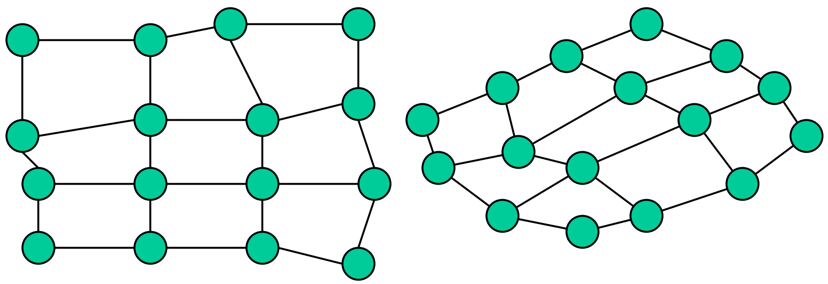
Utilización de Pesos Adaptivos.

Métrica Usual
Distancia Euclídea → Diferencias significativas en la varianza de las componentes de los patrones de entrada → Incorrectas Orientaciones

$$d^2[x(t), m_i(t)] = \sum_{j=1}^{N} \mu_{ij}^2 [x_j(t) - w_{ij}(t)]^2$$

$$e_{ij}(t+1) = (1-k_1)e_{ij}(t) + k_1 \mu_{ij} \left| x_j(t) - w_{ij}(t) \right|$$

$$e_i(t) = \frac{1}{N} \sum_{j=1}^{N} e_{ij}(t)$$

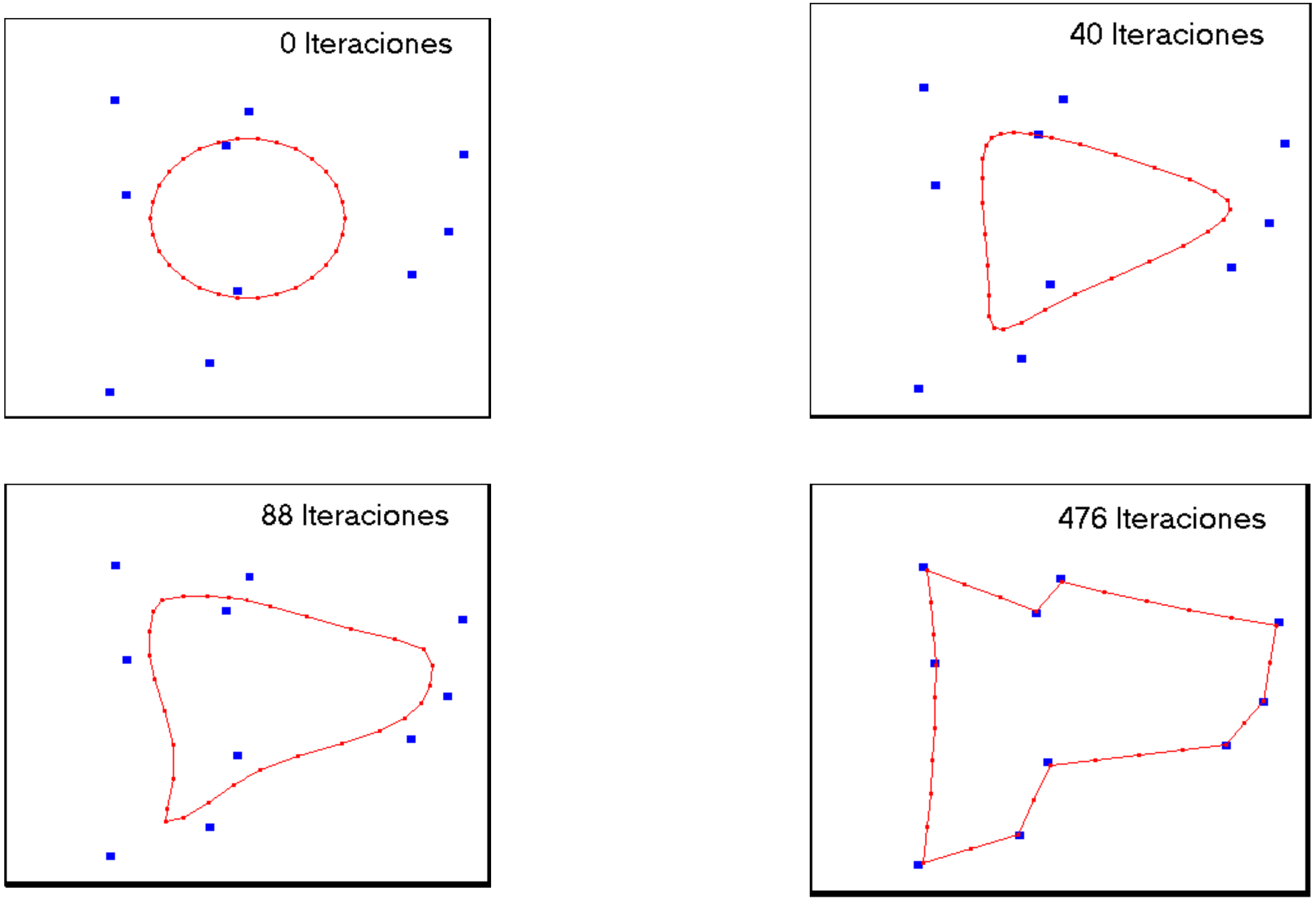


$$\mu_{ij}(t+1) = k_2 \mu_{ij}(t) \qquad 0 < k_2 < 1 \quad si \quad \mu_{ij} \left| x_j(t) - w_{ij}(t) \right| > e_i(t)$$

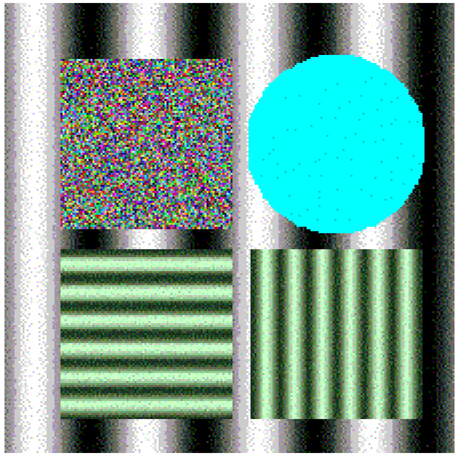
$$\mu_{ij}(t+1) = k_3 \mu_{ij}(t) \qquad k_3 > 1 \quad si \quad \mu_{ij} \left| x_j(t) - w_{ij}(t) \right| < e_i(t)$$

APLICACIONES:

Problema del Viajante



Segmentación de Imágenes



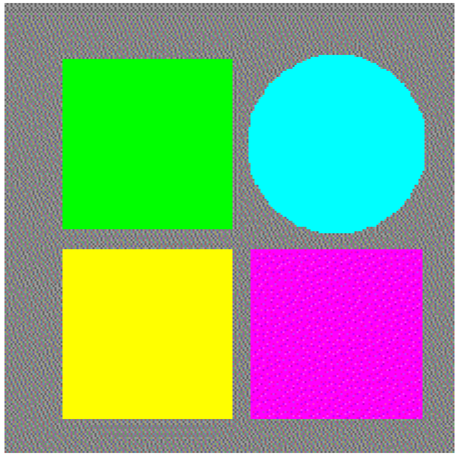
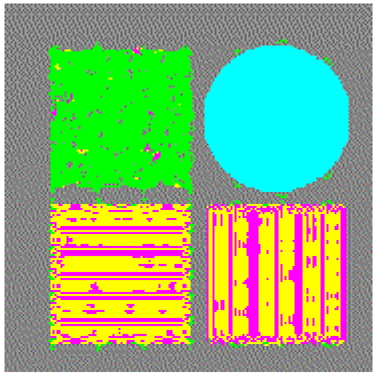
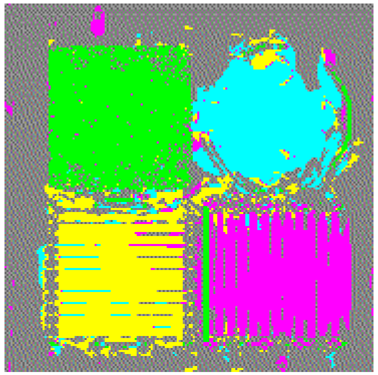


Imagen Original

Imagen Segmentada Ideal





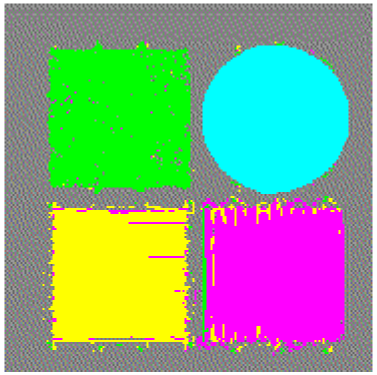
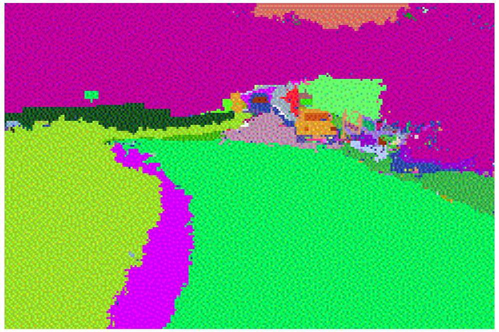
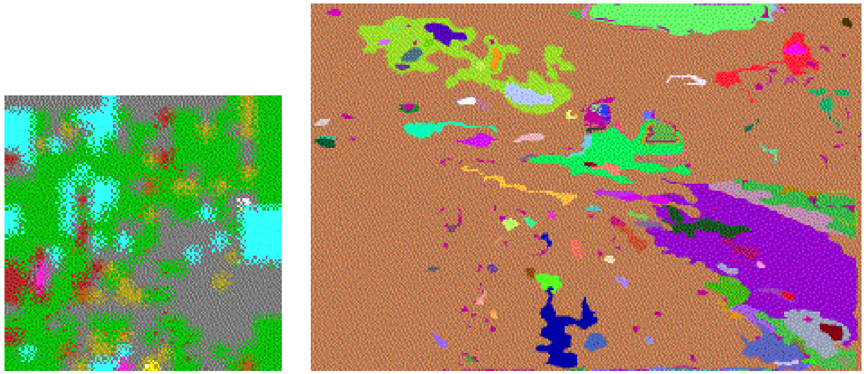


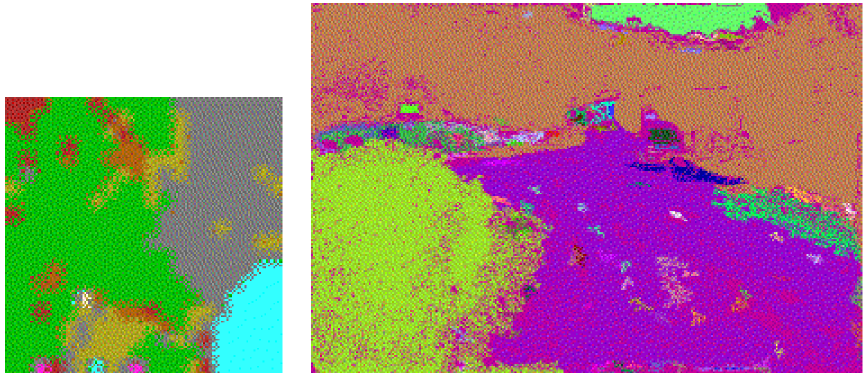
Imagen segmentada (Color)

Imagen segmentada (Textura)

Imagen Segmentada (Color y Textura)







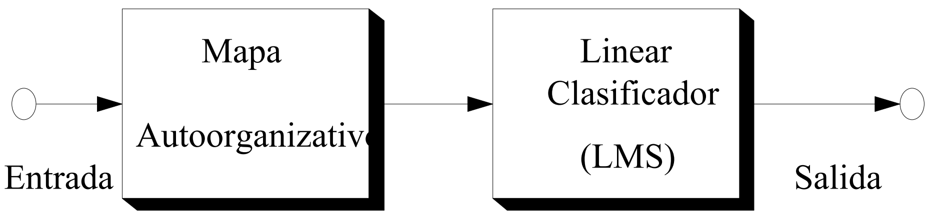
Segmentación Color

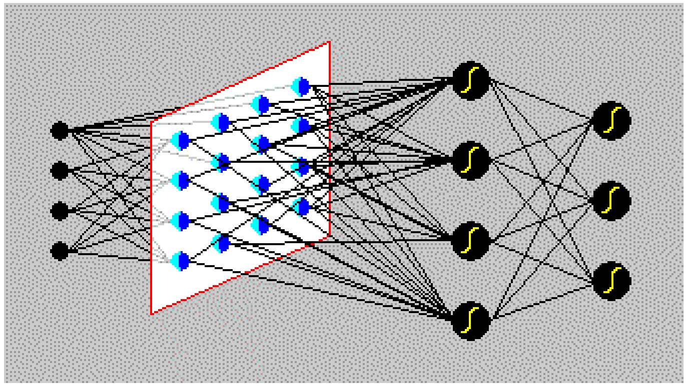
Segmentación Color y Textura

Clasificador de Mapa de Características. Learning Vector Quantization (LVQ)

- En problemas de clasificación de patrones, el requerimiento consiste en clasificar un conjunto de patrones de entrada en un número finito de clases. En estos problemas, una de las tareas más importantes consiste en delimitar las condiciones de separación de clases.

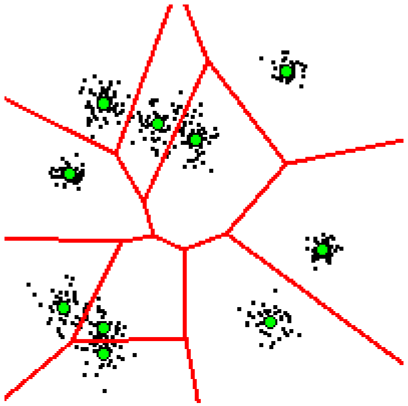
- Hasta el momento se han descrito métodos tanto en aprendizaje supervisado como en aprendizaje sin supervisar que trataban de resolver este problema. Diferentes estudios han llegado a demostrar que en ciertos casos un *método Híbrido* soluciona mejor este problema. La solución pasa por usar una combinación de un Mapa de características y un clasificador lineal con aprendizaje supervisado (Lippmann, 1989).





LVQ. Vector Quantization.

- ***Vector Quantization*** es una técnica por la cual el espacio de entradas es dividido en un número determinado de regiones y para cada una de ellas es definido un vector que la caracteriza. Un espacio continuo tiene que ser discretizado. La entrada al sistema son vectores n-dimensionales y la salida es una representación discreta del espacio de entradas.



Partición de un espacio continuo en 10 valores discretos

- Cuando al sistema se le presenta un nuevo vector de entrada, primero se le asigna una determinada región y después es representado por el vector característico de dicha región.

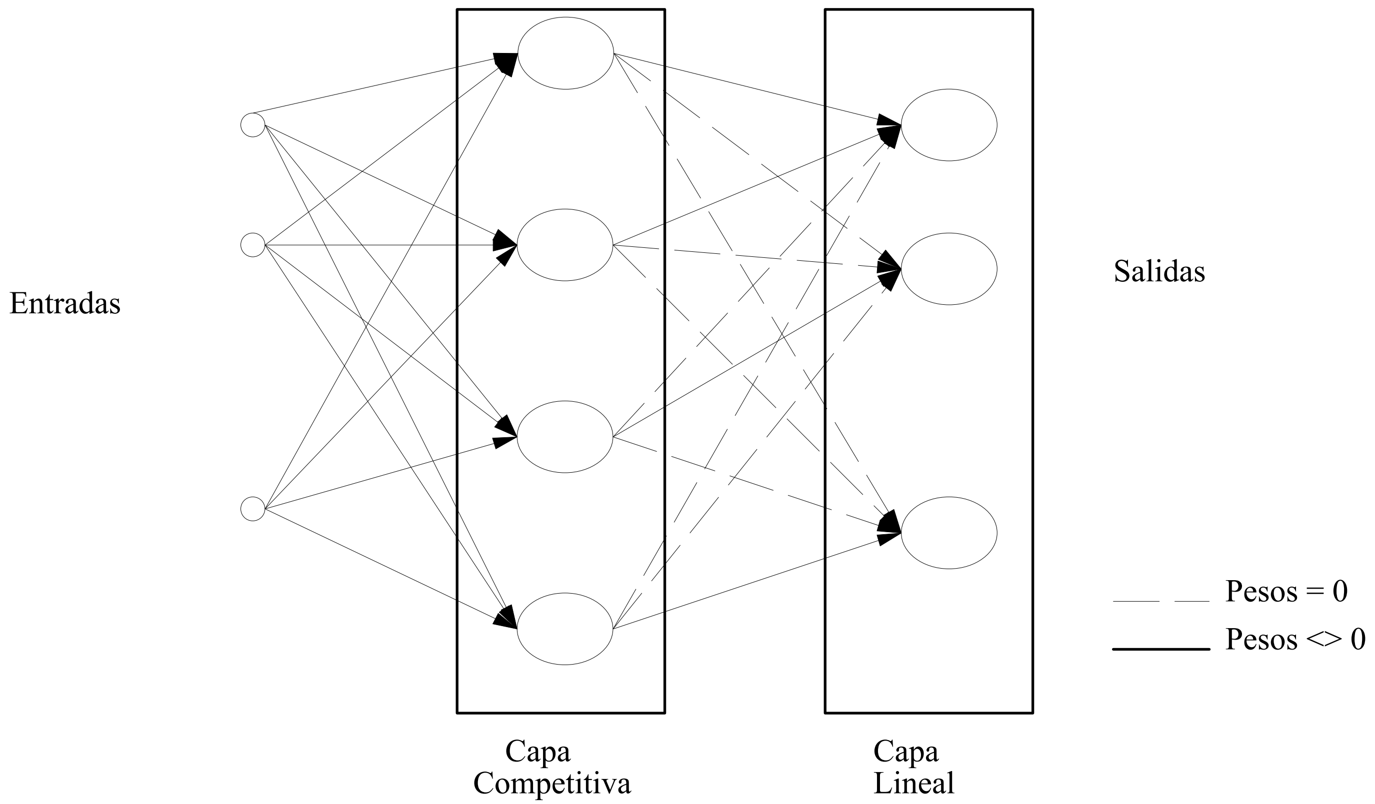
Esta técnica es frecuentemente utilizada en transmisión de datos, utilizando una versión codificada del vector de representación en lugar de los patrones de entrada.

- Cuando la medida de similaridad que se utiliza para asignar a un patrón de entrada a una determinada región es la Distancia Euclídea, el quantizer es denominado de ***Voronoi.*** Este, divide el espacio de entradas en ***Celdas de Voronoi***, y cada celda es representada por uno de los vectores de reconstrucción Wi. La i-esíma celda contiene aquellos puntos del espacio de entradas que están más cercanos al vector Wi que a cualquier otro Wj.

- La regla de aprendizaje competitivo basada en la distancia Euclídea puede se utilizada para posicionar un conjunto de vectores de representación (Kohonen 1989).

    - Este algoritmo de aprendizaje competitivo puede ser visto como un método aproximado para computar dichos vectores de manera sin supervisar.

Kohonen, diseñó versiones supervisadas de este método***, Learning Vector Quantization (LVQ)***, para problemas de ***clasificación adaptiva de patrones***. La información de clases se utiliza para afinar los vectores de representación: Mejora en la clasificación de regiones.



ESTRUCTURA QUE SIMULA LVQ

OPERACIÓN: LVQ es una técnica de aprendizaje supervisado. Un vector de entrada X es tomado aleatoriamente del espacio de entradas. Si las clases del vector X y la del W asociado coinciden, entonces el vector W es movido en la dirección de X. Por la otro parte, si las clases son diferentes, el vector W es alejado de X.

Las reglas del cambio de los pesos asociados a los PE que forman la región a la cual el Mapa Autoorganizativo da como ganadora son las siguientes:

$$Si \quad \varsigma_w = \varsigma_x$$

$$w(t+1) = w(t) + \mu(t)(x - w(t))$$

$$Si \quad \varsigma_w \neq \varsigma_x$$

$$w(t+1) = w(t) - \mu(t)(x - w(t))$$

Es deseable que el factor de ganancia $\mu(t)$ decrezca con el tiempo.

Los pesos de la capa lineal:

- Se inicializan a valores (1 y 0) antes del aprendizaje (Usual).
- Se modifican utilizando la Regla Delta.