

Tema 4

Aprendizaje no supervisado

Aprendizaje no supervisado

- Introducción
- Aprendizaje competitivo
- Mapas autoorganizados de Kohonen
- Otros algoritmos de agrupación

Introducción

- **Características principales de las RNA no supervisadas**

- No necesitan de un profesor externo
- La red descubre en los datos de entrada y de forma autónoma:
 - Características
 - Regularidades
 - Correlaciones
 - Categorías

Obtiene de forma automática la clasificación más natural de los datos

- Por tanto, muestran cierto grado de **Auto-organización**
- Sólo consigue resultados útiles si existe algún tipo de redundancia
 - sin redundancia sería imposible encontrar patrones o características en los datos

Introducción

- **Características principales (II)**

- Suelen requerir **menores tiempos de entrenamiento** que las supervisadas
- **Modelos más cercanos a estructuras neurobiológicas**. Tienden a imitar su comportamiento
- **Arquitectura simple**. Habitualmente son:
 - Una sola capa
 - Feed-forward
- Tipos fundamentales
 - Kohonen (SOM)
 - Grossberg (ART)

Introducción

- **Problemas abordables con estas redes**

- Agrupamiento (Clusterización)

- Determinar la existencia de clases y clasificar los patrones

- Prototipado

- Obtener prototipo (representante) asociado a la clase del patrón de entrada

- Extracción y relación de características

- Mapa topológico de los datos de entrada: que patrones parecidos produzcan respuestas similares en células cercanas (Ej: Mapa de pobreza mundial)

Modelo básico

- **Regla de Hebb**

- Esta regla contribuye a **decidir cómo y en qué grado modificar las conexiones sin depender de factores externos**

- Varias propuestas

$$\Delta w_{ij} = a_i \cdot a_j$$

$$\Delta w_{ij} = \Delta a_i \cdot \Delta a_j$$

Δa_i es la variación de la activación con respecto a su activación anterior

$$\Delta w_{ij} = \mu (a_i - \bar{a}_i)(a_j - \bar{a}_j) \quad \bar{a}_i \text{ es la media de las activaciones}$$

Modelo básico

- Características básicas de las RNA no supervisadas:
 - Interacción lateral: El nivel de activación de una célula o neurona no depende sólo de la entrada sino de su ubicación y de la actividad de su vecindario
 - Aprendizaje competitivo: Las neuronas compiten entre ellas

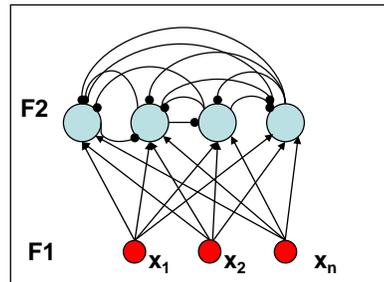
Aprendizaje competitivo

- **Características**
 - Tipo de Aprendizaje no supervisado utilizado en modelos de RNA
 - Objetivo: categorizar los datos de entrada: Datos parecidos han de ser clasificados (agrupados) como pertenecientes a la misma categoría
 - Cada categoría representada por una célula o neurona (también llamada prototipo)
 - Hay una capa de competición compuesta por tantas neuronas o células como categorías pueda haber en los datos. Cada neurona se corresponde con un prototipo
 - El sistema debe relacionar cada neurona (prototipo) con los datos de entrada que representa

Aprendizaje competitivo

• Arquitectura

- Dos capas: **F1 capa de entrada** y **F2 capa de competición**
- Cada neurona de la capa F1 está conectada con todas las neuronas de la capa F2 a través de conexiones ponderadas (con pesos)
- F2 tiene conexiones laterales inhibitorias entre todas las neuronas de su capa, excepto con ella misma, en que la conexión es excitatoria
- Cada neurona se refuerza a sí misma: Una vez que las neuronas de F2 han sido activadas por las entradas de F1, ésta deja de actuar y la señal se propaga por F2: todas las neuronas intentarán impedir que las demás tengan un valor alto (conex. inhibit) e intentarán tener ellas mismas un valor alto (conex. excitat).



Aprendizaje competitivo

Procedimiento

1. Se recibe el estímulo o entrada en F1.
2. Se propaga la señal hasta F2 y se calcula el valor de activación para cada neurona de F2.
3. Se inhiben las conexiones entre la capa F1 y la F2.
 - Se propaga la señal por la capa F2, calculándose los nuevos valores de activación de las neuronas.
 - Cuando sólo haya una neurona (la ganadora) con un valor de salida mayor que cero ir al paso 4.
4. Restablecer las conexiones entre las capas F1 y F2.
 - **Aprendizaje:** Calcular los nuevos valores para los pesos de las conexiones entre la capa F1 y la neurona ganadora en el paso 3.

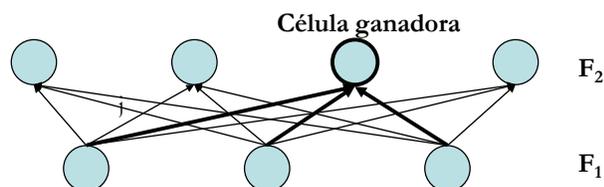
Aprendizaje competitivo

- La capa F2 se estabiliza cuando todas las salidas tienen valor 0 excepto una que será la que al principio ha recibido la entrada más alta de F1 (la que ha inhibido más al resto y se ha reforzado más a ella misma)
- Las neuronas de la capa F2 compiten por la entrada: sólo una consigue ganar la competición (neurona ganadora)
- En F2 los pesos son fijos

Aprendizaje competitivo

• Aprendizaje

- La neurona ganadora representa al prototipo que es asignado al dato de entrada
- Las conexiones entre la capa F1, la de la entrada, y la célula ganadora son reforzadas
- Es por esto que se llama "el que gana se lo lleva todo" (winner takes all)

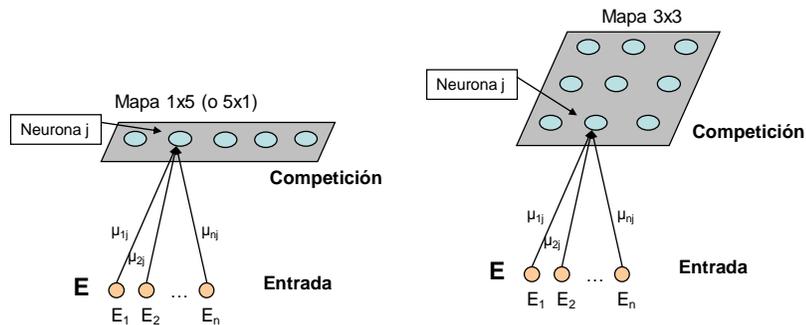


• Inconveniente

- Es necesario establecer a priori el número de categorías

Mapas autoorganizados de Kohonen

- **Problemas de agrupación (clasificación no supervisada)**
- **Arquitectura**
 - RNA de dos capas: Entrada y Competición
 - Las neuronas en la capa de competición se pueden organizar:
 - En una recta (mapa unidimensional)
 - En un plano (mapa bidimensional)



Aprendizaje no supervisado

13

Mapas autoorganizados de Kohonen

- n atributos de entrada $E=\{E_1, \dots, E_n\}$
- m neuronas en la capa de competición, entonces los pesos serán:

$$\begin{pmatrix} \mu_{11} & \mu_{12} & \dots & \mu_{1m} \\ \mu_{21} & \mu_{22} & \dots & \mu_{2m} \\ \dots & \dots & \dots & \dots \\ \mu_{n1} & \mu_{n2} & \dots & \mu_{nm} \end{pmatrix}$$

μ_{ij} es el peso de la conexión entre la neurona i de la capa de entrada y la neurona j de la capa de competición

- Cada neurona j ($j=1, \dots, m$) de la capa de competición tiene entonces un vector asociado:

$$\mu_j = \{\mu_{1j}, \mu_{2j}, \dots, \mu_{ij}, \dots, \mu_{nj}\}$$

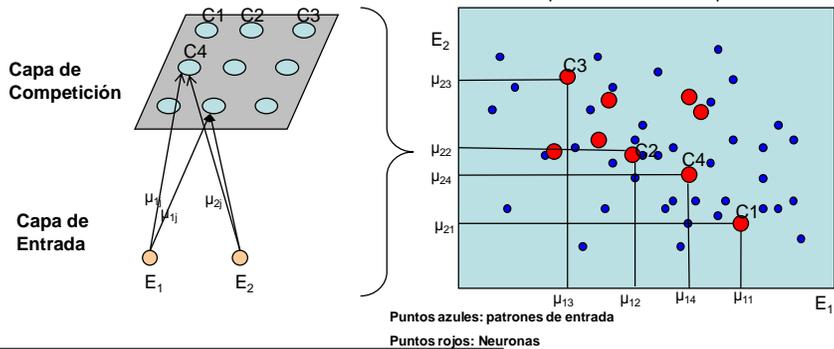
Aprendizaje no supervisado

14

Mapas autoorganizados de Kohonen

- **Representación geométrica**

- Los datos de entrada y las neuronas de la c.c. se pueden representar como puntos en un espacio n dimensional
- Cada neurona de la c.c. se representa en el espacio de entrada utilizando los pesos asociados a esa neurona. (A las neuronas de la c.c. se le llaman también prototipos)



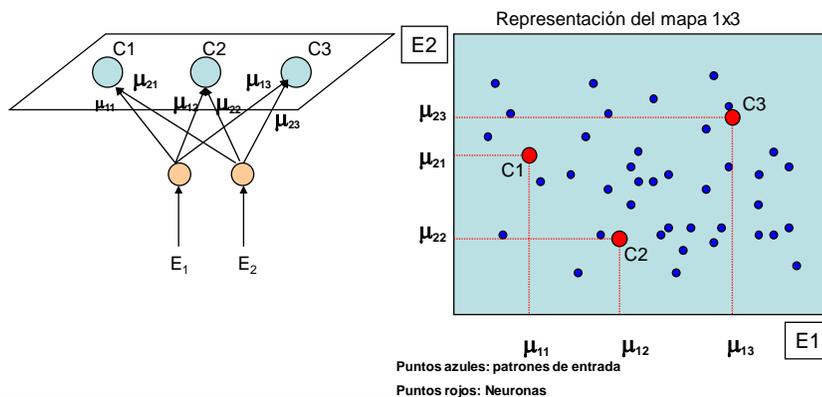
Aprendizaje no supervisado

15

Mapas autoorganizados de Kohonen

- **Representación geométrica**

- Lo mismo para un mapa unidimensional



Aprendizaje no supervisado

16

Mapas autoorganizados de Kohonen

- La **activación** de cada neurona j de la capa de competición (denotada como τ_j) se calcula utilizando la distancia del vector de entrada al vector de pesos asociado a cada neurona:

$$\tau_j = d(E, \mu_j)$$

- Se pueden elegir diversas distancias.
 - El producto escalar: En este caso, la salida de la capa de competición será la habitual en los modelos supervisados (suma ponderada de las entradas por los pesos de las conexiones)

$$\tau_j = d(E, \mu_j) = \sum_{i=1}^n E_i \mu_{ij}$$

- La función más utilizada es la **distancia euclídea**

$$\tau_j = \sqrt{\sum_{i=1}^n (E_i - \mu_{ij})^2}$$

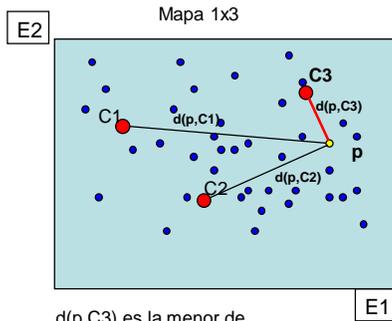
Mapas autoorganizados de Kohonen

- **Funcionamiento**
 - Se recibe el **vector de entrada**
 - Se propaga por las conexiones hasta llegar a la capa de competición
 - Cada neurona de la c.c. produce una salida al comparar la entrada con sus pesos
 - Se selecciona aquella que produzca una **salida más pequeña** (célula ganadora, la más cercana)

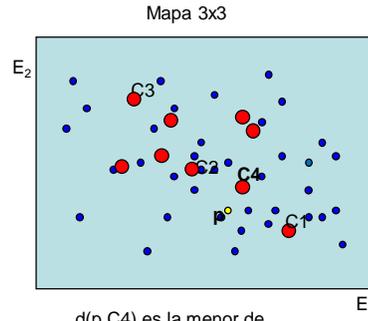
Mapas autoorganizados de Kohonen

• Representación geométrica

- Se calcula la distancia de cada de entrada a las neuronas de la capa de competición (prototipos) y se elige la ganadora:



$d(p,C3)$ es la menor de todas, luego C3 es la neurona ganadora para el punto p



$d(p,C4)$ es la menor de todas, luego C4 es la neurona ganadora para el punto p

Mapas autoorganizados de Kohonen

$$\mu_{ij}(t) = \mu_{ij}(t-1) + \Delta\mu_{ij}$$

$$\Delta\mu_{ij} = \alpha(t) \delta_j(t) (e_i(t) - \mu_{ij}(t))$$

tasa de aprendizaje

Coordenada i de la neurona j $i=1, \dots, n$

$$\delta_j = \begin{cases} 1 & \text{si } j \text{ ganadora} \\ 0 & \text{en caso contrario} \end{cases}$$

si no se tiene en cuenta el vecindario

Mapas autoorganizados de Kohonen

• Tasa de Aprendizaje $\alpha(t)$

- El valor de α es decrementado una cantidad constante pequeña β tras cada ciclo completo de todos los patrones de aprendizaje:

$$\alpha(t+1) = \alpha(t) - \beta$$

Mediante la asignación del parámetro β se pueden determinar el número total de ciclos de aprendizaje que sería:

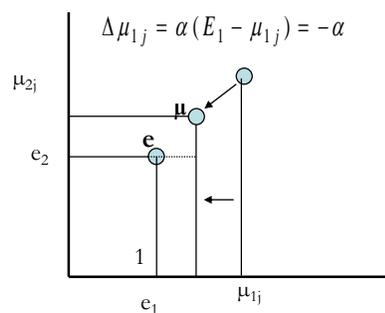
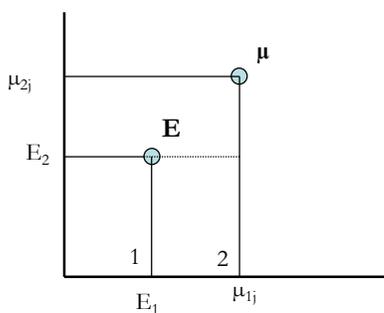
$$\text{Iteraciones} = \frac{\alpha(t)}{\beta}$$

- El valor de α es decrementado siguiendo un esquema logarítmico (grandes variaciones al principio y disminución más lenta después)

Mapas autoorganizados de Kohonen

• Representación geométrica del aprendizaje

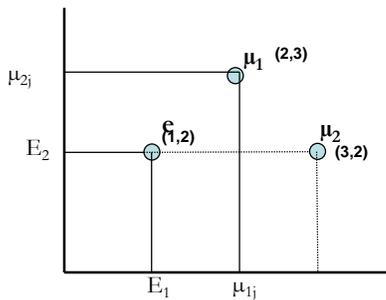
- Se calcula la diferencia, por coordenadas, entre el dato de entrada y la neurona ganadora
- La neurona se acerca al dato de entrada



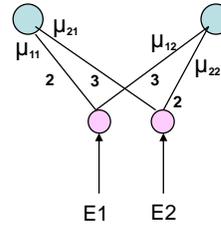
Mapas autoorganizados de Kohonen

Ejemplo

- Dos neuronas $\mu_1=(2,3)$ y $\mu_2=(3,2)$
- Un ejemplo de entrada: $E=(1,2)$



Aprendizaje no supervisado

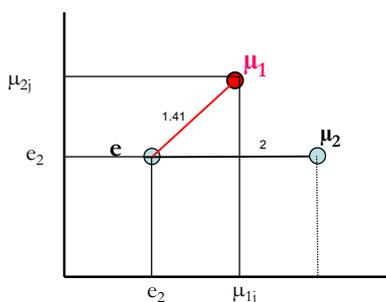


Determinar la neurona ganadora y ajustar sus pesos.
Tasa de aprendizaje 0.2

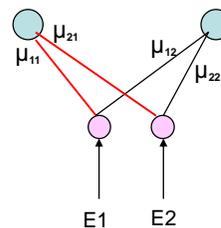
23

Mapas autoorganizados de Kohonen

- Calcular distancias
- $d(e, \mu_1)=1.41$, $d(e, \mu_2)=2$
- μ_1 es la neurona ganadora



Aprendizaje no supervisado



$$\Delta \mu_{ij} = \alpha (e_i - \mu_{ij})$$

$$\Delta \mu_{11} = 0.2(1 - 2) = -0.2$$

$$\Delta \mu_{21} = 0.2(2 - 3) = -0.2$$

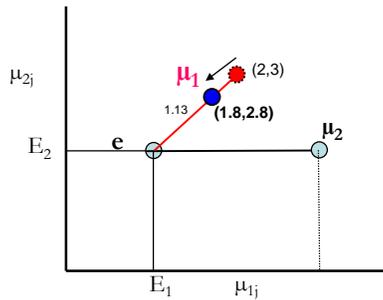
$$\mu_{11} = 2 - 0.2 = 1.8$$

$$\mu_{21} = 3 - 0.2 = 2.8$$

24

Mapas autoorganizados de Kohonen

- Ejemplo: Situación final



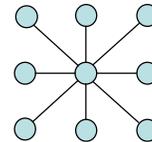
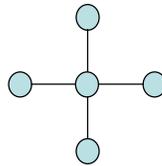
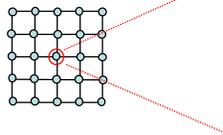
Aprendizaje no supervisado

25

Mapas autoorganizados de Kohonen

Vecindario: generalmente se incluye el concepto de vecindario en el aprendizaje.

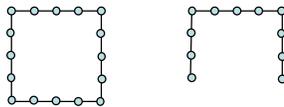
- Vecindarios bidimensionales



Las neuronas de la capa de competición se organizan en una estructura bidimensional. En la figura, cada neurona tiene cuatro vecinos. Podría tener 8 si se consideran las diagonales.

- Vecindarios unidimensionales

- Las neuronas pueden organizarse en una estructura unidimensional, abierta o cerrada



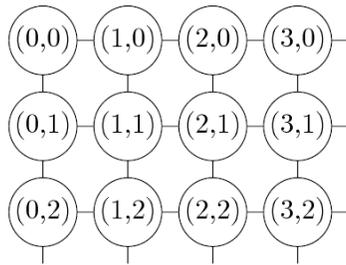
Cada neurona tendrá dos vecinos, cuatro vecinos, ...

Aprendizaje no supervisado

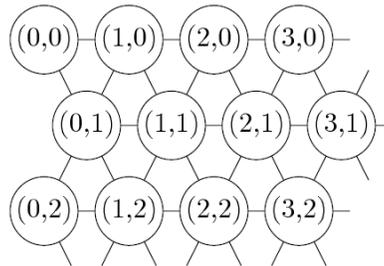
26

Mapas autoorganizados de Kohonen

Vecindarios en SOM_PAK



Rectangular



Hexagonal

Mapas autoorganizados de Kohonen

• Ley de aprendizaje con vecindario

- Se define un término de vecindario entre neuronas i y j $\sigma(c_i, c_j)$

$$\sigma(c_i, c_i) = 1$$

$$\sigma(c_i, c_j) = \text{número de saltos en la capa de competición} + 1$$

- El esquema en el aprendizaje es:

$$\Delta\mu_{ik} = \begin{cases} \frac{\alpha(t)}{\sigma(c_j, c_k)} (e_i(t) - \mu_{ik}(t)) & \begin{array}{l} c_j \text{ ganadora,} \\ \sigma(c_j, c_i) < \theta \end{array} \\ 0 & \text{en caso contrario} \end{cases}$$

$i=1, \dots, n$; k representa el índice de la neurona ganadora y sus vecinas

- Donde θ es el límite del vecindario
- Habitualmente vecindario dinámico que decrece con el tiempo

Mapas autoorganizados de Kohonen

- **Aprendizaje con vecindario**

- Se calcula la diferencia, por coordenadas, entre el dato de entrada y la neurona ganadora
- La neurona ganadora se acerca al dato de entrada
- Las neuronas del vecindario se acercan al dato de entrada en menor medida que la neurona ganadora.
- Mientras más lejos en el vecindario, menos se acercará a la neurona ganadora. La ley es inversamente proporcional a $\sigma(c_i, c_j)$

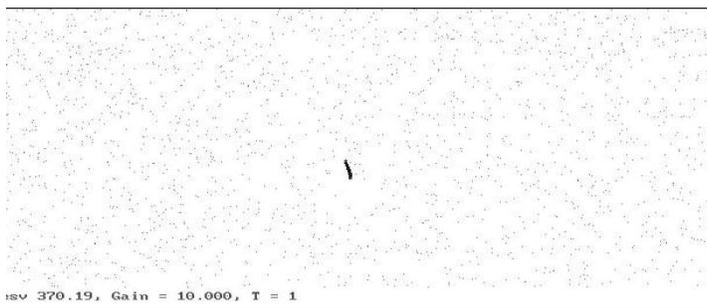
Mapas autoorganizados de Kohonen

- **Método de Kohonen**

1. Inicializar pesos.
Asignar a los pesos valores pequeños aleatorios.
2. Presentar una nueva entrada.
(El conjunto de aprendizaje se presenta cíclicamente hasta llegar a la convergencia de la red. Actualizar α)
3. Propagar el patrón de entrada hasta la capa de competición. Obtener los valores de salida de las neuronas de dicha capa.
4. Seleccionar la neurona ganadora C cuya salida sea menor.
5. Actualizar las conexiones entre la capa de entrada y la neurona C, así como las de su vecindad, según su grado de vecindad.
6. Si α por encima de cierto umbral volver al paso 2, en caso contrario FIN.

Mapas autoorganizados de Kohonen

- **Ejemplo: distribuciones uniformes**
 - Distribución uniforme, iteración 1:

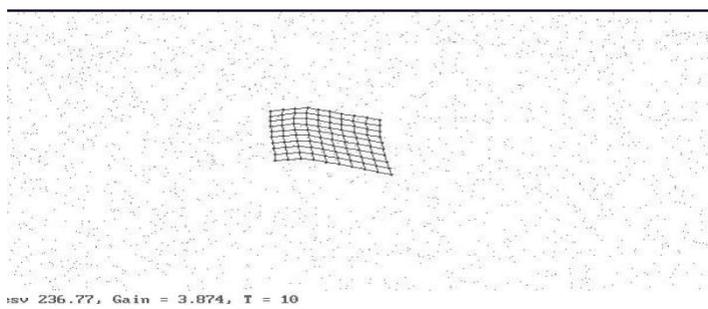


Aprendizaje no supervisado

31

Mapas autoorganizados de Kohonen

- **Ejemplo: distribuciones uniformes**
 - Distribución uniforme, iteración 10:

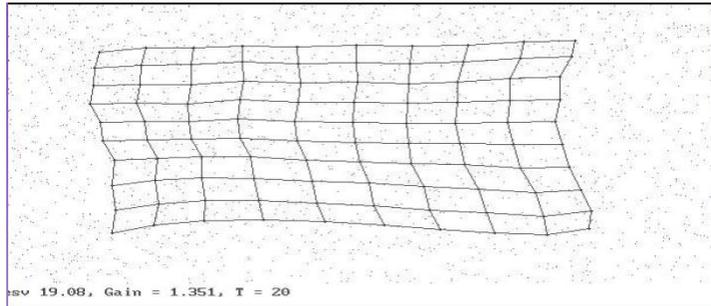


Aprendizaje no supervisado

32

Mapas autoorganizados de Kohonen

- **Ejemplo: distribuciones uniformes**
 - Distribución uniforme, iteración 20:

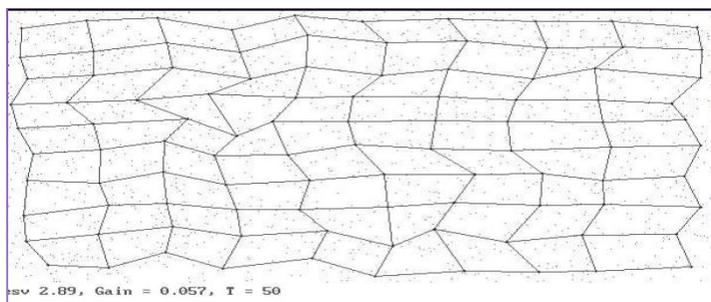


Aprendizaje no supervisado

33

Mapas autoorganizados de Kohonen

- **Ejemplo: distribuciones uniformes**
 - Distribución uniforme, iteración 50:

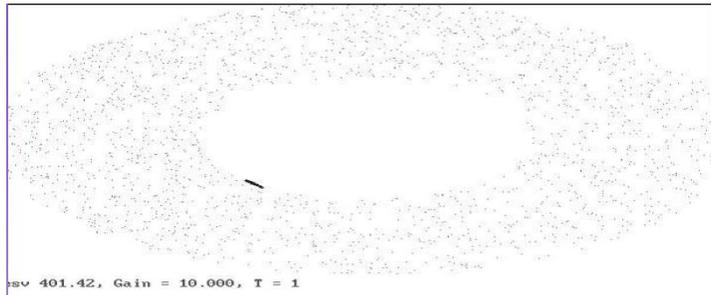


Aprendizaje no supervisado

34

Mapas autoorganizados de Kohonen

- **Ejemplo: distribuciones toroidales**
 - Distribución toroidal, iteración 1:

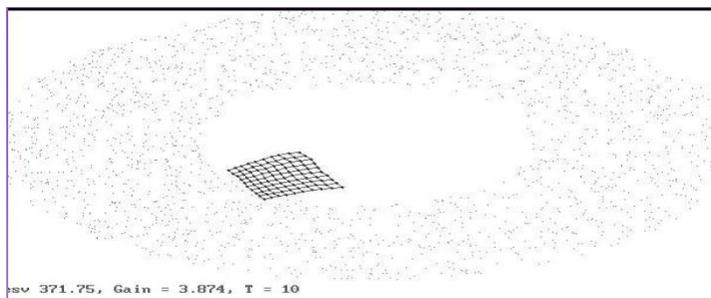


Aprendizaje no supervisado

35

Mapas autoorganizados de Kohonen

- **Ejemplo: distribuciones toroidales**
 - Distribución toroidal, iteración 10:



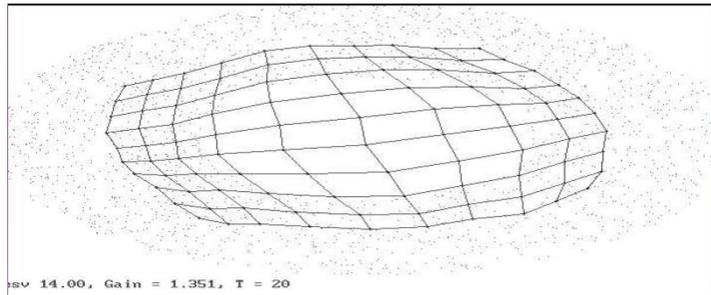
Aprendizaje no supervisado

36

Mapas autoorganizados de Kohonen

- **Ejemplo: distribuciones toroidales**

- Distribución toroidal, iteración 20:



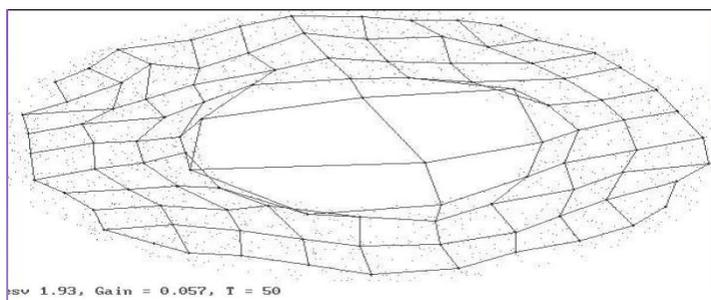
Aprendizaje no supervisado

37

Mapas autoorganizados de Kohonen

- **Ejemplo: distribuciones toroidales**

- Distribución toroidal, iteración 50:



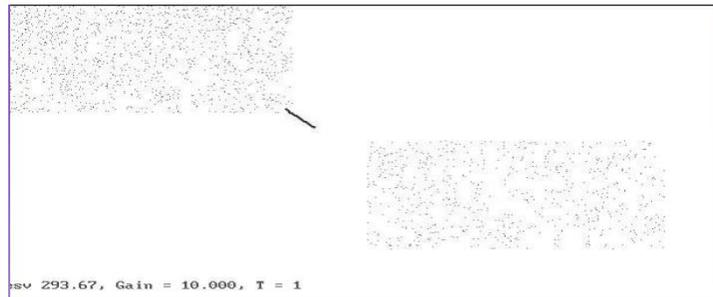
Aprendizaje no supervisado

38

Mapas autoorganizados de Kohonen

- **Ejemplo: distribuciones biuniformes**

- Distribución biuniforme, iteración 1:



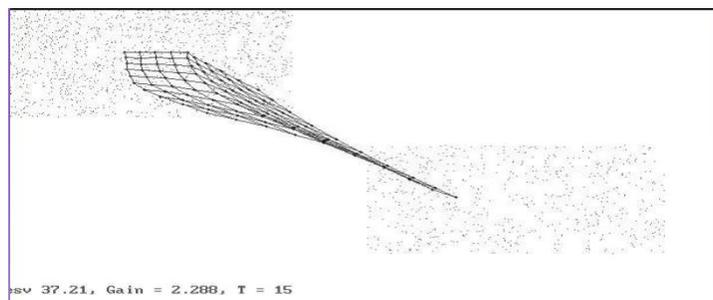
Aprendizaje no supervisado

39

Mapas autoorganizados de Kohonen

- **Ejemplo: distribuciones biuniformes**

- Distribución biuniforme, iteración 15:



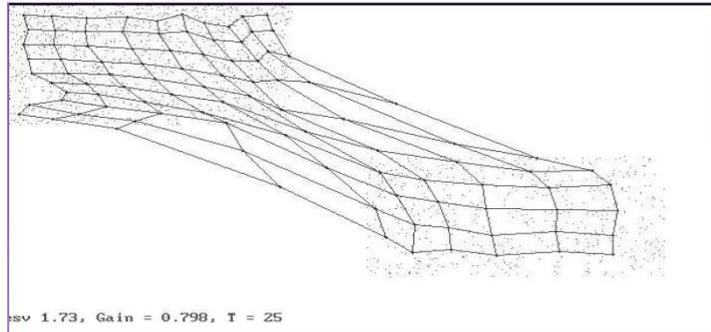
Aprendizaje no supervisado

40

Mapas autoorganizados de Kohonen

- **Ejemplo: distribuciones biuniformes**

- Distribución biuniforme, iteración 25:



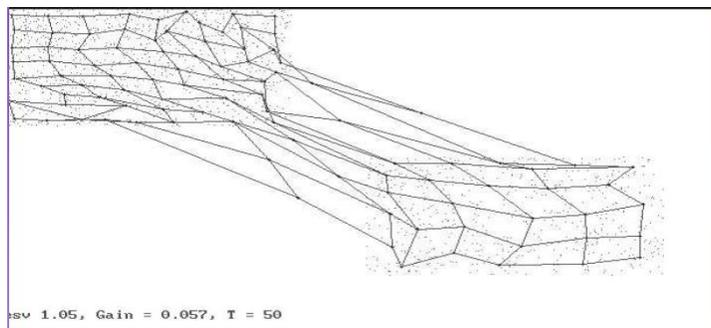
Aprendizaje no supervisado

41

Mapas autoorganizados de Kohonen

- **Ejemplo: distribuciones biuniformes**

- Distribución biuniforme, iteración 50:

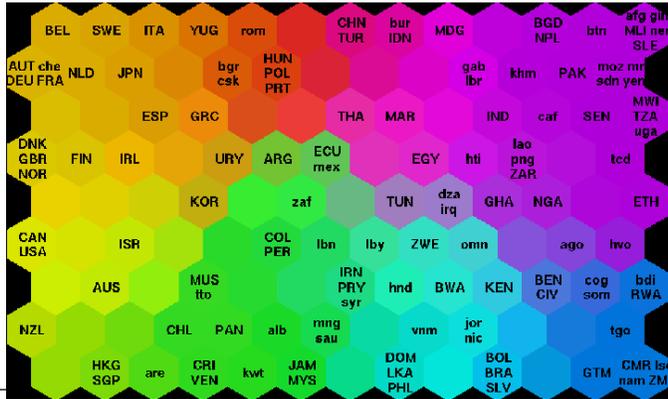


Aprendizaje no supervisado

42

Mapas autoorganizados de Kohonen

- Ejemplo: mapa de pobreza mundial
 - Generado a partir de 39 dimensiones

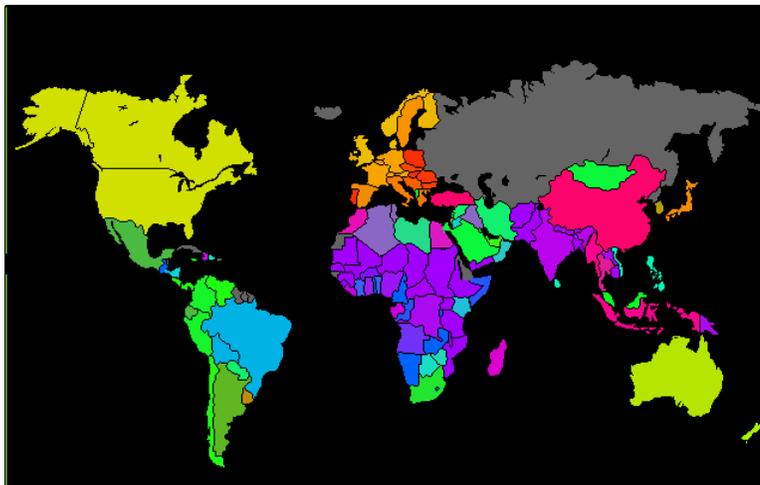


Aprendizaje no supervisado

<http://www.cis.bu.ac.uk/research/som-research/worldmap.html>

43

Mapas autoorganizados de Kohonen



Aprendizaje no supervisado

44