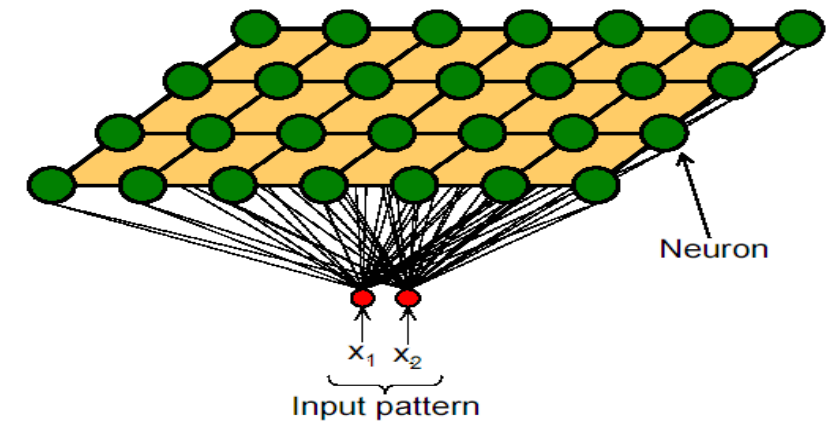


Modelos de Soporte No Supervisado

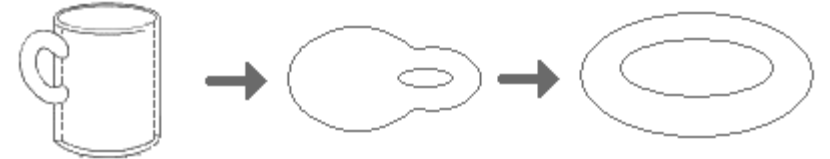
SOM/Kohonen Clustering

SOM/Kohonen

- Introducido por el Prof. Teuvo Kohonen en 1982
- También conocido como mapa de características Kohonen
- Red neuronal no supervisada
- Herramienta de agrupamiento de alta dimensión y datos complejos
- El mapa de auto-organización (SOM) es un algoritmo de red neuronal artificial no supervisado
- Cada peso es representativo de una determinada entrada
- Los patrones de entrada se muestran a todas las neuronas simultáneamente
- Aprendizaje competitivo: se elige la neurona con la respuesta más grande.



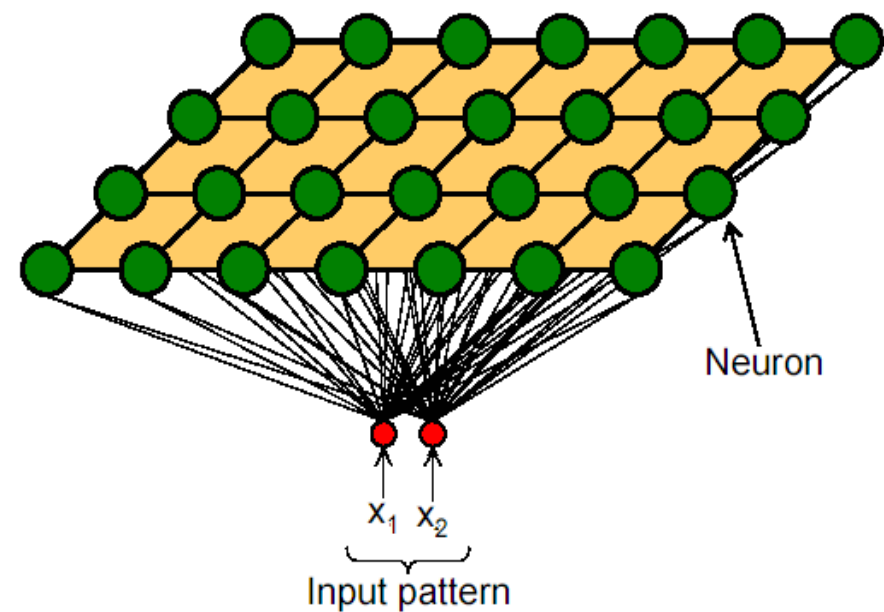
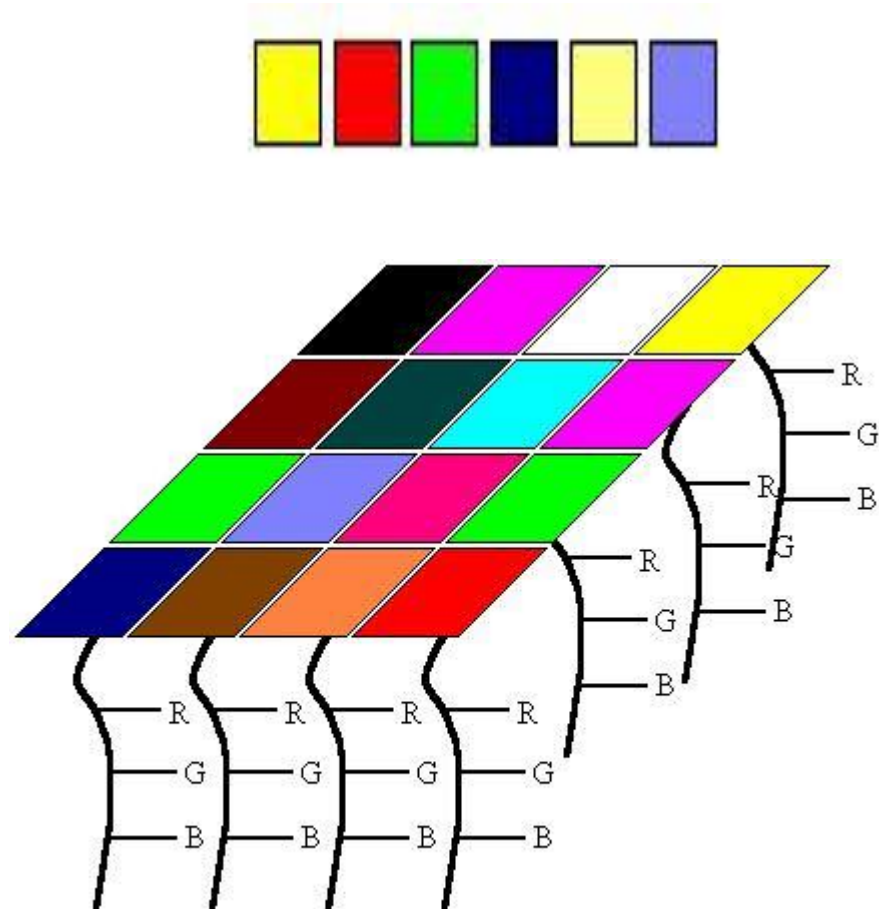
SOM/Kohonen



- Mantiene la topología del conjunto de datos
- El entrenamiento se produce a través de la competencia entre las neuronas
- Imposible asignar nodos de red a clases de entrada específicas por adelantado
- Se puede usar para detectar similitudes y grados de similitud
- Se supone que el patrón de entrada cae en agrupaciones distintas suficientemente grandes
- Inicialización de vector de peso aleatorio

SOM/Kohonen

- Datos de muestra
- Pesos
- Nodos de salida



SOM/Kohonen: Estructura del mapa

- Rejilla bidimensional o unidimensional
- Cada punto de la rejilla representa un nodo de salida
- La rejilla se inicializa con vectores aleatorios

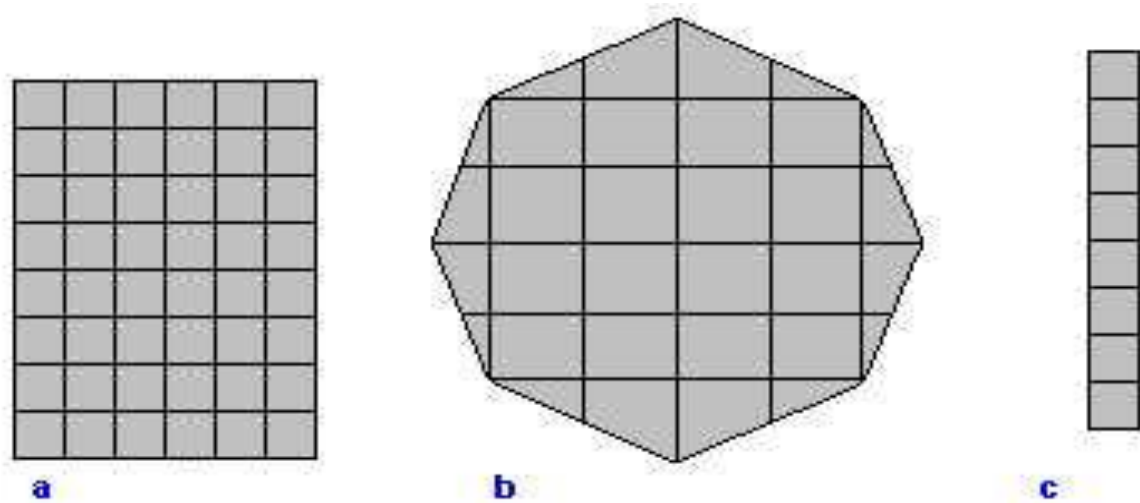
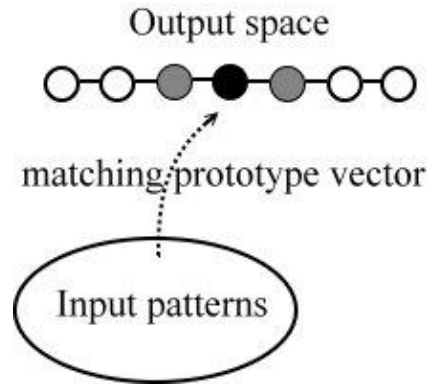


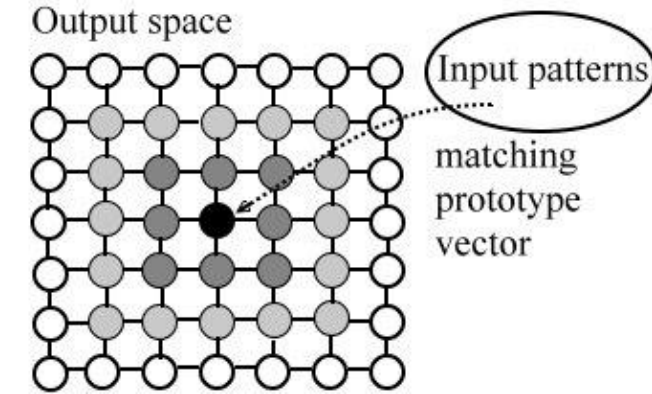
Illustration of Topological Maps

- Illustration of the SOM model with one or two-dimensional map.

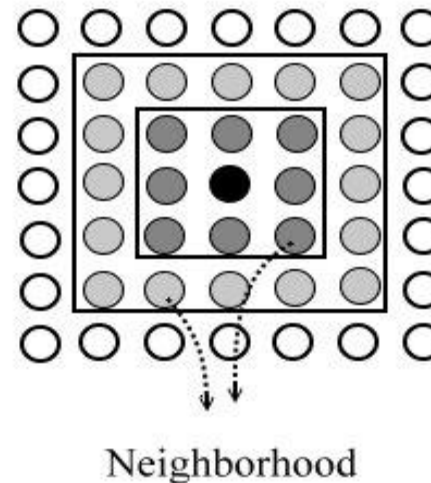
Line:



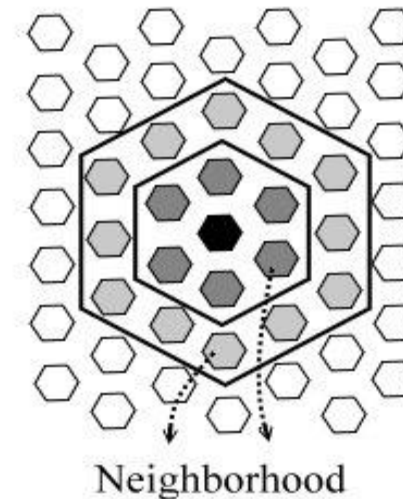
Plane:



Rectangular:



Hexagonal:



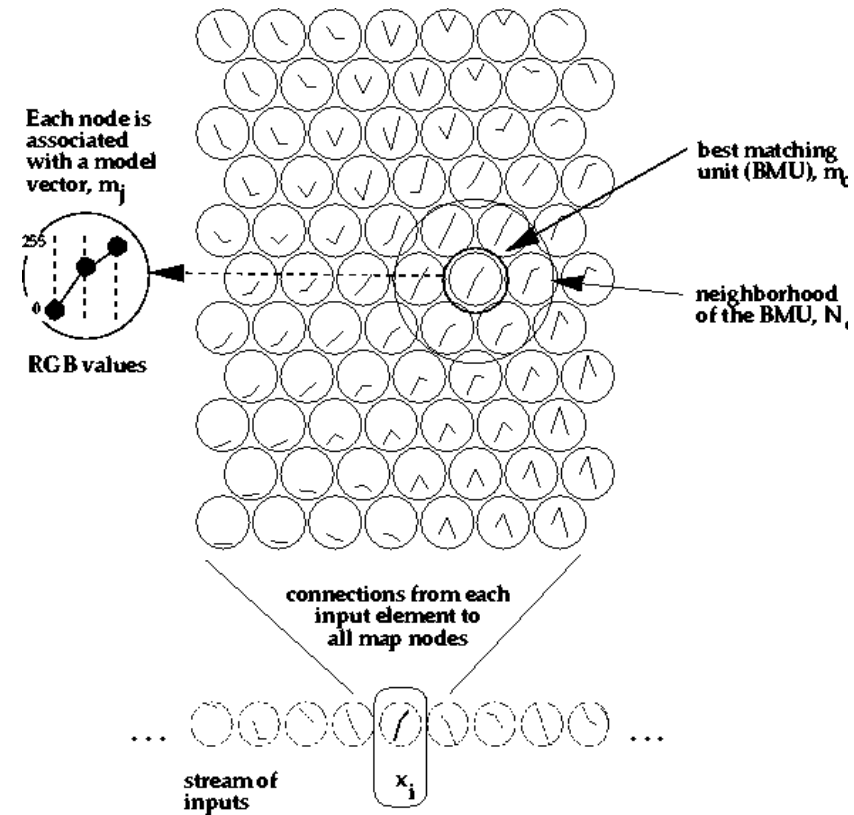
- Example of the SOM model with the rectangular or hexagonal map.

SOM/Kohonen: Algoritmo de entrenamiento

- Inicializar mapa
- Para t de 0 a 1
- Seleccione una muestra
 - Obtener la mejor unidad de coincidencia
 - Escalar vecinos
 - Aumentar t una pequeña cantidad
- End for

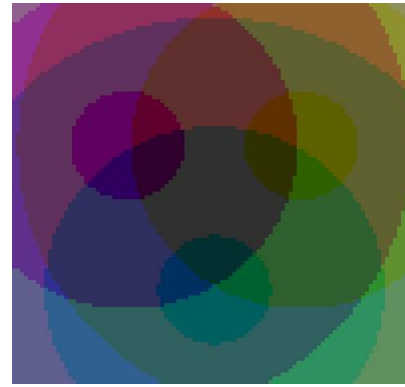
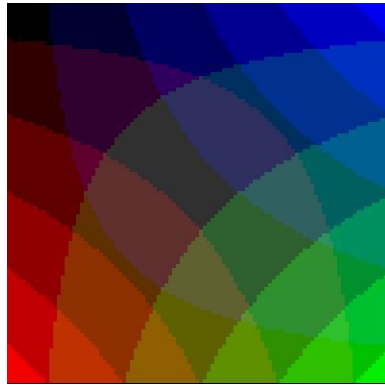
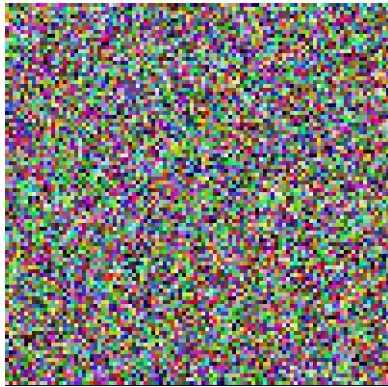
$$m_i(t+1) = m_i(t) + \alpha(t)[x(t) - m_i(t)]$$

$$\forall i \in N_c(t)$$



SOM/Kohonen: Inicializar los pesos

- Los SOM son computacionalmente muy costosos
- Buena inicialización
- Menos iteraciones
- Calidad del mapa



SOM/Kohonen: Obtenga la mejor unidad de emparejamiento

- Cualquier método para la distancia vectorial i. mi.
- Vecino más cercano
- El vecino más lejano
- Distancia entre los medios
- Distancia entre medianas
- El método más común es la distancia euclidiana.

$$\sqrt{\sum_{i=0}^n \mathbf{x}_i^2}$$

Algorithm for Kohonen's SOM

- Let the map of size M by M , and the weight vector of neuron i is \mathbf{m}_i
- Step 1: Initialize all weight vectors $\mathbf{m}_i(0)$ *randomly or systematically*.
- Step 2: A vector \mathbf{x} is randomly chosen from the training data.
Then, compute the Euclidean distance d_i between \mathbf{x} and neuron i .

$$d_i = \| \mathbf{x} - \mathbf{m}_i(t) \|, \quad 1 \leq i \leq M^2$$

- Step 3: Find the best matching neuron (winning node) c .

$$d_c = \| \mathbf{x} - \mathbf{m}_c(t) \| = \min\{ \| \mathbf{x} - \mathbf{m}_i(t) \| \}, \quad \forall i$$

- Step 4: Update the weight vectors of the winning node c and its neighborhood as follows.

$$\mathbf{m}_i(t+1) = \mathbf{m}_i(t) + \alpha(t) h_{c,i}(t) [\mathbf{x} - \mathbf{m}_i(t)]$$

where $0 \leq \alpha(t) \leq 1$ is an adaptive function which decreases with time.

- Step 5: Iterate the Step 2-4 until the sufficiently accurate map is acquired.

SOM/Kohonen: Vecinos de escala

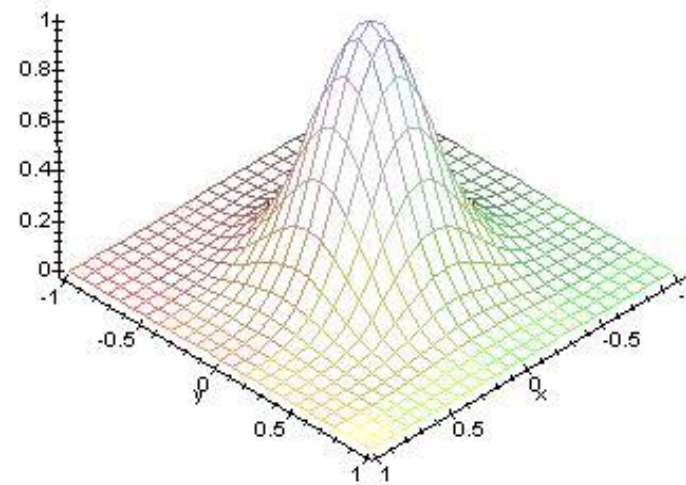
- La determinación de los vecinos
- Tamaño del vecindario
- Disminuye con el tiempo
- Efecto en los vecinos
- Aprendizaje

$$\forall i \in N_c(t),$$

$$m_i(t+1) = m_i(t) + \alpha(t)[x(t) - m_i(t)]$$

otherwise ,

$$m_i(t+1) = m_i(t)$$



Neighborhood Kernel

- The $h_{c,i}(t)$ is a neighborhood kernel centered at the winning node c , which decreases with time and the distance between neurons c and i in the topology map.

$$h_{c,i}(t) = \exp\left(-\frac{\|r_c - r_i\|^2}{\sigma^2(t)}\right)$$

where r_c and r_i are the coordinates of neurons c and i .

The $\sigma(t)$ is a suitable decreasing function of time,

e.g.
$$\sigma(t) = \sigma_0 \left(1 - \frac{t}{\text{maxstep}}\right)$$

SOM/Kohonen: Ventajas y desventajas

- Muy fácil de entender
- Funciona bien
- Costoso computacionalmente
- Cada SOM es diferente



SOM/Kohonen: Prueba de convergencia

- Prueba completa solo para una dimensión
- Muy trivial
- Casi todas las pruebas parciales se basan en Cadenas de Markov

Dificultades :

- No hay definición para "Una configuración ordenada correctamente"
- Resultado comprobado: no es posible asociar una "función potencial de disminución global" con este algoritmo.