Trabajo Práctico N°4 Aprendizaje No Supervisado

Grupo 5

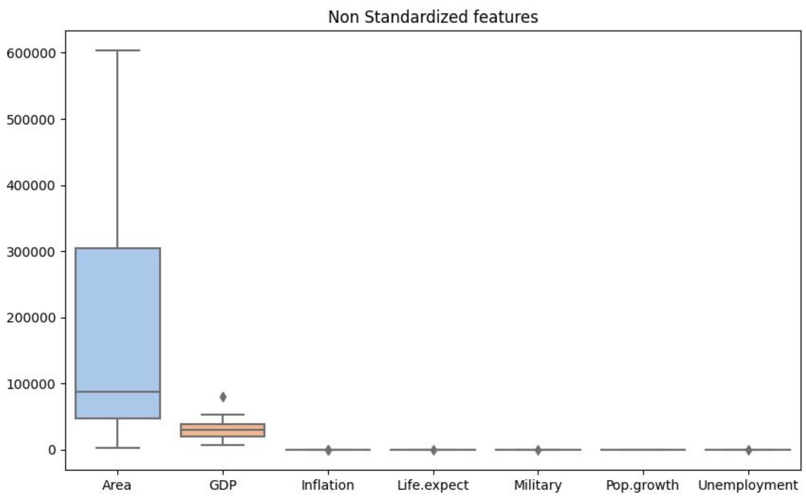
Gonzalo Baliarda
Franco Nicolás Estevez
Ezequiel Agustin Perez
Leandro Ezequiel Rodriguez
Lucas Agustín Vittor

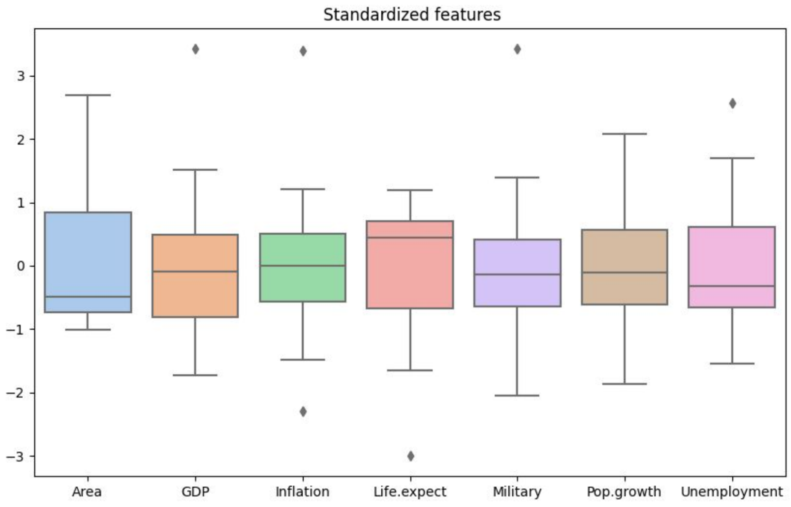
Ejercicio 1

Países de Europa.

Dataset

Características económicas, sociales y geográficas de 28 países de Europa.

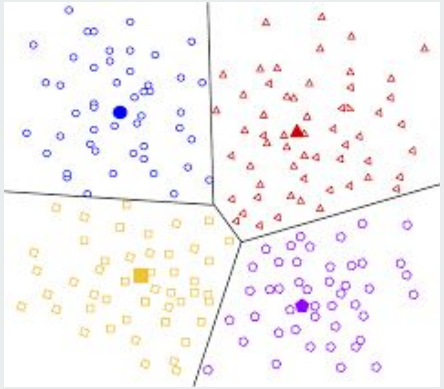


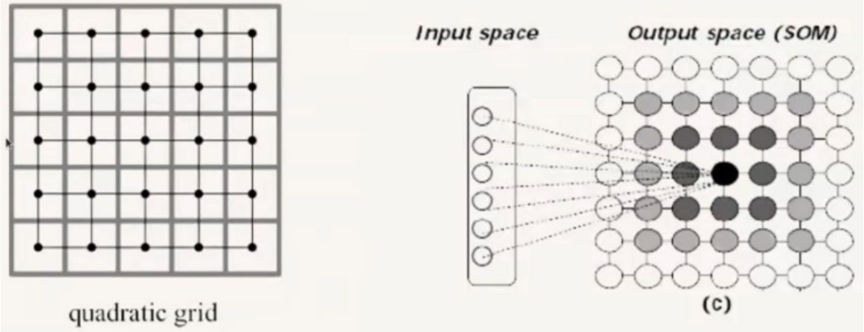


Red de Kohonen

Red de Kohonen

Permite identificar patrones en los inputs, y agruparlos según estos patrones.





Parámetros del modelo

- Inicialización de pesos con samples aleatorios de los inputs.
- Radio de vecinos (mapa de k.k neuronas):

$$R_0 = \frac{k}{2} \qquad R(i) = R_0 * (1 - \frac{i}{epochs}) \qquad R_{min} = 1$$

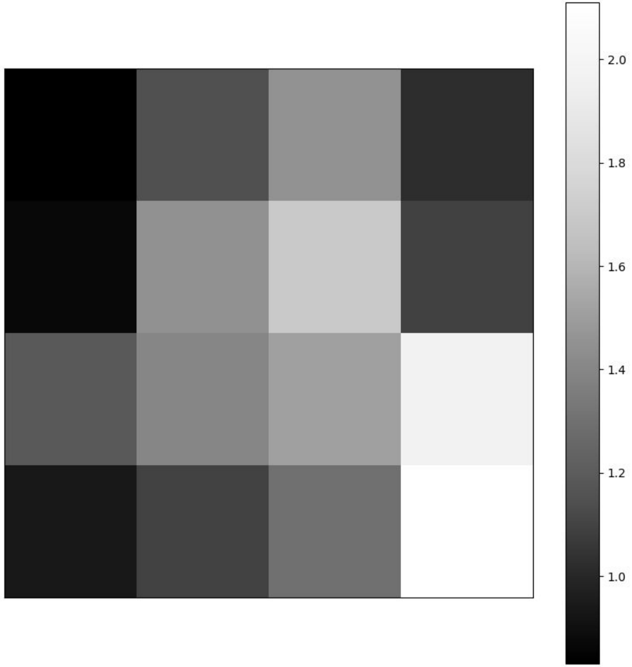
- Tasa de aprendizaje:

$$\eta_0 = 1 \qquad \eta(i) = \eta_0 * (1 - \frac{i}{epochs})$$

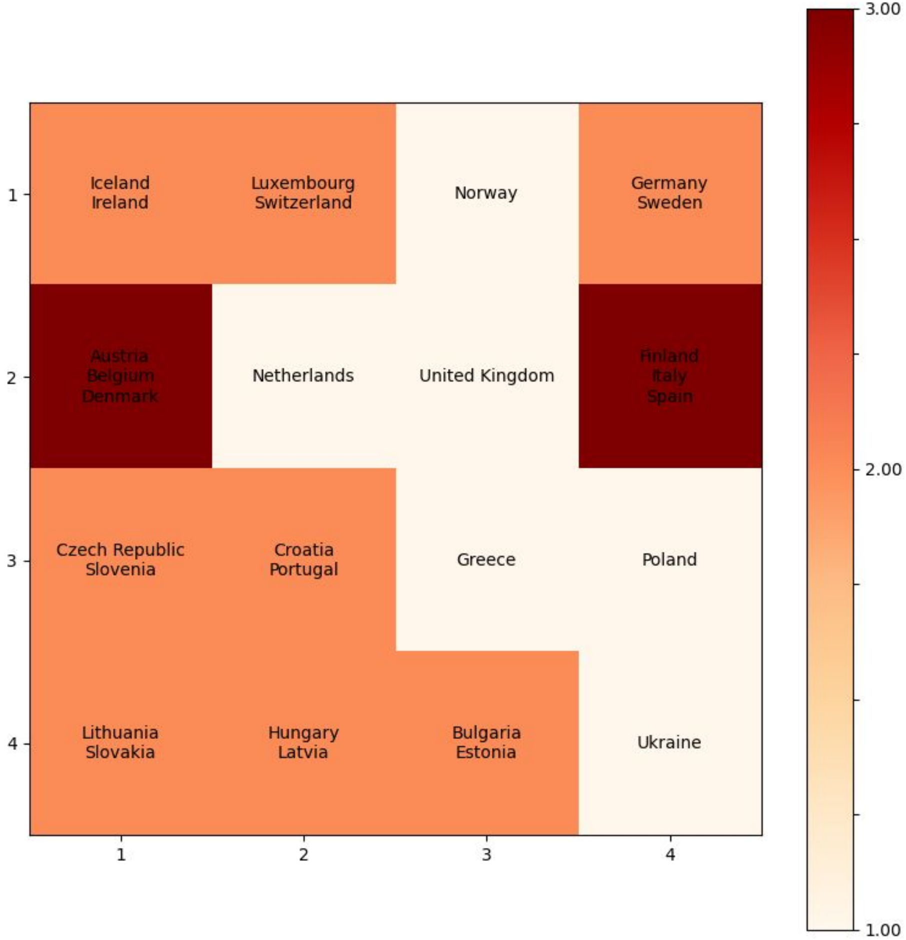
Entrenamiento

- Similitud euclídea.
- Cantidad de épocas: 10.000
- Mapa de $k^2 = 16$ neuronas.
- Se selecciona un input al azar en cada época.

Resultados



2.0
1.8
1.6
1.4
1.2
1.0

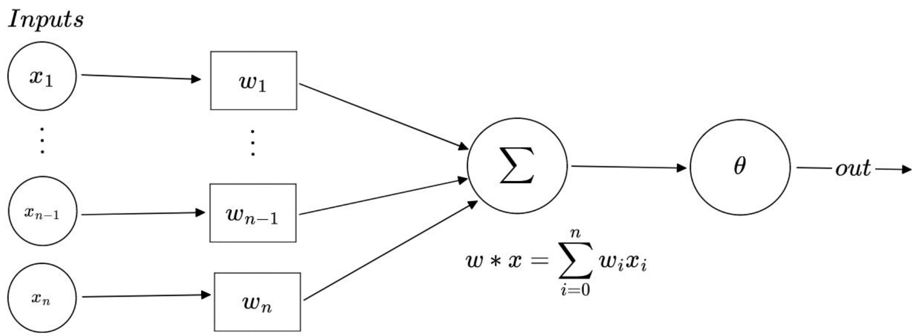


Iceland
Ireland
Luxembourg
Switzerland
Norway
Germany
Sweden
Austria
Belgium
Denmark
Netherlands
United Kingdom
Finland
Italy
Spain
Czech Republic
Slovenia
Croatia
Portugal
Greece
Poland
Lithuania
Slovakia
Hungary
Latvia
Bulgaria
Estonia
Ukraine
1
2
3
4
3.00
2.00
1.00

Modelo de Oja

Red de Oja

Permite el aprendizaje no supervisado de las conexiones sinápticas de una red neuronal



Regla de Oja

$$\Delta w = \eta O^{\mu} x^{\mu}$$

$$w_j^{n+1} = \frac{w_j^n + \Delta w}{(\sum_{j=1}^{N} (w_j^n + \Delta w)^2)^{0.5}}$$

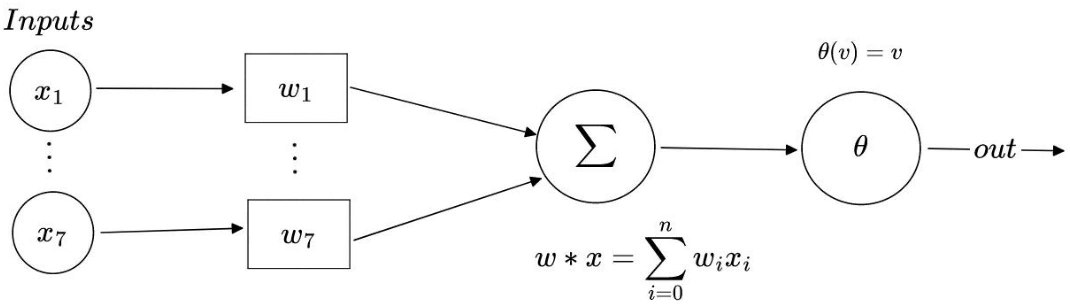
Perceptrón Lineal

$$\Delta w = \eta(\zeta^{\mu} - O^{\mu}) x^{\mu}$$

$$w_j^{n+1} = w_j^n + \Delta w$$

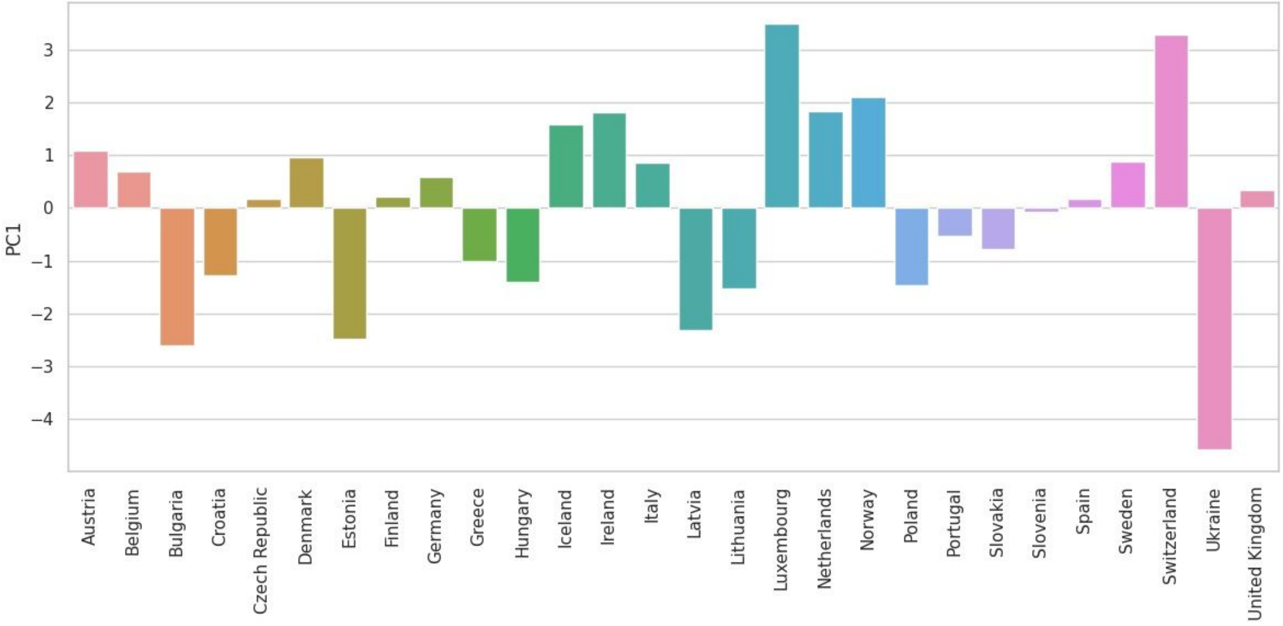
Red de Oja: arquitectura

$$\eta = 0.001$$



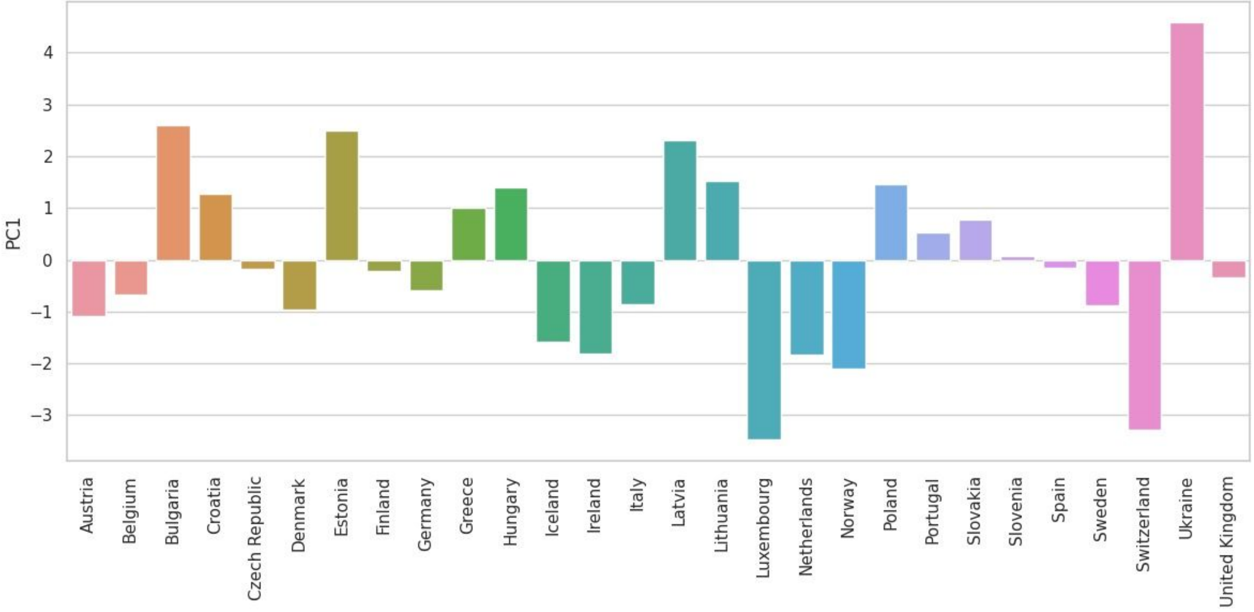
Componente Principal

- 500 iteraciones



Componente Principal librería

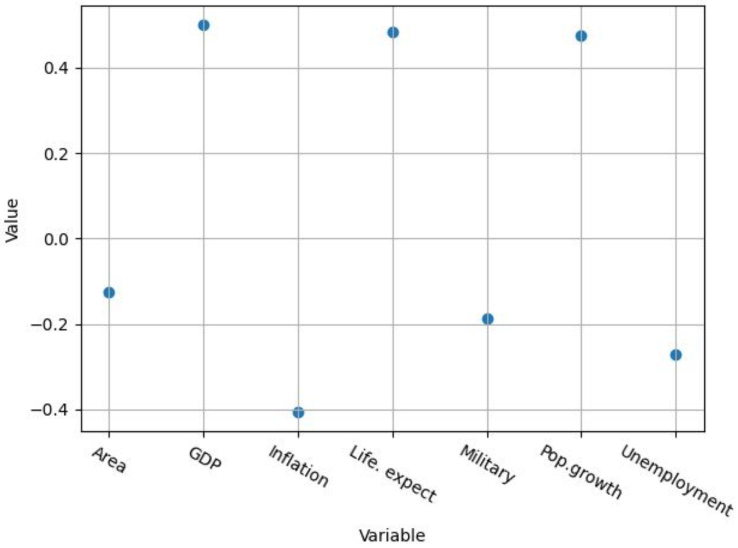
- Librería scikit-learn



Componente Principal

Implementación	
Variable	**Valor**
Area	-0.1248739
GDP	0.50050586
Inflation	-0.40651815
Life.expect	0.48287333
Military	-0.18811162
Pop.growth	0.47570355
Unemployment	-0.27165582

Librería scikit-learn	
Variable	**Valor**
Área	0.124874
GDP	-0.500506
Inflation	0.406518
Life.expect	-0.482873
Military	0.188112
Pop.growth	-0.475704
Unemployment	0.271656

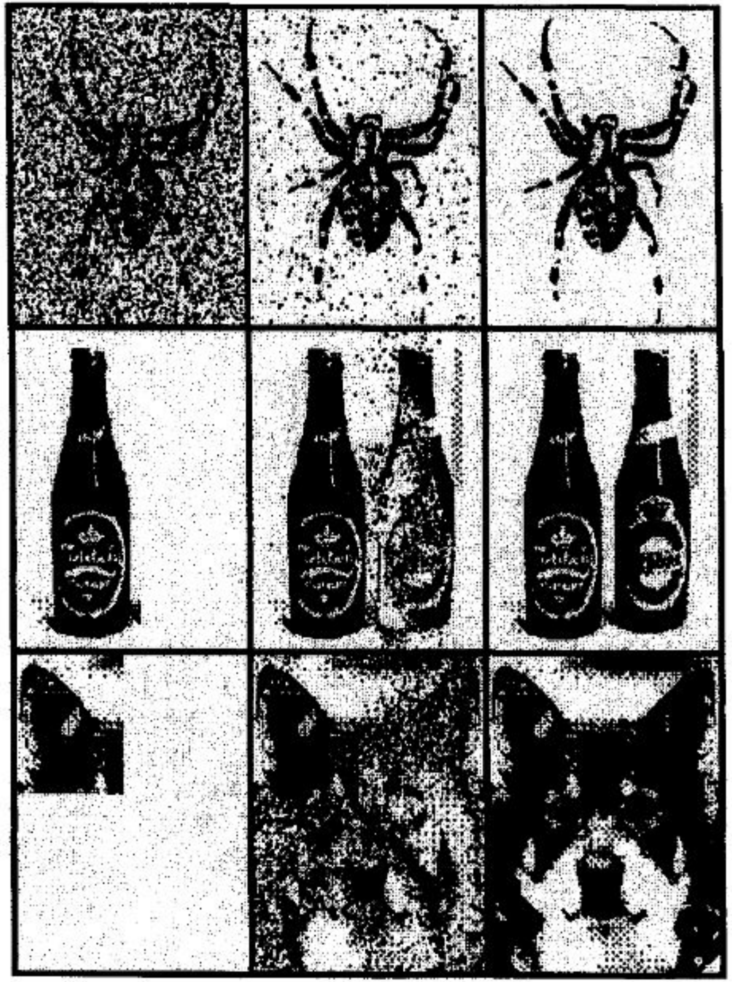


Ejercicio 2

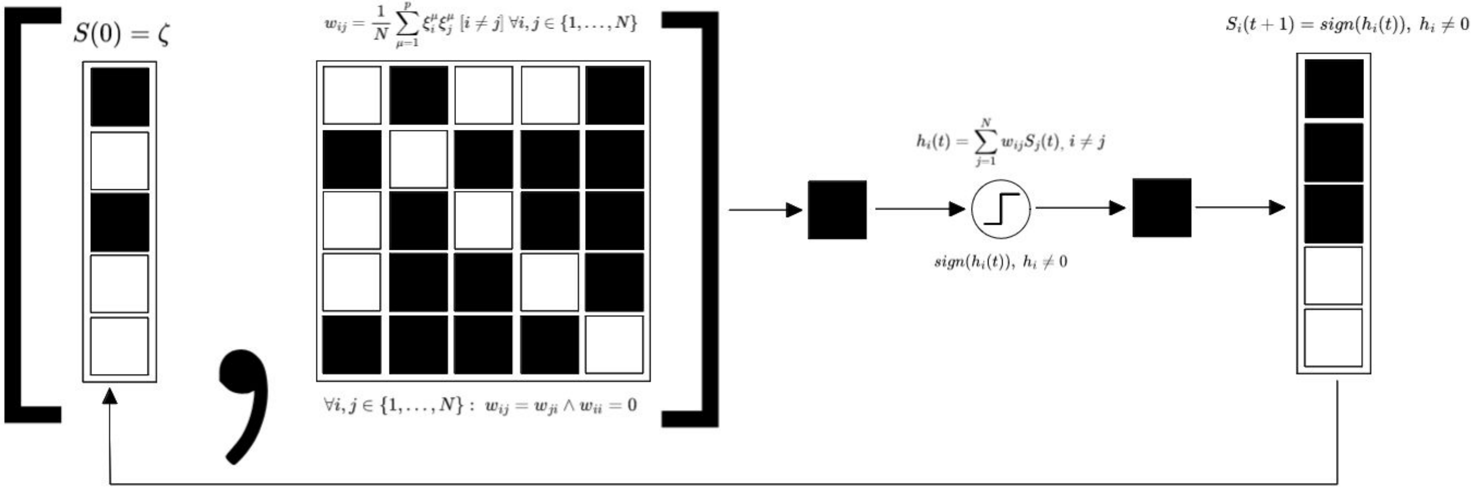
Reconstrucción de patrones.

Red de Hopfield

Permite asociar un patrón de consulta binario (con perturbaciones) con alguno de los patrones almacenados



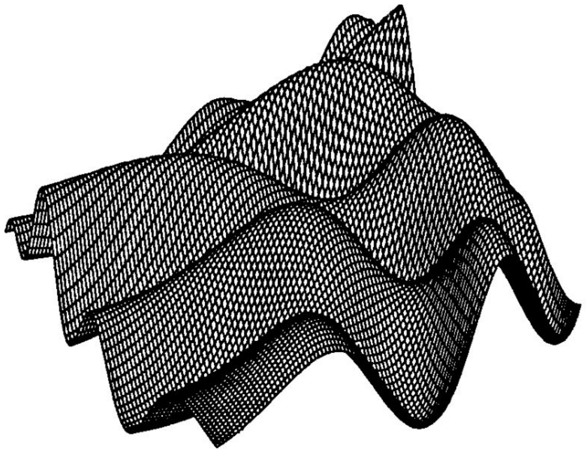
Modelo

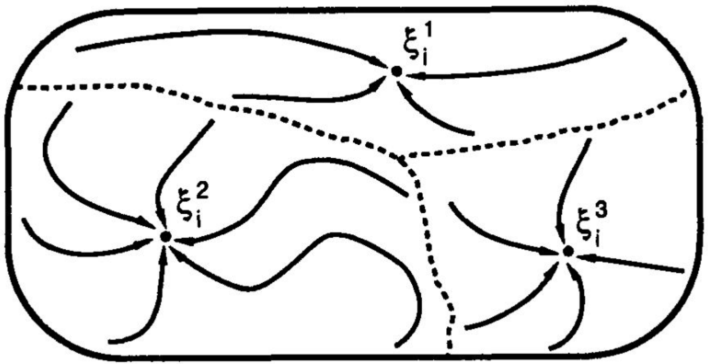


$$\text{Stable Network} \iff S_i = S_{i-1} \; \forall i \in \{1, \ldots, N\}$$

$$\text{Network Energy: } H(w, t) = -\sum_{j>i} w_{ij} S_i(t) S_j(t)$$

Estados espurios

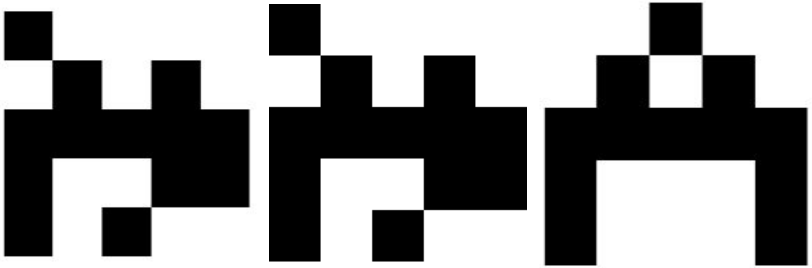




$$\text{Network Energy: } H(w,t) = -\sum_{j>i} w_{ij} S_i(t) S_j(t)$$

Convergencia

$$\text{Patterns} = \{"a", "j", "k", "s"\}$$



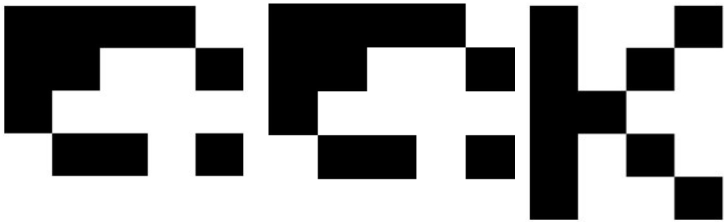
$$\text{Noise} = 0.2$$

$$\text{Network Energy: } H(w,t) = -\sum_{j>i} w_{ij} S_i(t) S_j(t)$$

t	H(w, t)
0	-2.560
1	-11.359

Convergencia

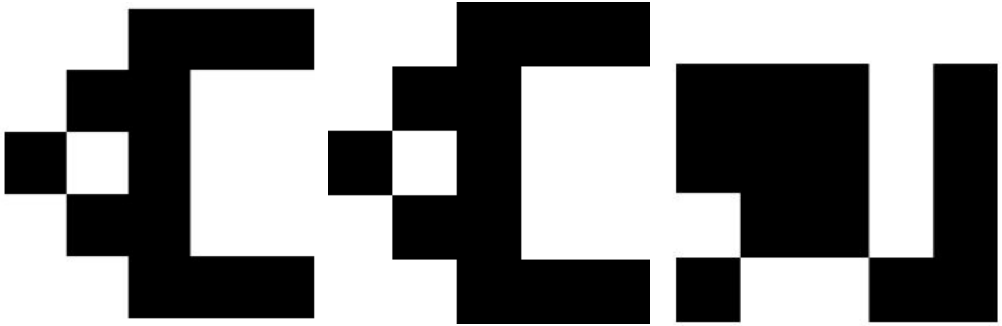
$$\text{Patterns} = \{\text{"a"}, \text{"j"}, \text{"k"}, \text{"s"}\}$$



$$\text{Noise} = 0.6$$

t	H(w, t)
0	1.120
1	-1.600
2	-6.720
3	-6.880
4	-9.359
5	-11.199

Invertir letras



$$\text{Patterns} = \{\text{"a"}, \text{"j"}, \text{"k"}, \text{"s"}\}$$

Spurious States

We have shown that the Hebb prescription (2.9) gives us (for small enough p) a dynamical system that has attractors—local minima of the energy function—at the desired points ξ_i^μ. These are sometimes called the **retrieval states**. But we have *not* shown that these are the only attractors. And indeed there are others.

First of all, the reversed states $-\xi_i^\mu$ are minima and have the same energy as the original patterns. The dynamics and the energy function both have a perfect symmetry, $S_i \leftrightarrow -S_i$ for all i. This is not too troublesome for the retrieved patterns; we could agree to reverse all the remaining bits when a particular "sign bit" is -1 for example.

Second, there are stable **mixture states** ξ_i^{mix}, which are not equal to any single pattern, but instead correspond to linear combinations of an odd number of patterns [Amit et al., 1985a]. The simplest of these are symmetric combinations of three stored patterns:

$$\xi_i^{\text{mix}} = \operatorname{sgn}(\pm\,\xi_i^{\mu_1} \pm \xi_i^{\mu_2} \pm \xi_i^{\mu_3})\,. \tag{2.32}$$

All eight sign combinations are possible, but we consider for definiteness the case where all the signs are chosen as $+$'s. The other cases are similar. Observe that on average ξ_i^{mix} has the same sign as $\xi_i^{\mu_1}$ three times out of four; only if $\xi_i^{\mu_2}$ and $\xi_i^{\mu_3}$ both have the opposite sign can the overall sign be reversed. So ξ_i^{mix} is Hamming distance $N/4$ from $\xi_i^{\mu_1}$, and of course from $\xi_i^{\mu_2}$ and $\xi_i^{\mu_3}$ too; the mixture states lie at points equidistant from their components. This also implies that $\sum_i \xi_i^{\mu_1}\xi_i^{\text{mix}} = N/2$ on average. Now to check the stability of (2.32), still with all $+$ signs, we can repeat the calculation of (2.11) and (2.12), but this time pick out the three special μ's:

$$h_i^{\text{mix}} = \frac{1}{N}\sum_{j\mu} \xi_i^\mu \xi_j^\mu \xi_j^{\text{mix}} = \frac{1}{2}\xi_i^{\mu_1} + \frac{1}{2}\xi_i^{\mu_2} + \frac{1}{2}\xi_i^{\mu_3} + \text{cross-terms}\,. \tag{2.33}$$

Thus the stability condition (2.10) is indeed satisfied for the mixture state (2.32). Similarly 5, 7, ... patterns may be combined. The system does not choose an *even* number of patterns because they can add up to zero on some sites, whereas the units have to be ± 1.

Third, for large p there are local minima that are not correlated with any finite number of the original patterns ξ_i^μ [Amit et al., 1985b]. These are sometimes called **spin glass states** because of a close correspondence to spin glass models in statistical mechanics. We will meet them again in Section 2.5.

So the memory does not work perfectly; there are all these additional minima in addition to the ones we want. The second and third classes are generally called **spurious minima**. Of course we only fall into one of them if we start close to it, and they tend to have rather small basins of attraction compared to the retrieval states. There are also various tricks that we will consider later, including finite temperature and biased patterns, that can reduce or remove the spurious minima.

Limitaciones

$$h_i^\nu = \xi_i^\nu + \frac{1}{N}\sum_i \sum_{\mu \neq \nu} \xi_i^\mu \xi_j^\mu \xi_j^\nu$$

$$\{\xi_i^\mu\}_1^p \text{ quasi-orthogonals}$$

$$C \equiv \frac{p}{N} \leq 0.15$$

Ejercicio

$$C \equiv \frac{4}{5 * 5} = 0.16 \nleq 0.15$$

Gracias