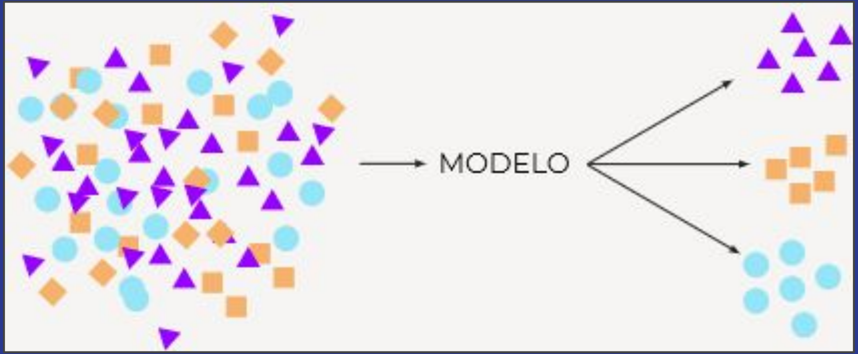
Trabajo Práctico N°4
Aprendizaje No Supervisado



Grupo 7:

- Andres Podgorny
- Hugo Lichtenberger
- Ouss Slaitane
- Nicolás Birsa
- Valentin Ye Li

Ej1 Paises europeos

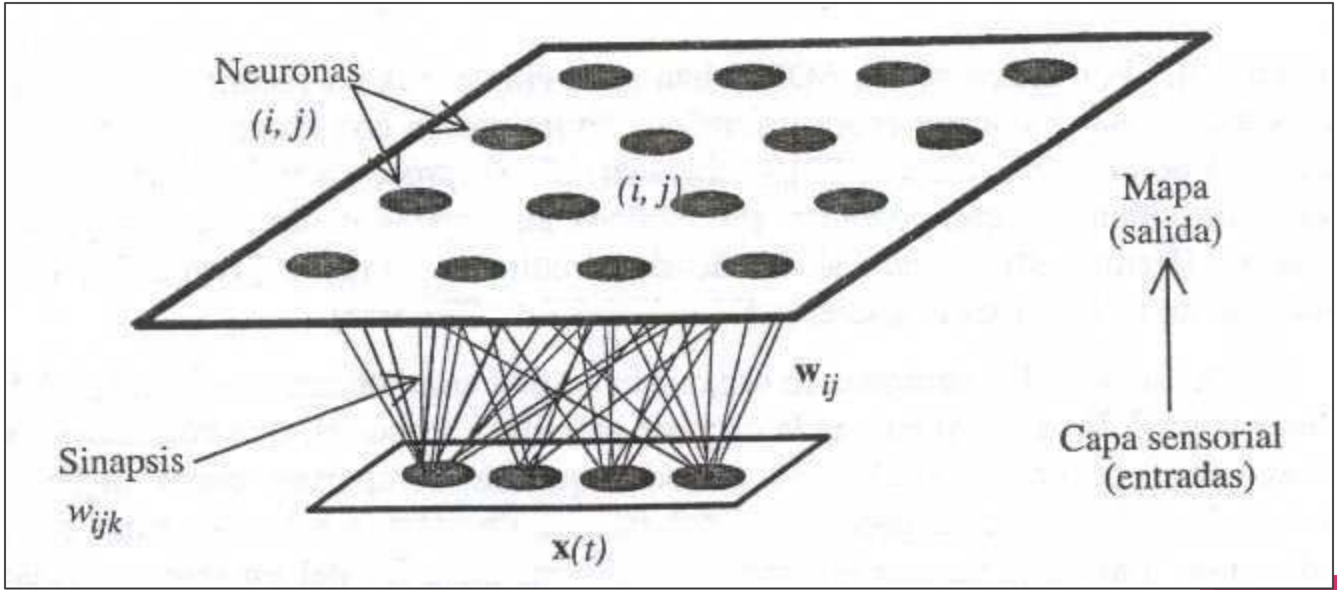
Consigna

Dada una lista de 28 países de Europa con datos de cada país correspondientes a sus características económicas, sociales y geográficas, se busca

- Estudiar su asociación usando Kohonen **(1.A)**

- Analizar las componentes principales usando Oja **(1.B)**

1.A Modelo de Kohonen



Variación del radio y eta

Se parte de un valor inicial, para luego reducirlos a lo largo de las ejecuciones hasta aproximarse a un valor final cercano al mínimo aceptable.

$$\eta(i) = \eta_0 \frac{N-i}{N} \quad siendo\; \eta(0) < 1$$

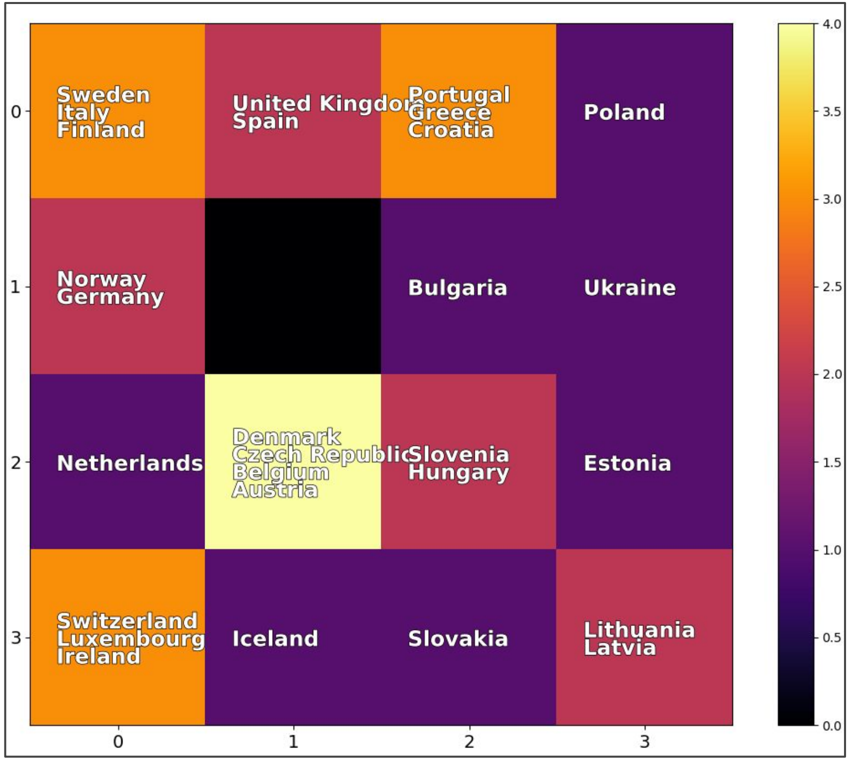
$$r(i) = r_0 + (1 - r_0)\frac{i}{N}$$

con

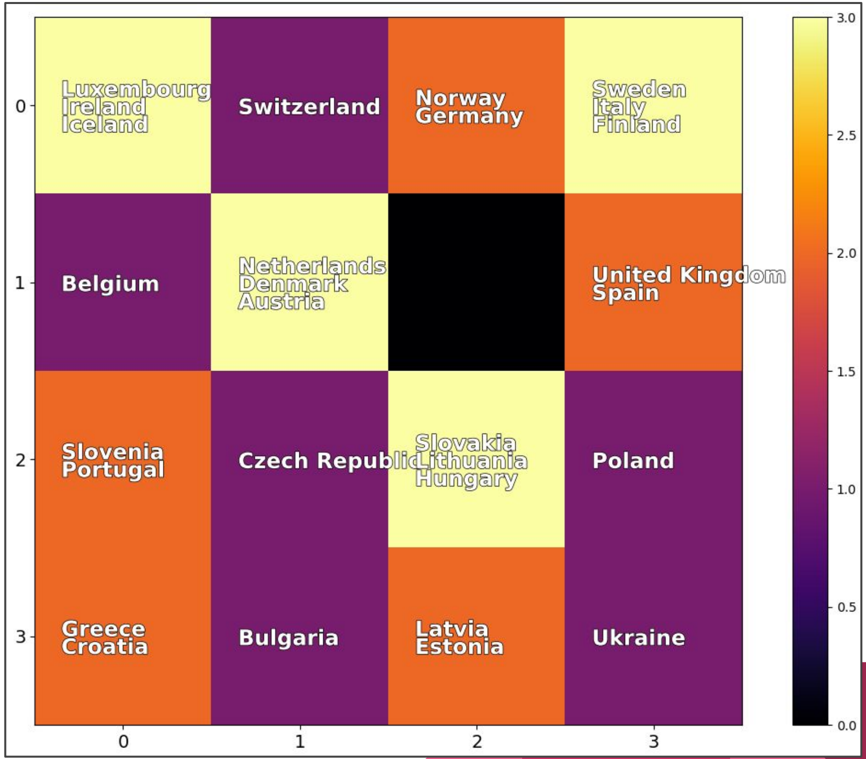
$$N = Número\ de\ iteraciones$$

$$i = Iteración\ actual$$

Pesos aleatorios

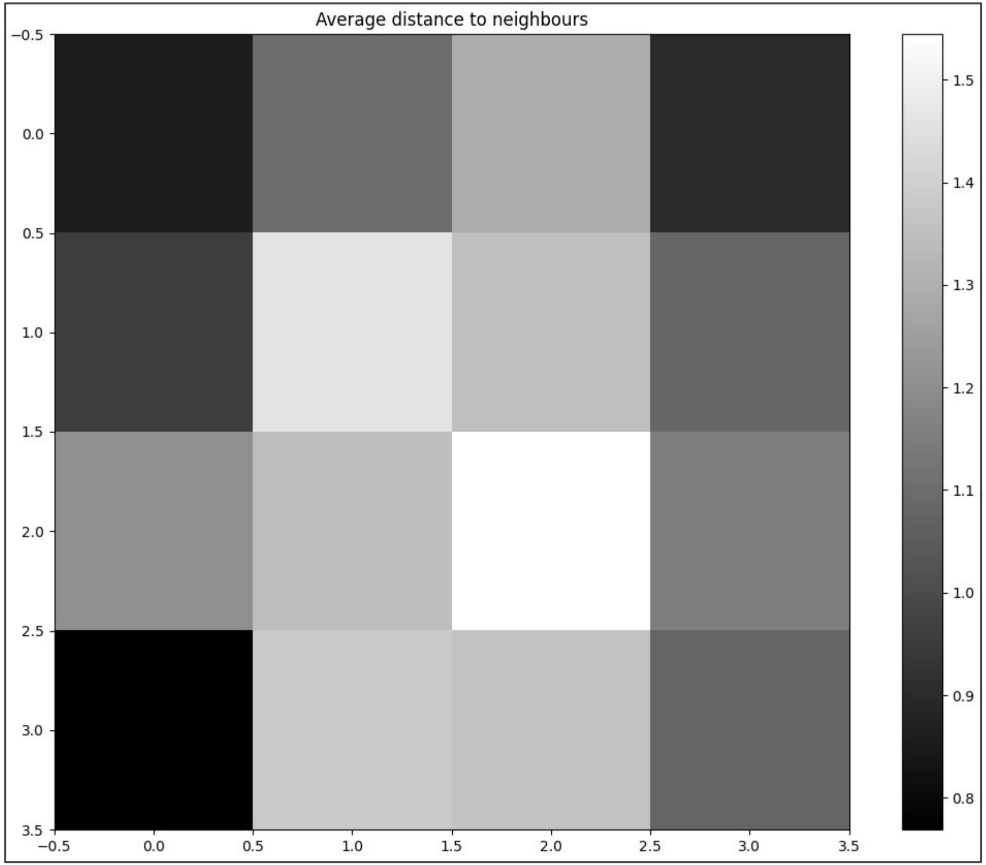


Pesos del input



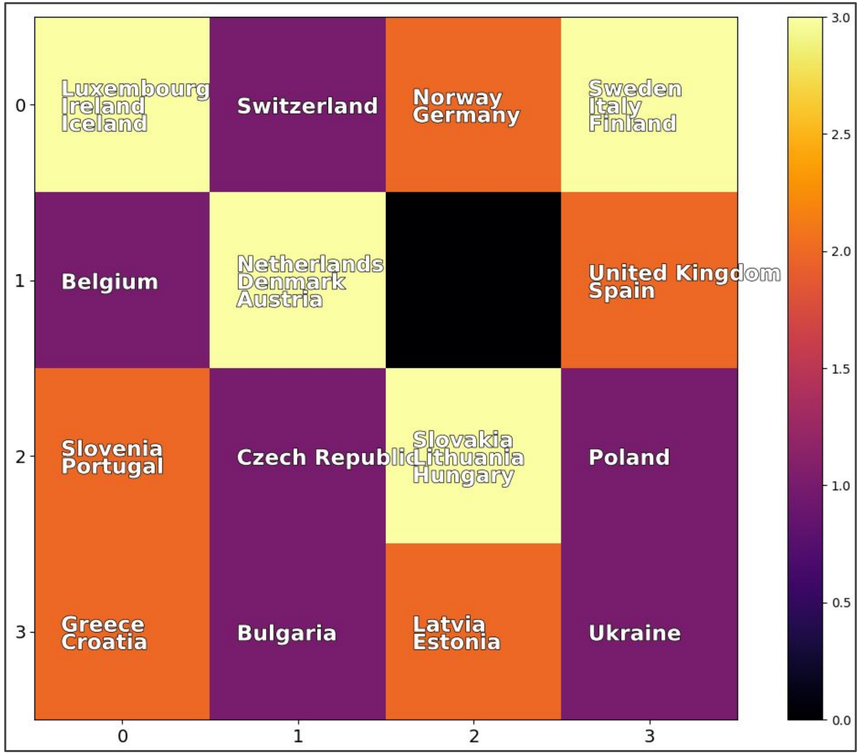
r_0=1.8 | η_0=0.9 | Grilla 4x4 | 500 épocas

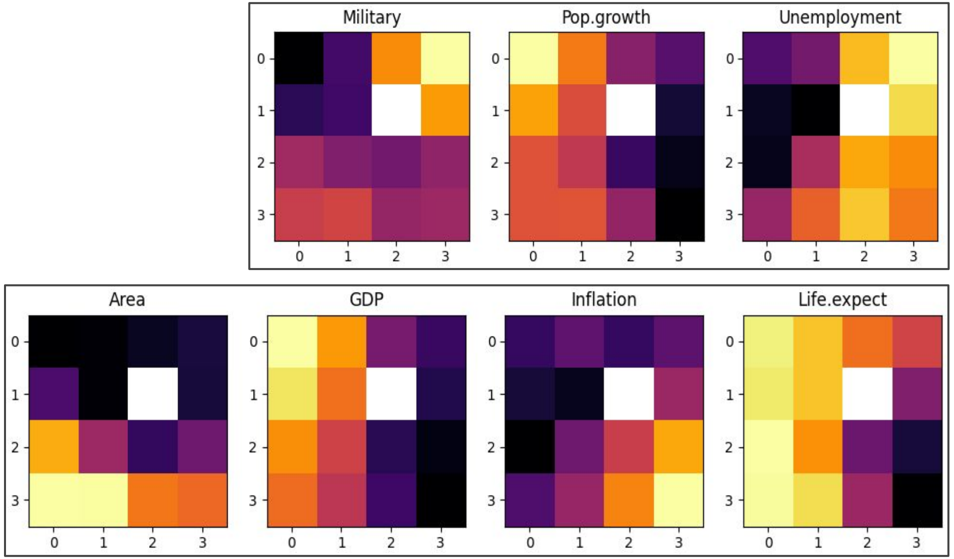
Distancias entre neuronas vecinas



η_0=0.9
r_0=1.8
500 épocas
Grilla 4x4
Con pesos iniciales del input

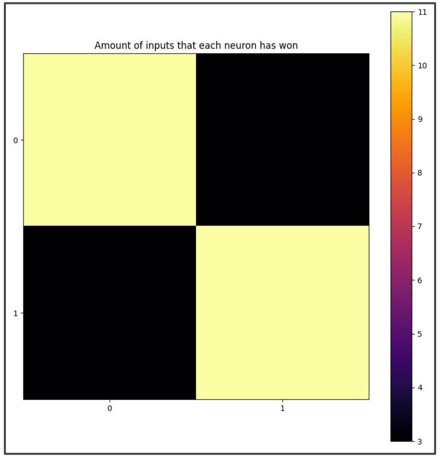
Promedio de las variables



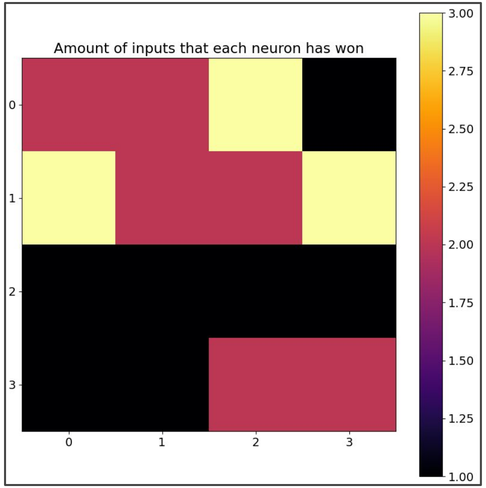


η_0=0.9
r_0=1.8
500 épocas
Grilla 4x4
Con pesos iniciales del input

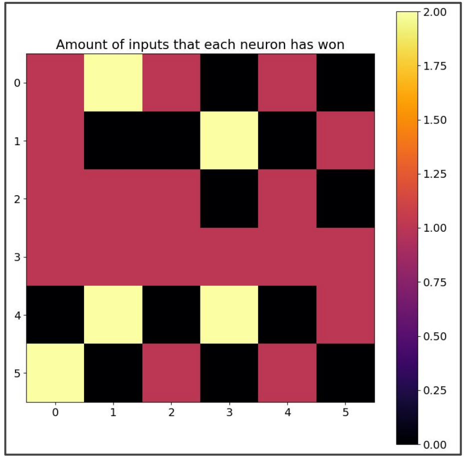
Tamaño de la grilla



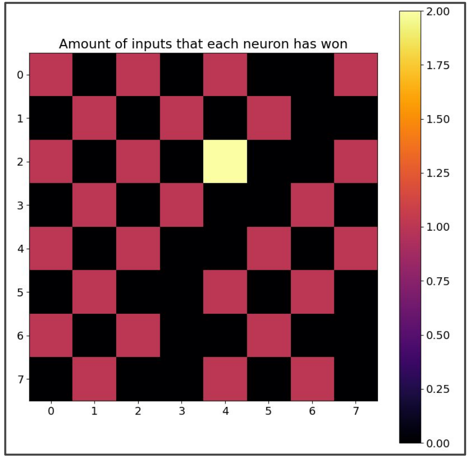
2x2



4x4



6x6



8x8

η_0=0.9
r_0=1.8
500 épocas
Con pesos iniciales del input

Para 2x2 hay un aglutinamiento de países en 2 neuronas, mientras que con 6x6 apenas se pueden agrupar países de a pares con varias neuronas muertas.

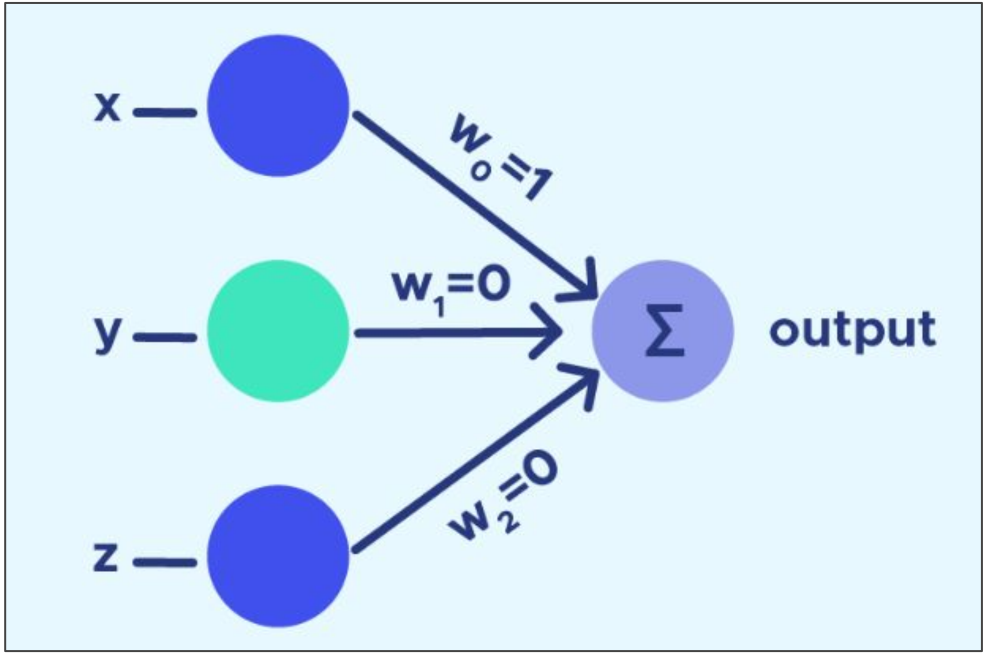
4x4 es un buen balance entre ambos casos.

Conclusiones

- Para el caso analizado, un tamaño de grilla **4x4** logra agrupar países con características similares sin caer en que éstos estén o muy juntos o muy dispersos.

- Inicializar la grilla con valores iniciales no asegura que, al finalizar el algoritmo, no hayan neuronas muertas.

1.B Oja

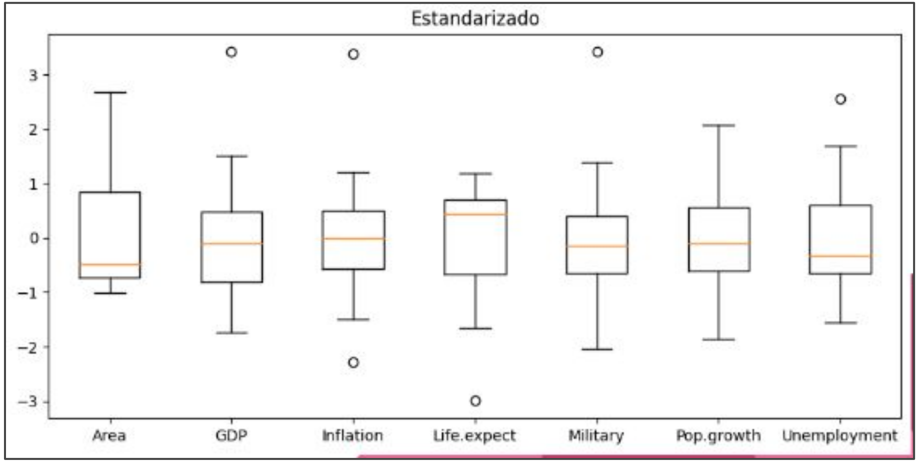


$$y(\mathbf{x}) = \sum_{j=1}^{m} x_j w_j$$

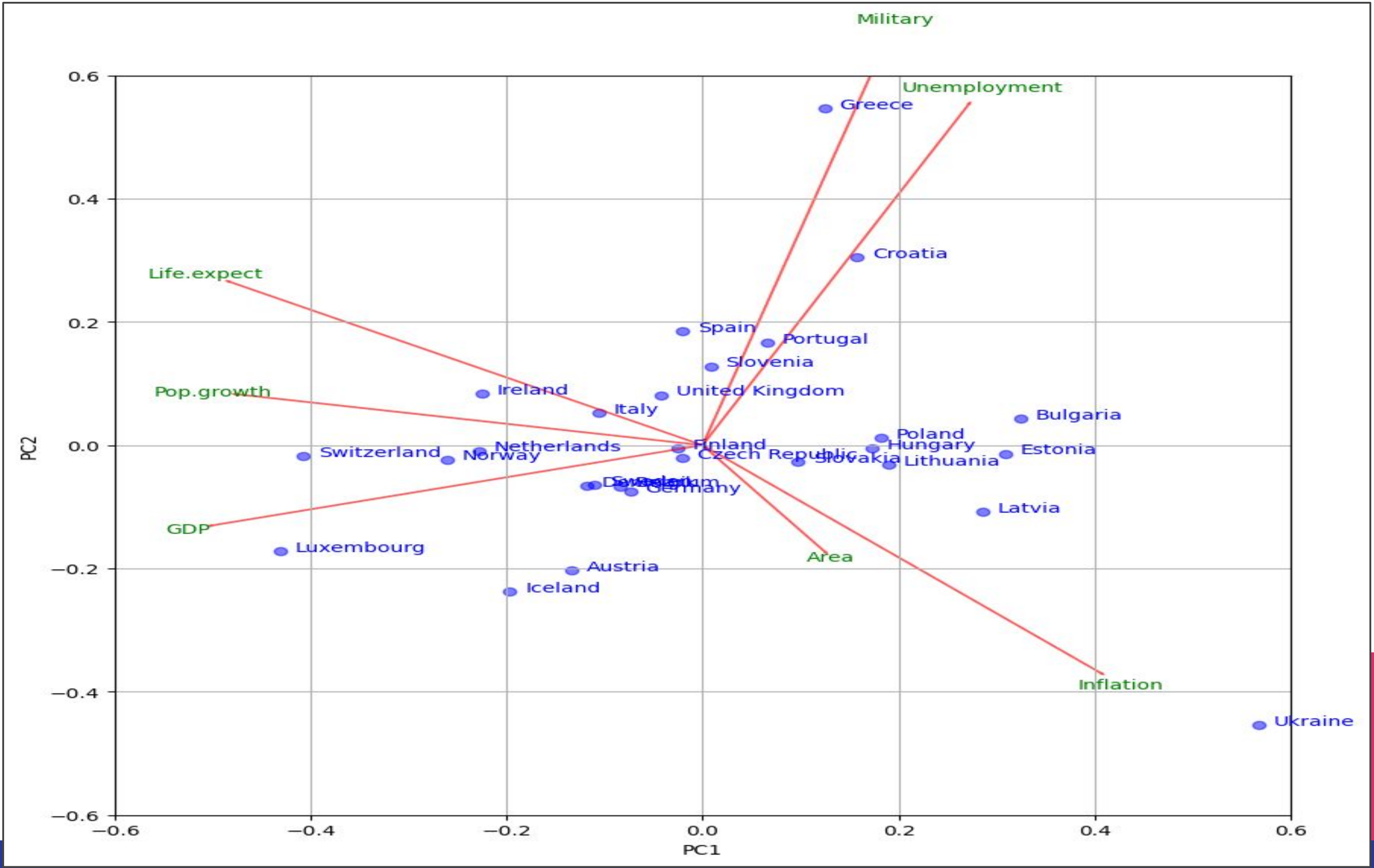
$$\Delta \mathbf{w} = \mathbf{w}_{n+1} - \mathbf{w}_n = \eta\, y_n (\mathbf{x}_n - y_n \mathbf{w}_n)$$

SKLearn

Variables con y sin estandarización



SKLearn



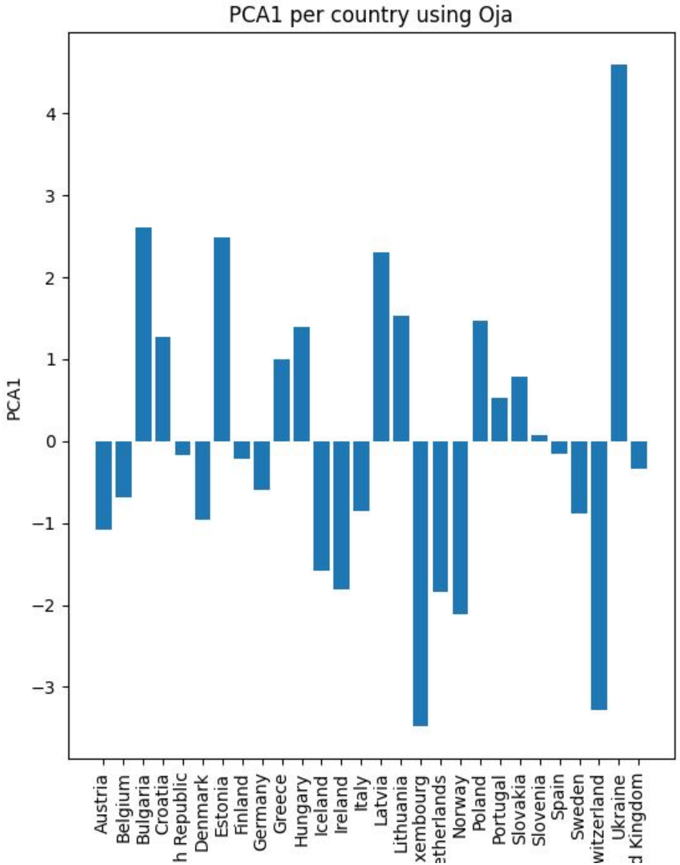
Military
Unemployment
Life.expect
Pop.growth
GDP
Area
Inflation
Greece
Croatia
Spain
Portugal
Slovenia
Ireland
United Kingdom
Italy
Bulgaria
Poland
Hungary
Finland
Czech Republic
Slovakia
Lithuania
Estonia
Switzerland
Netherlands
Norway
Germany
Latvia
Luxembourg
Austria
Iceland
Ukraine
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
PC2
−0.6
−0.4
−0.2
0.0
0.2
0.4
0.6
PC1

Valores con Oja y Sklearn

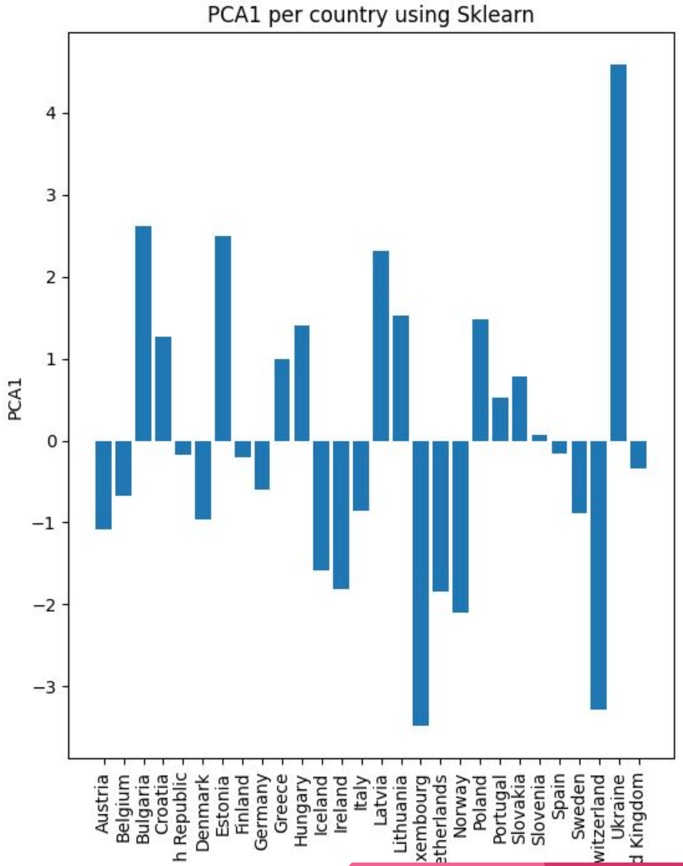
VARIABLE	OJA	SKLEARN
INFLATION	0.407222	0.406518
UNEMPLOYMENT	0.271308	0.271656
MILITARY	0.187514	0.188112
ÁREA	0.125589	0.124874
POPULATION GROWTH	-0.475552	-0.475704
LIFE EXPECTANCY	-0.483021	-0.482873
GDP	-0.500443	-0.500506

```
learning rate: 0.0001 epochs: 10000
```

Oja


PCA1 per country using Oja
PCA1
Austria
Belgium
Bulgaria
Croatia
h Republic
Denmark
Estonia
Finland
Germany
Greece
Hungary
Iceland
Ireland
Italy
Latvia
Lithuania
xembourg
etherlands
Norway
Poland
Portugal
Slovakia
Slovenia
Spain
Sweden
witzerland
Ukraine
d Kingdom

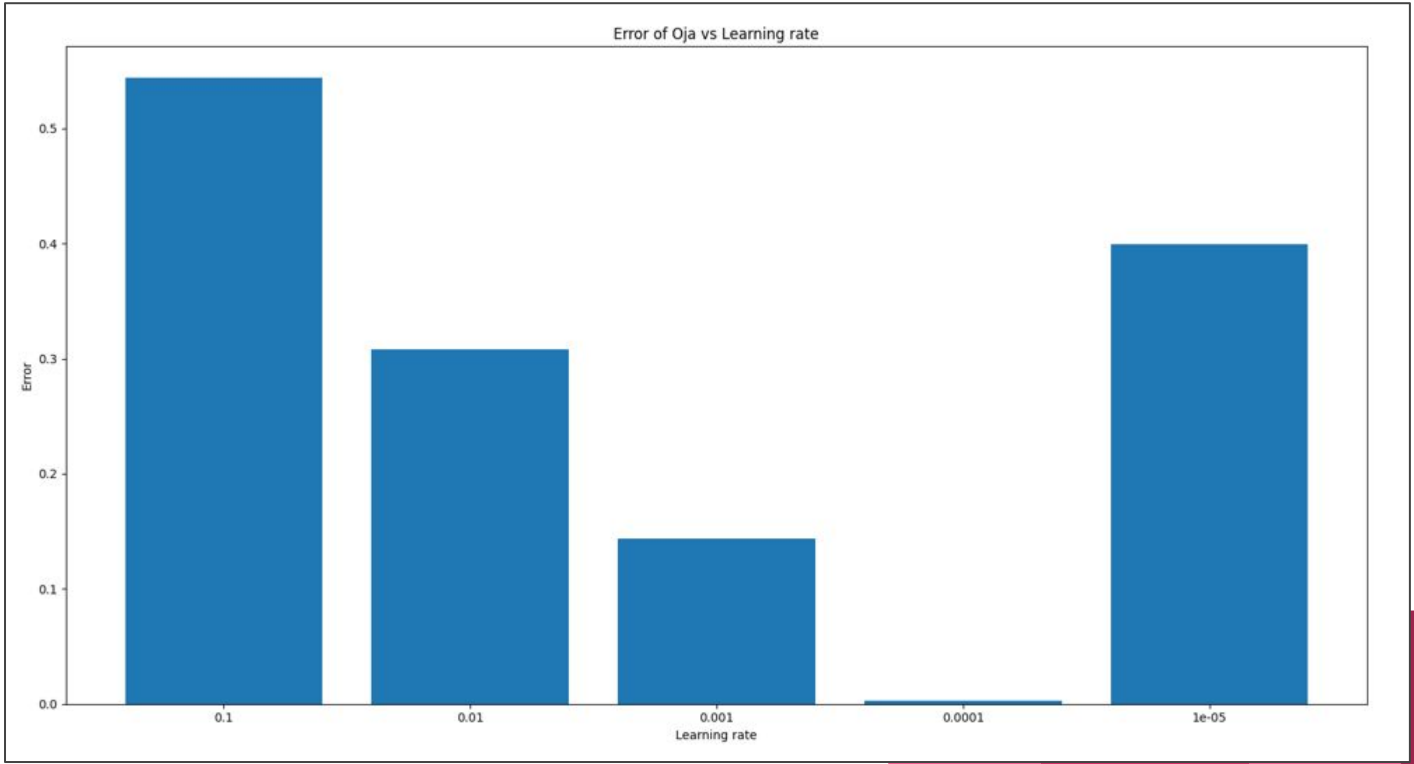
SKLearn


PCA1 per country using Sklearn
PCA1
Austria
Belgium
Bulgaria
Croatia
h Republic
Denmark
Estonia
Finland
Germany
Greece
Hungary
Iceland
Ireland
Italy
Latvia
Lithuania
xembourg
etherlands
Norway
Poland
Portugal
Slovakia
Slovenia
Spain
Sweden
witzerland
Ukraine
d Kingdom

Cambio en la tasa de aprendizaje

1000 épocas
Promedio de 5 ejecuciones

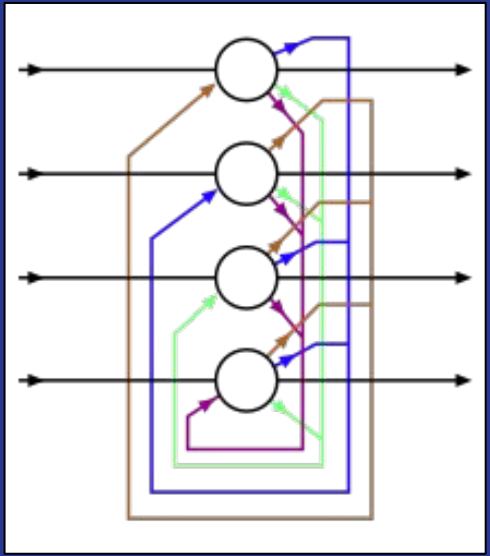
El error comparado con Sklearn



Conclusiones

- Importante estandarizar el dataset

- Tasa de aprendizaje influye mucho en los resultados



2. Hopfield

Consigna

Construir letras del abecedario usando **{-1, 1}** con matrices de **5x5**, e implementar el método de Hopfield para asociar matrices de consulta, generalmente con ruido, con patrones de letras almacenados.

$$\begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ -1 & -1 & -1 & 1 & -1 \\ -1 & -1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 & -1 \\ 1 & 1 & 1 & -1 & -1 \end{pmatrix}$$

Matriz

```
*  *  *  *  *
         *
         *
*        *
*  *  *
```

Patrón

Estructura de estado

Se guardan **4** patrones de letras en la red de *Hopfield* de acuerdo a la **regla del 15%**, excepto cuando se pruebe qué pasa variando este valor.

Estas letras se almacenan como vectores de **25** posiciones con los mismos valores **{-1, 1}**, y se usan para calcular la ortogonalidad entre patrones.

Métrica de ortogonalidad

Ya que ningún conjunto de letras resultaba puramente ortogonal, usamos un **score** para medir la ortogonalidad de dicho conjunto, donde:

$$\boxed{score(\{a, b, c, d\}) = |a * b| + |a * c| + |a * d| + |b * c| + |b * d| + |c * d|}$$

En este caso, mientras **menor sea el score, más ortogonal será el conjunto** y, por ende, mejores resultados con Hopfield.

También lo usaremos para comparar vectores *muy ortogonales* con otros *poco ortogonales*, y sus diferencias en resultados.

Conjuntos de prueba

Para los análisis a realizar se usaron 3 conjuntos de letras, cada uno representativo de un nivel de ortogonalidad según el score usado:

- **[G, R, T, V]** -> score: 6 (el más ortogonal)
- **[O, Q, T, X]** -> score: 58 ("algo" ortogonal)
- **[A, F, P, R]** -> score: 110 (el menos ortogonal)

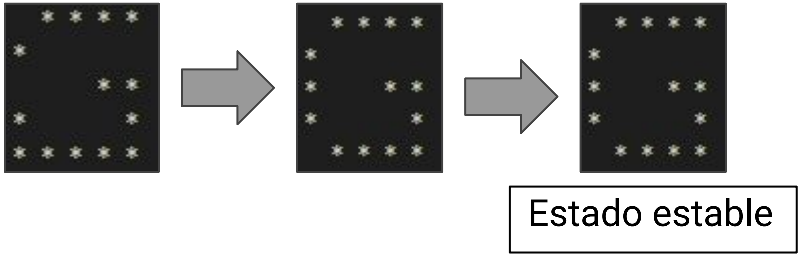
Además, sus consultas sin ruido (de letras del conjunto) son:

- Ortogonal: todas correctas
- "Algo" ortogonal: **O,Q -> Q | T -> T | X -> espúreo**
- Menos ortogonal: todas devuelven **P**

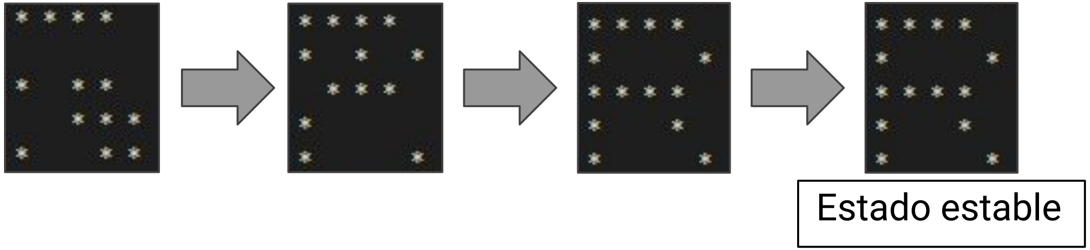
Ejemplos de consultas

Conjunto ortogonal **[G, R, T, V]**

Input: G
Ruido: 0,1



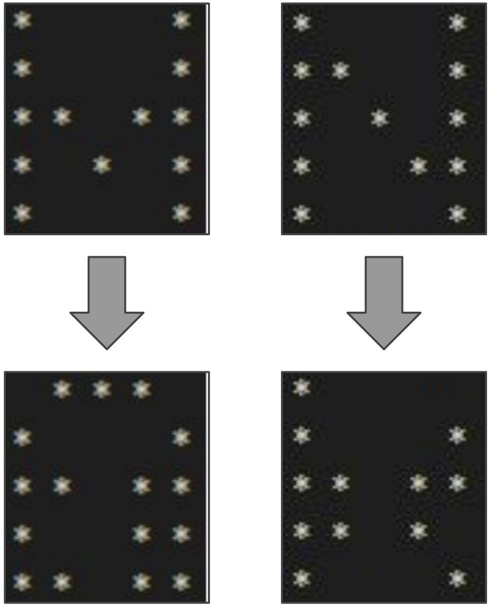
Input: R
Ruido: 0,3



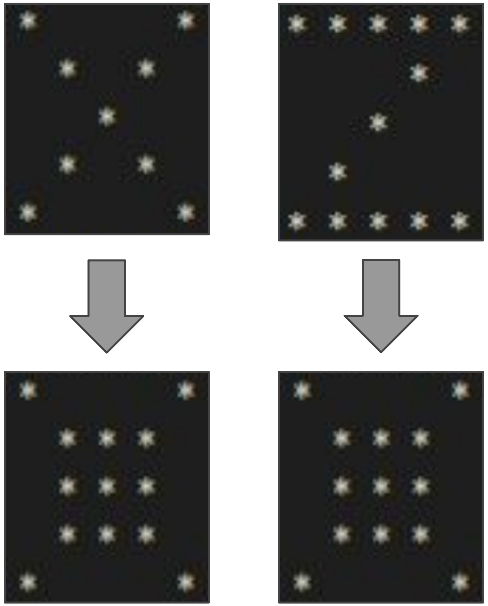
Input: V
Ruido: 0,5

Estados espúreos

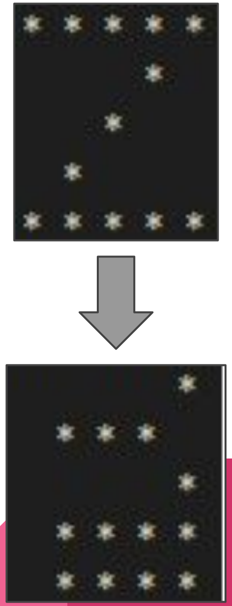
Más ortogonal
[G, R, T, V]



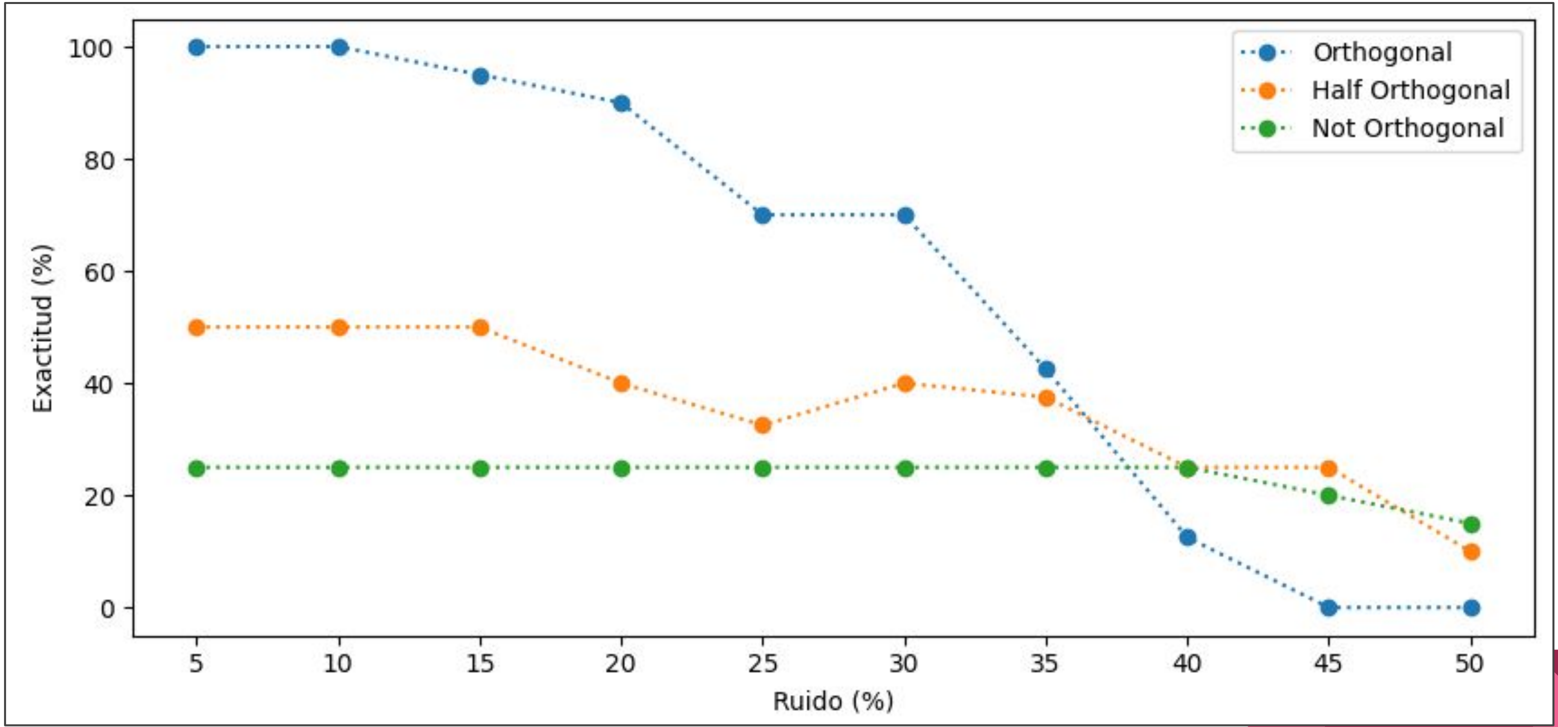
“Algo” ortogonal
[O, Q, T, X]



Nada ortogonal
[A, F, P, R]

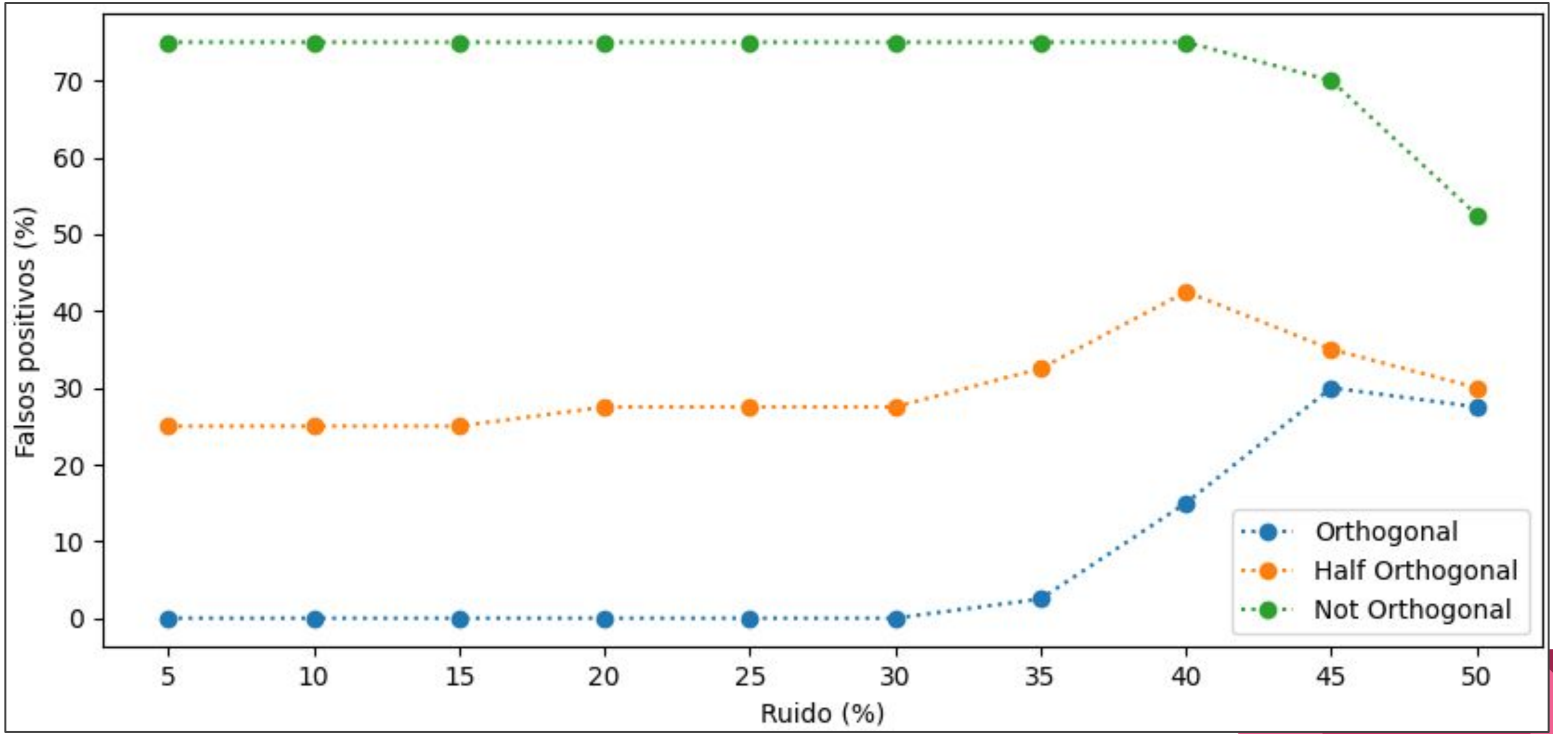


Ruido - Exactitud



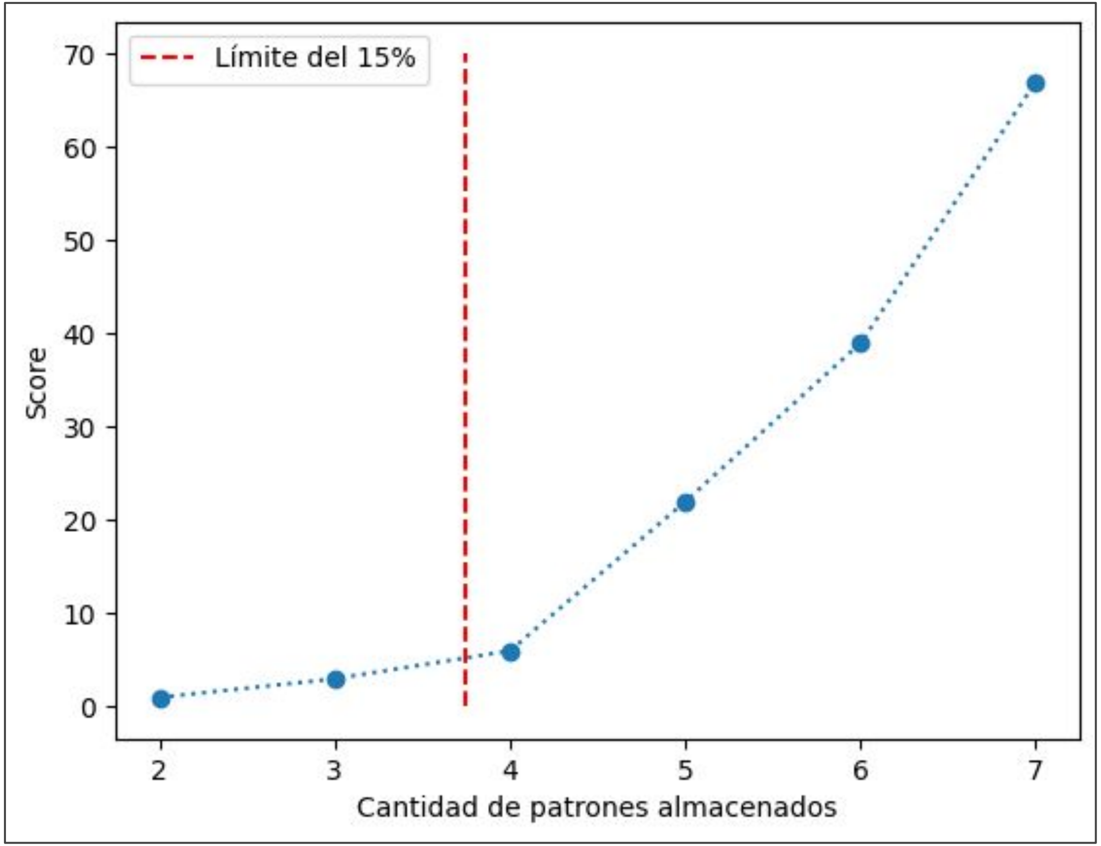
40 realizaciones por valor y conjunto, con 10 por letra

Ruido - “Falsos positivos”



40 realizaciones por valor y conjunto, con 10 por letra

Cantidad de patrones

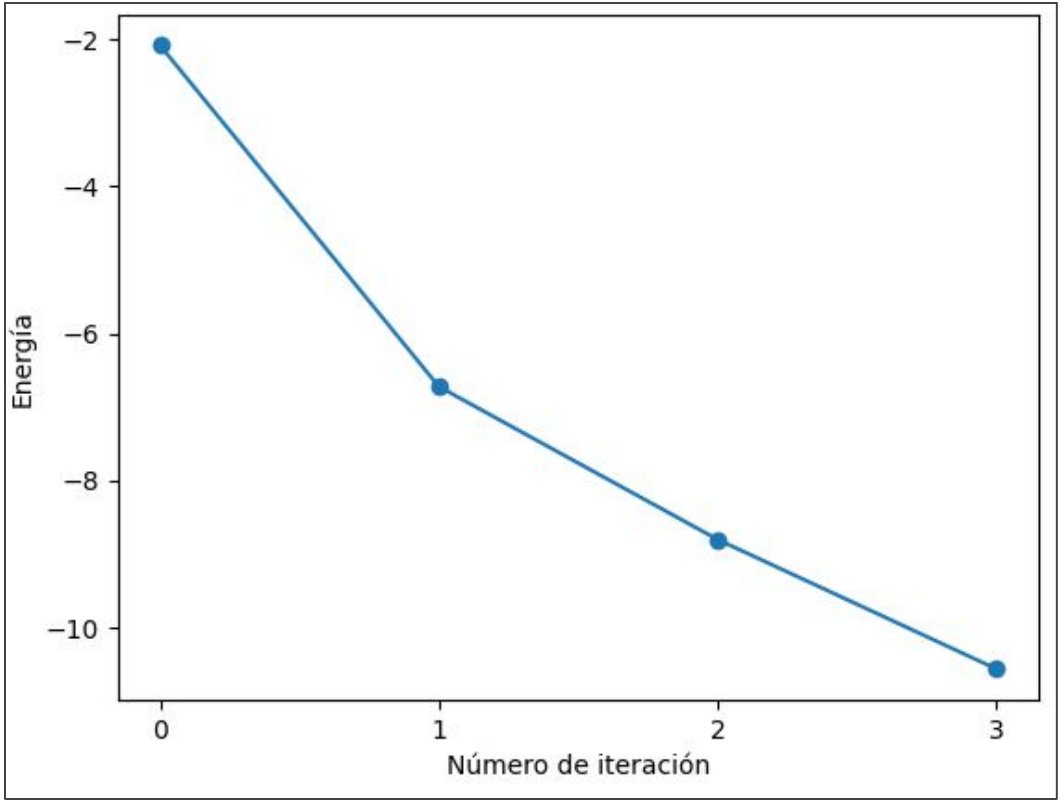


Se usa el conjunto más ortogonal para cada tamaño de conjunto.

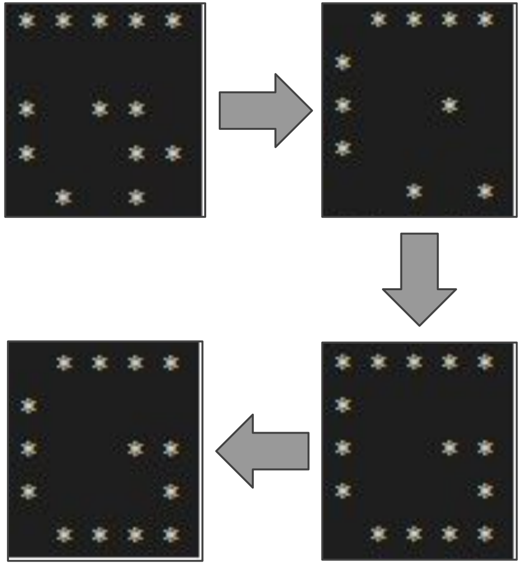
El conjunto es relativamente ortogonal para tamaños menores o cercanos a la regla del 15%.

Luego, el score de ortogonalidad se dispara, cayendo en los casos antes vistos de baja ortogonalidad.

Función de energía



Más ortogonal
[G, R, T, V]



Letra G con 30% de ruido

Conclusiones

- La presencia de estados espúreos es independiente de la ortogonalidad del conjunto.
- El modelo es susceptible al ruido, siendo mayor este efecto mientras más ortogonal sea el conjunto usado.
- En el caso estudiado, sobrepasar el límite teórico de 15% significa no poder encontrar conjuntos *relativamente* ortogonales.

¡Gracias!