



SIA - TP4

Modelos de aprendizaje no supervisado

Grupo 1

Alberto Bendayan

Tobias Ves Losada

Cristian Tepedino



Introducción



Csv de países de Europa

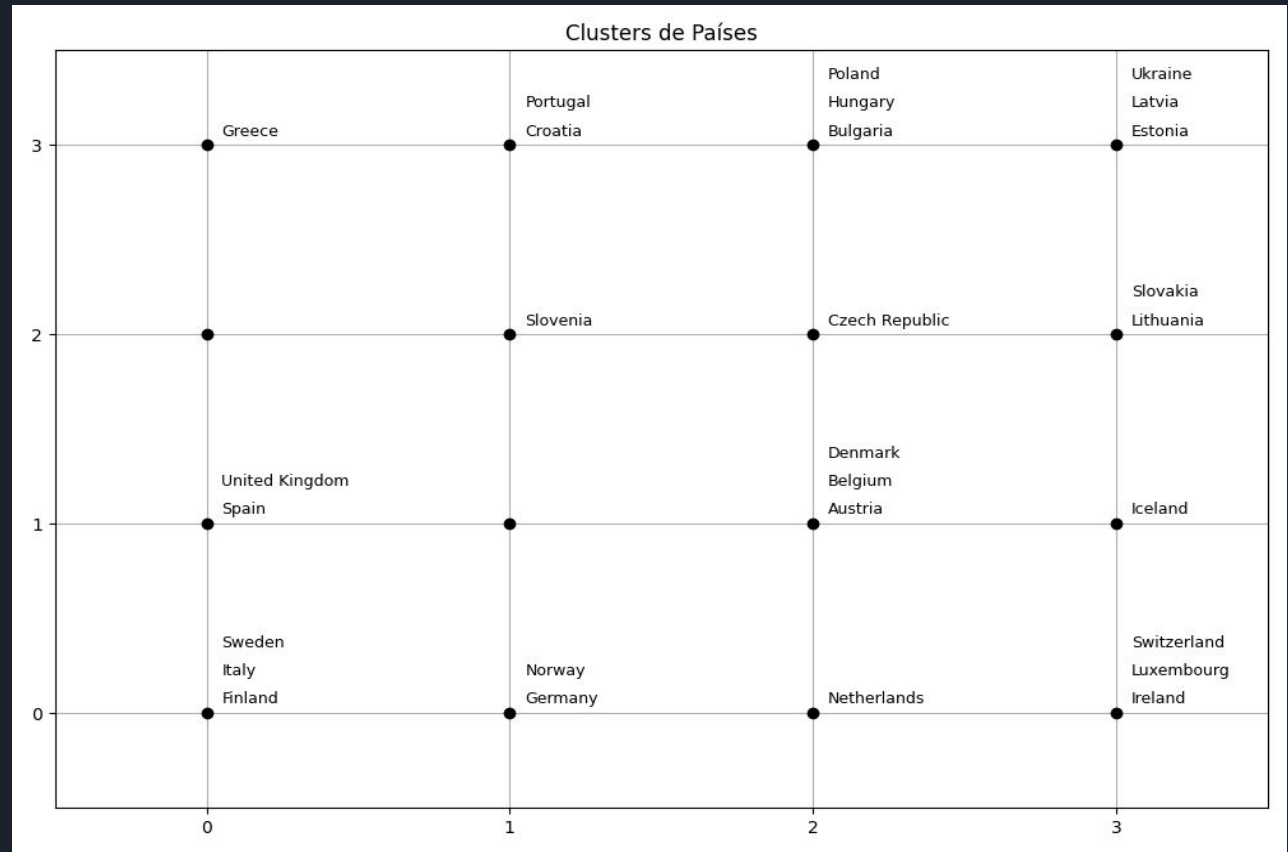
28 países:

- Nombre del país
- Área
- Producto bruto interno
- Inflación anual
- Expectativa de vida media en años
- Presupuesto militar
- Tasa de crecimiento poblacional
- Tasa de desempleo

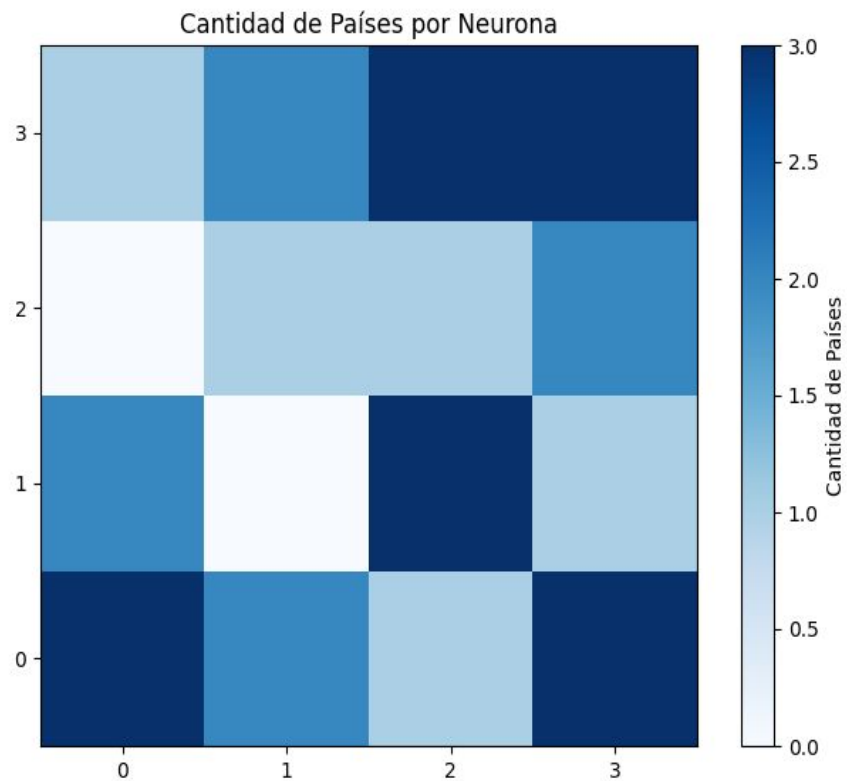
A decorative graphic on the left side of the slide. It consists of a blue parallelogram and a light green parallelogram, both tilted at an angle. The blue shape is in the foreground, and the green shape is partially behind it. They are set against a dark blue background with diagonal stripes of varying shades.

Red de Kohonen

Learning Rate inicial: 0,1
Iteraciones: 10000

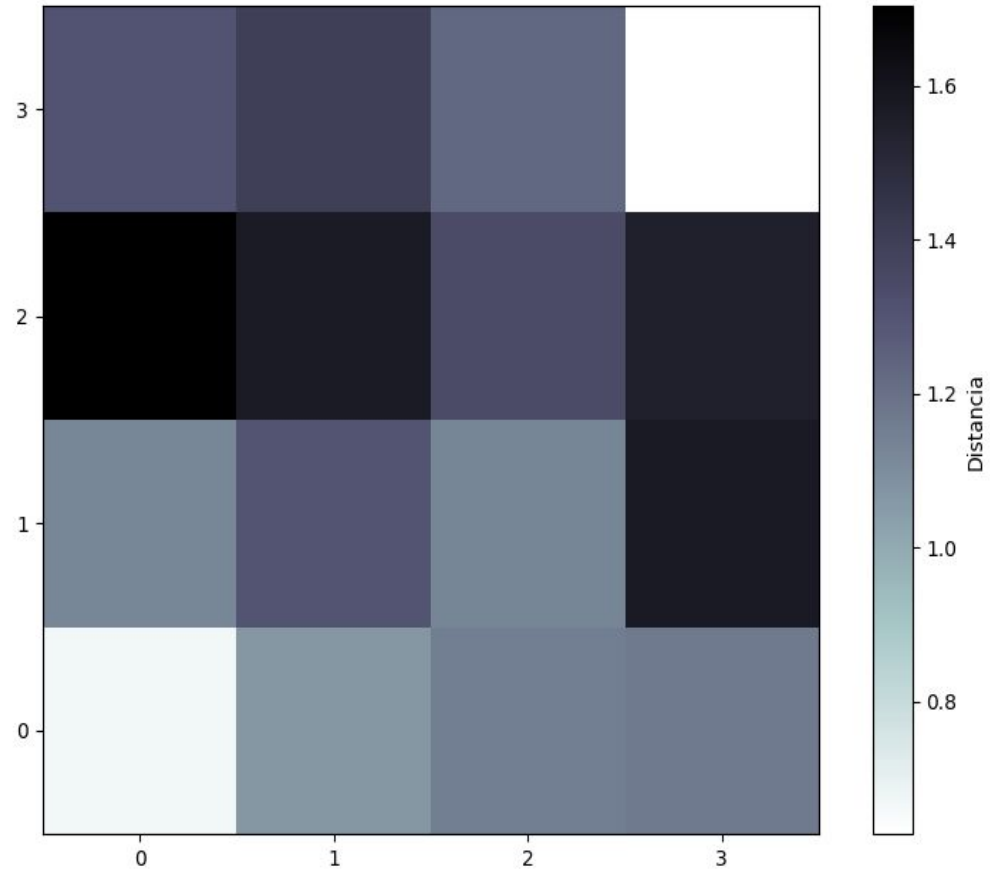


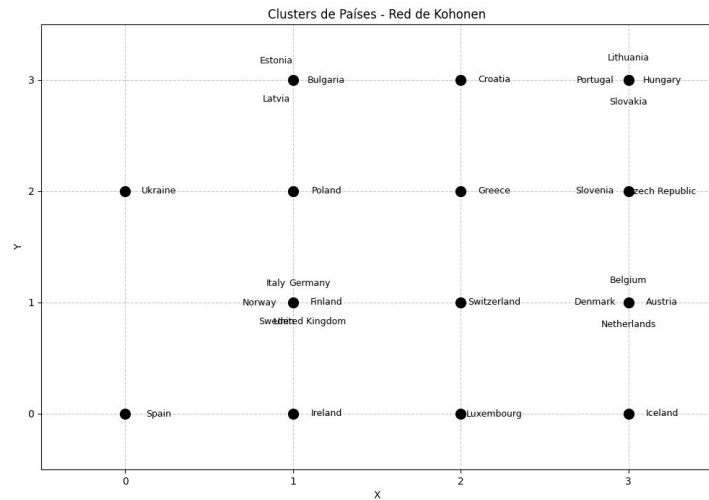
Learning Rate inicial: 0,1
Iteraciones: 10000



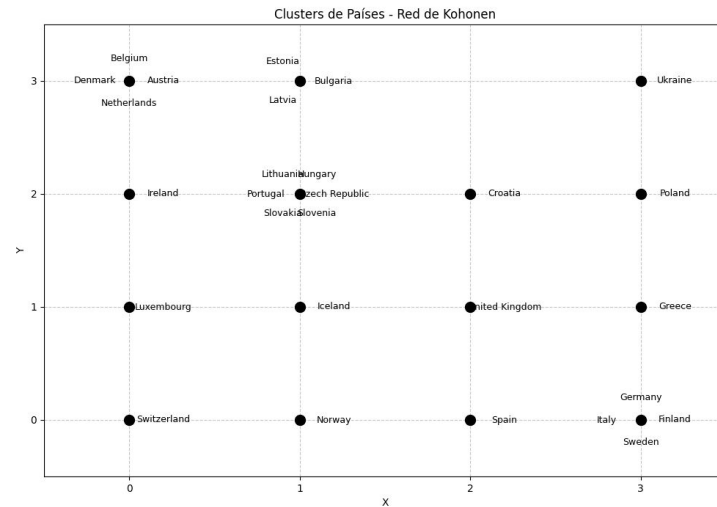


Learning Rate inicial: 0,1
Iteraciones: 10000

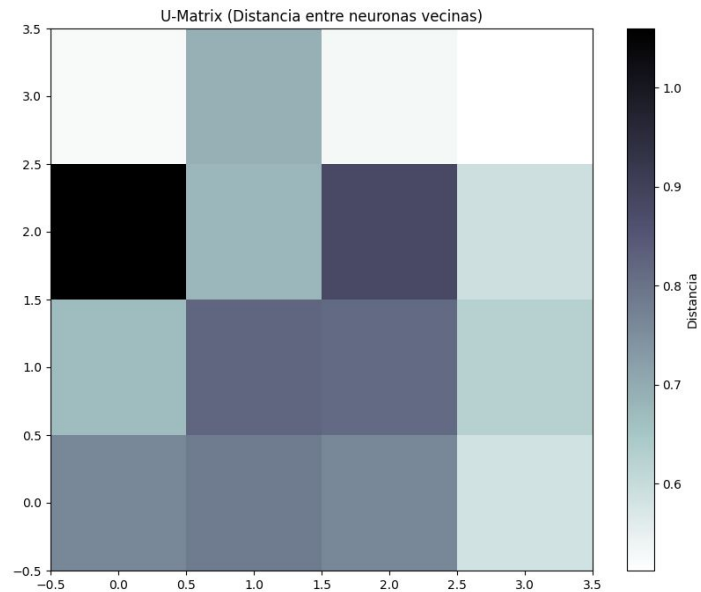




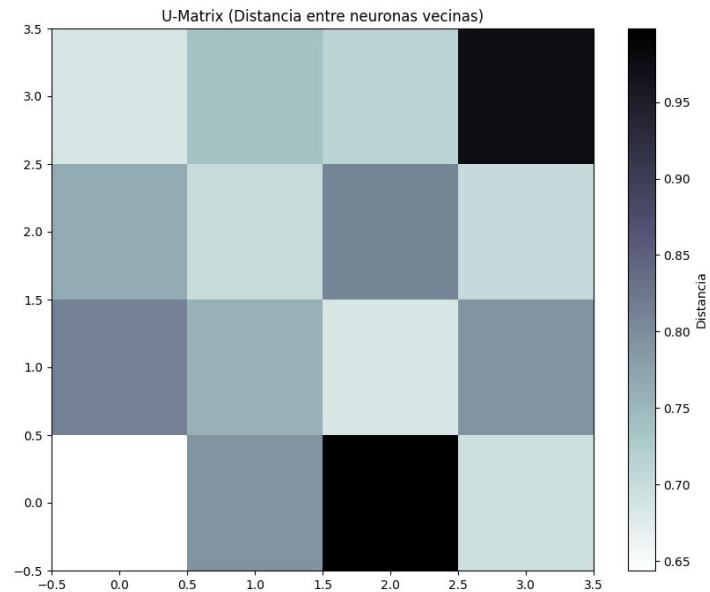
Learning Rate inicial: 0,4



Learning Rate inicial: 0,8

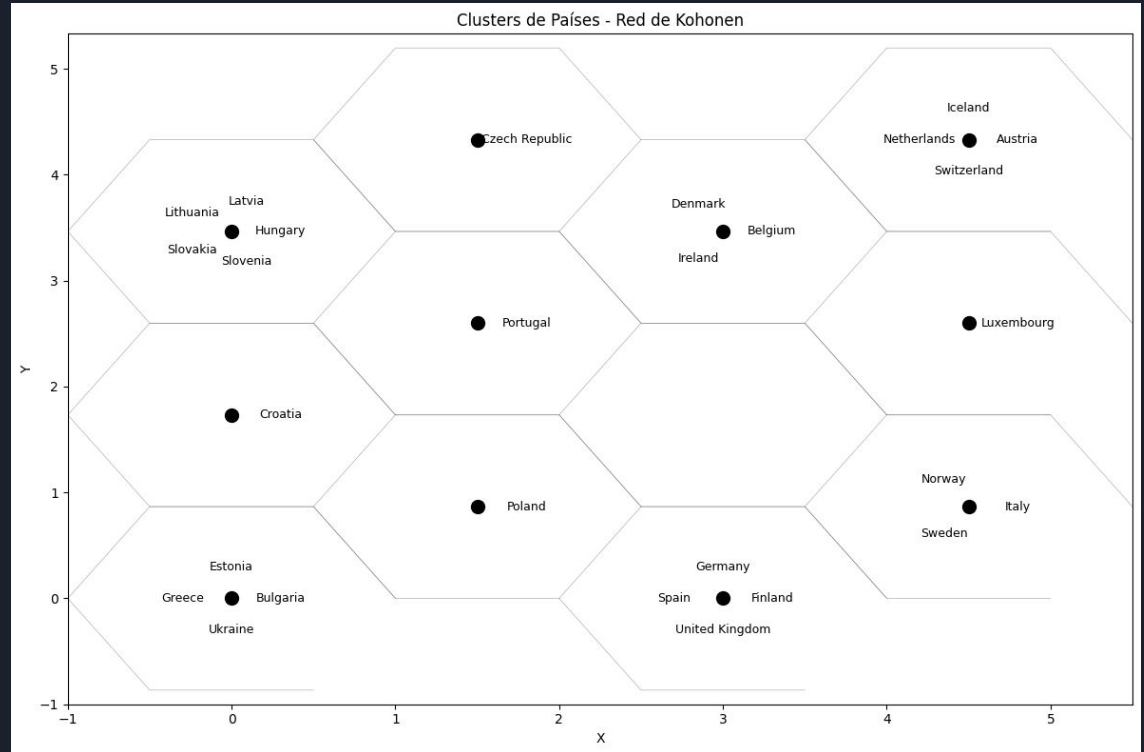



Learning Rate inicial: 0,4



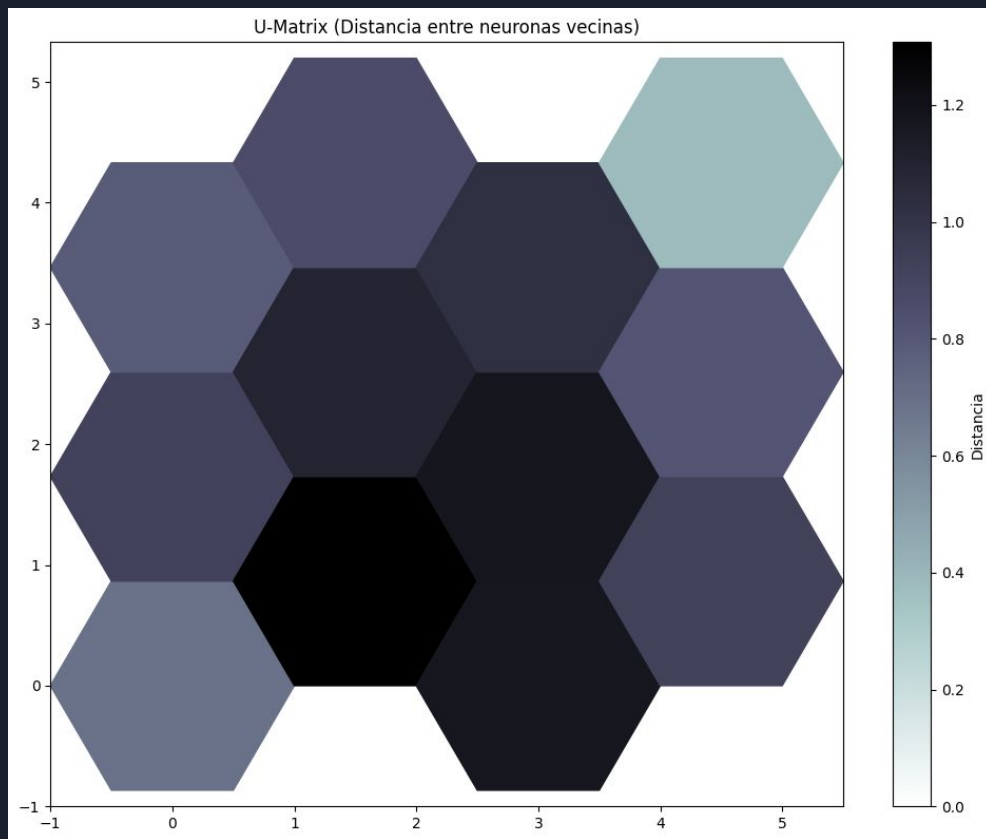
Learning Rate inicial: 0,8

Learning Rate inicial: 0,05
Iteraciones: 10000





Learning Rate: 0,05
Iteraciones: 10.000





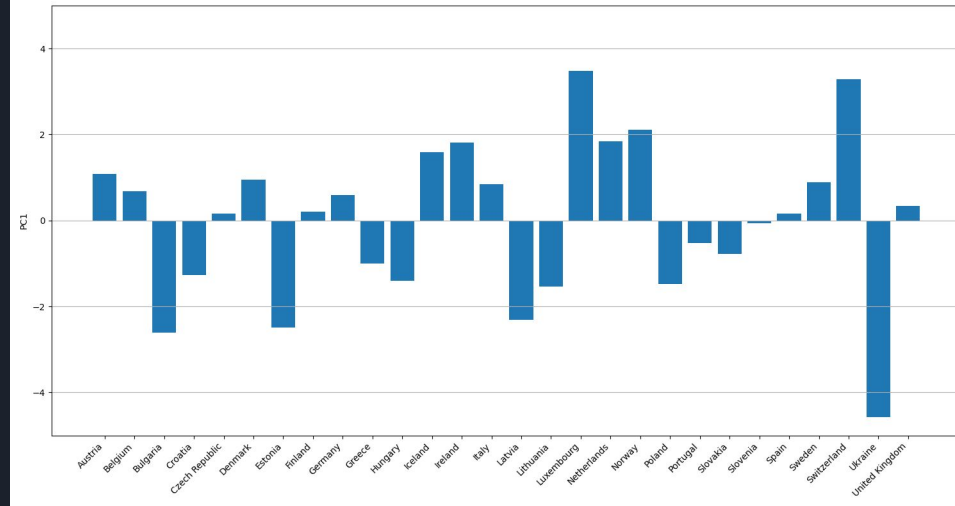
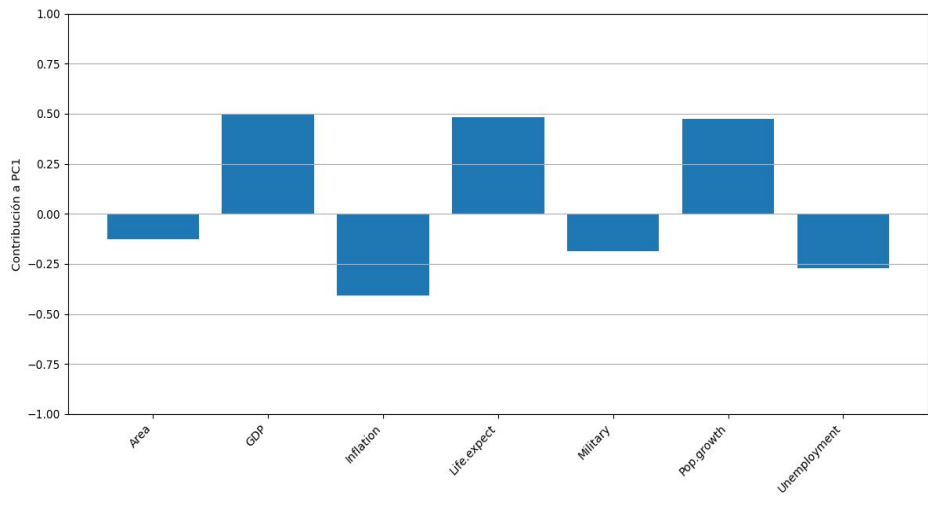
Conclusiones

- Grupos bien definidos (Ucrania y España siempre solos)
- Regiones cercanas y opuestas
- Kohonen como herramienta agrupadora poderosa
- Buenos resultados variando el Learning Rate

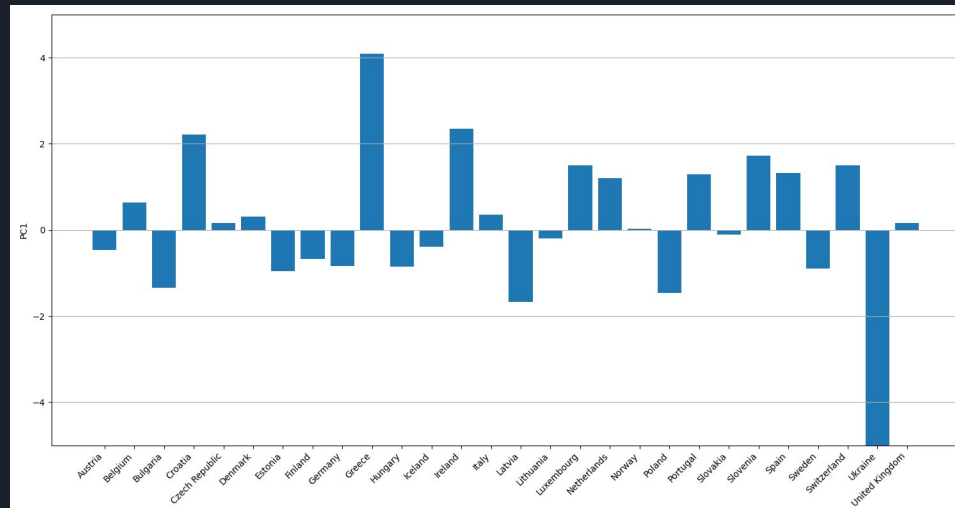
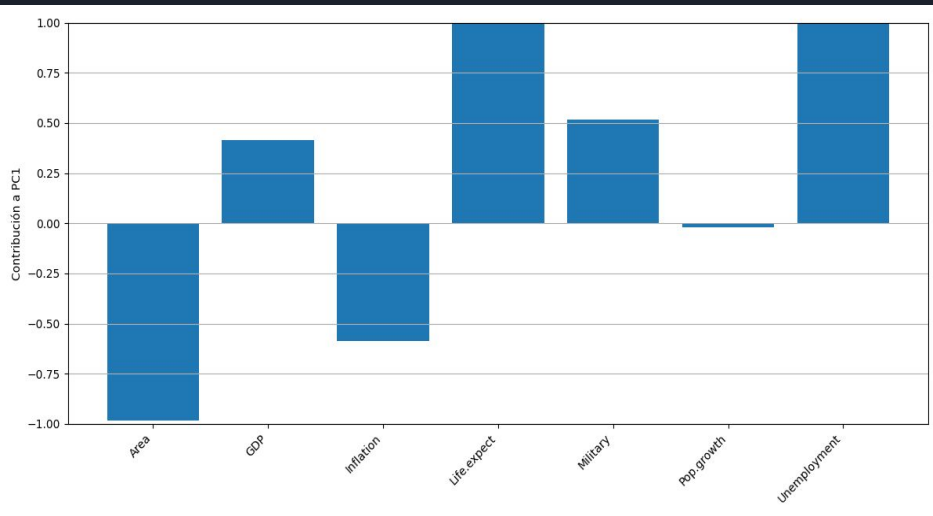


Modelo de Oja

PC1 por librería



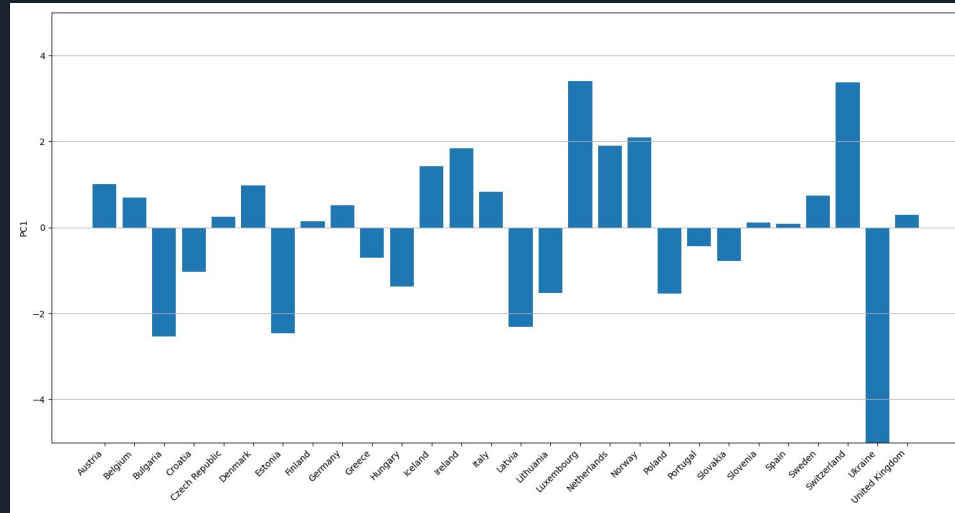
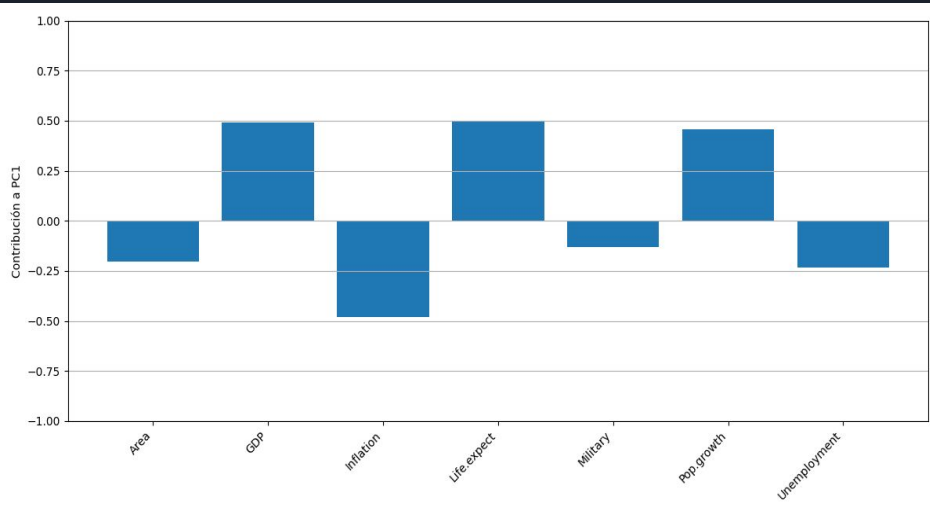
PC1 por regla de Oja



Learning rate constante
 $\eta(0) = 0.1$

E = 0.646

PC1 por regla de Oja

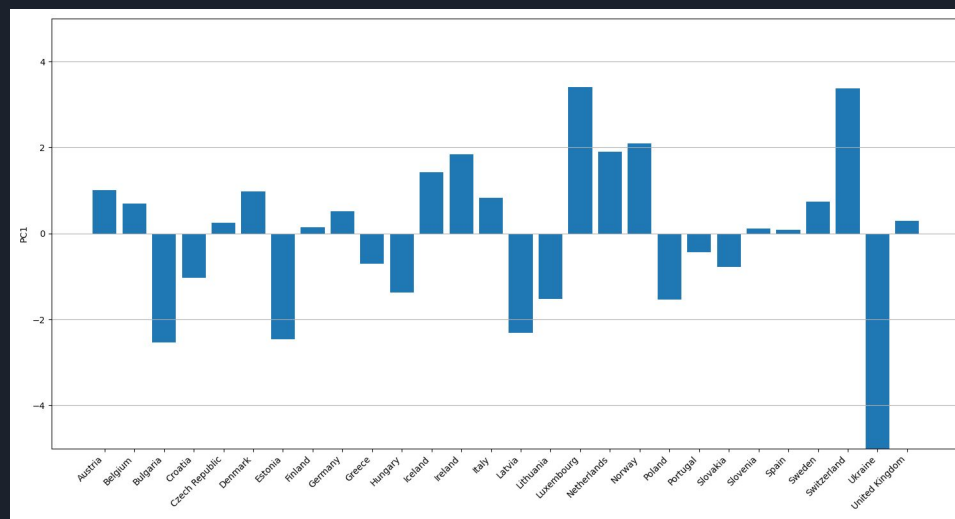
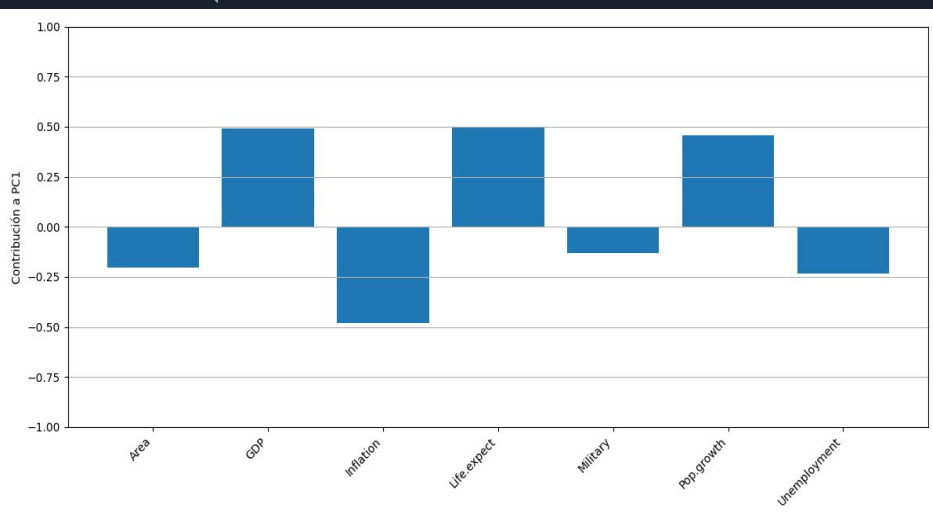


Learning rate constante
 $\eta(0) = 0.01$

$\text{Cos}(\theta) = 0.99185$

$E = 8.14 \cdot 10^{-3}$

PC1 por regla de Oja

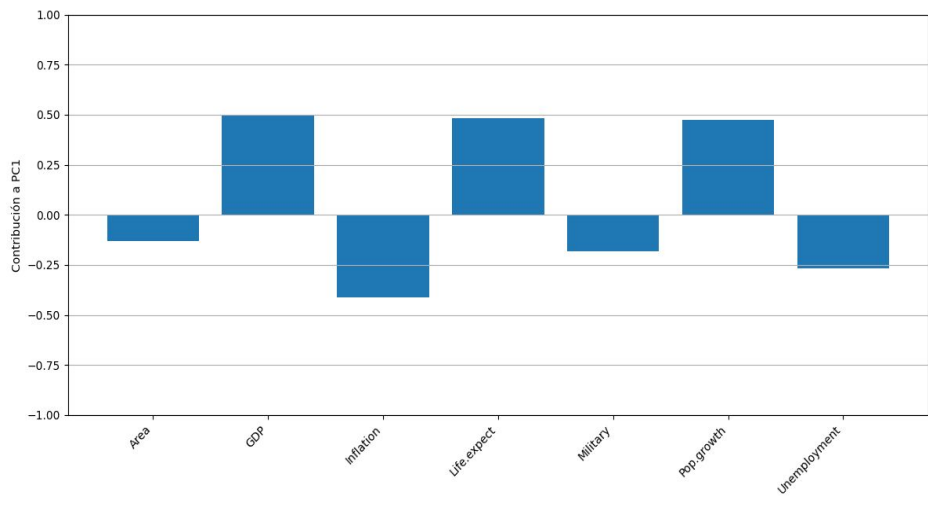


Learning rate constante
 $\eta(0) = 0.01$

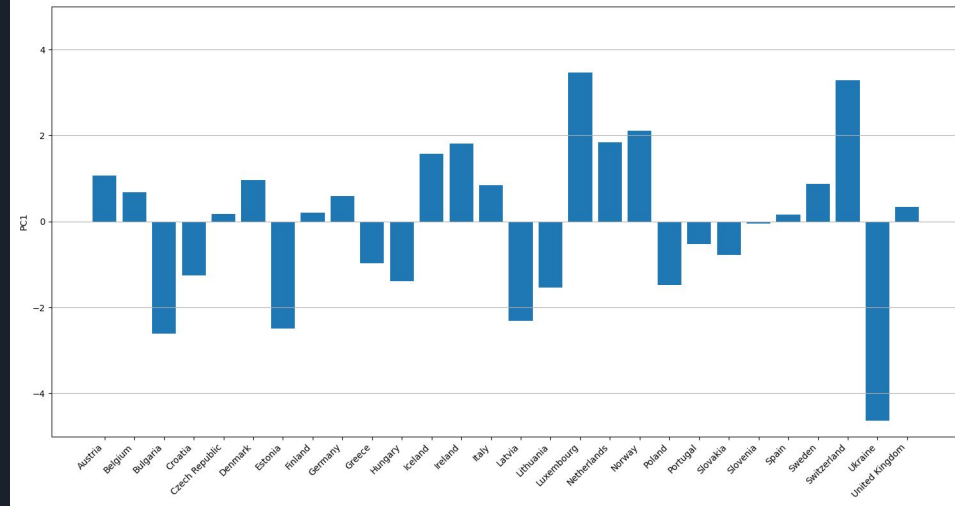
$\text{Cos}(\theta) = -0.99185$

$E = 8.14 \cdot 10^{-3}$

PC1 por regla de Oja

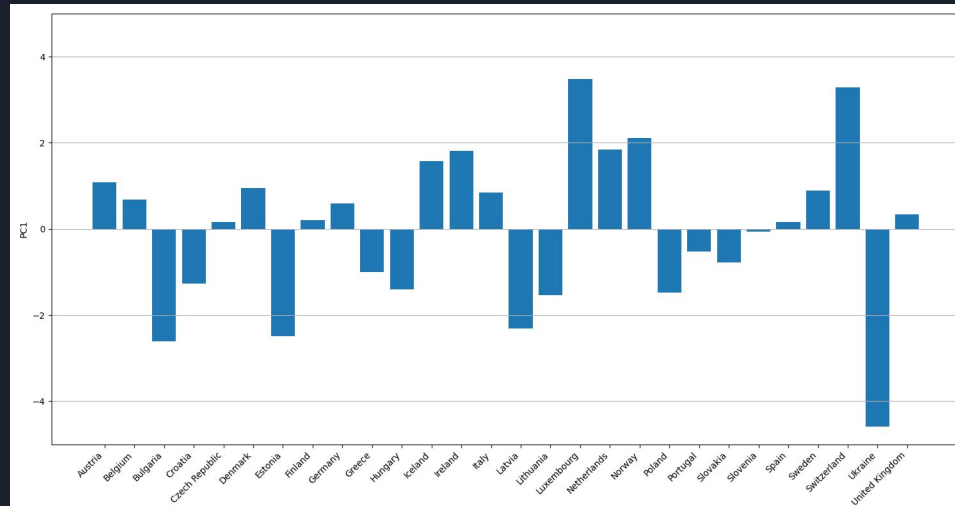
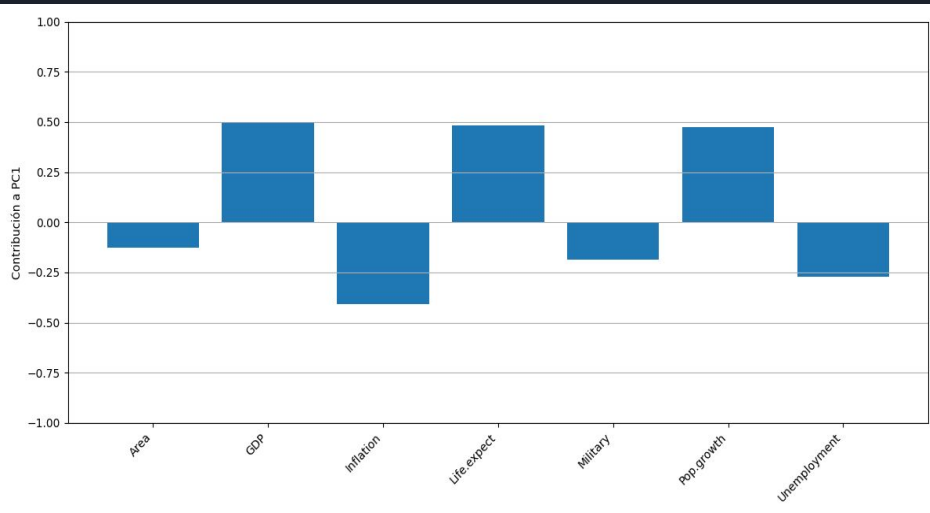


Learning rate constante
 $\eta(0) = 0.001$



$$E = 7.63 \cdot 10^{-5}$$

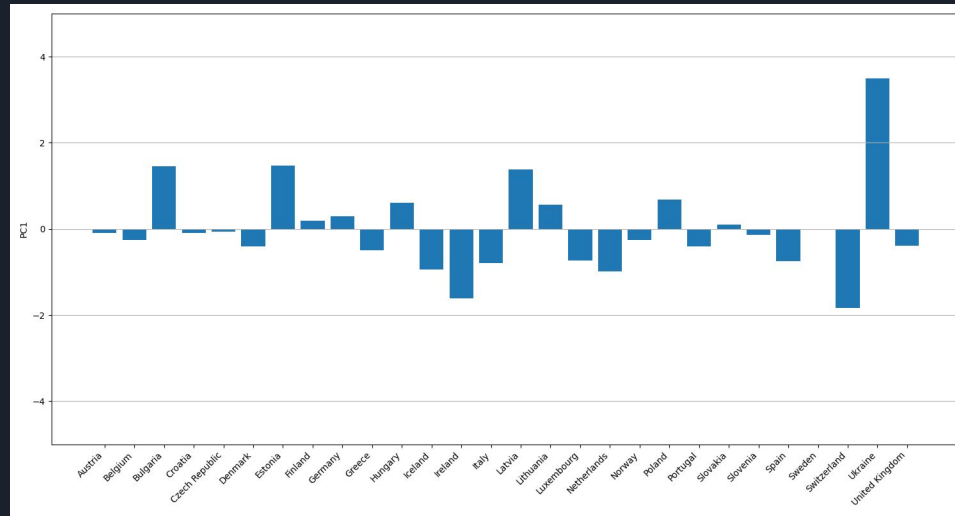
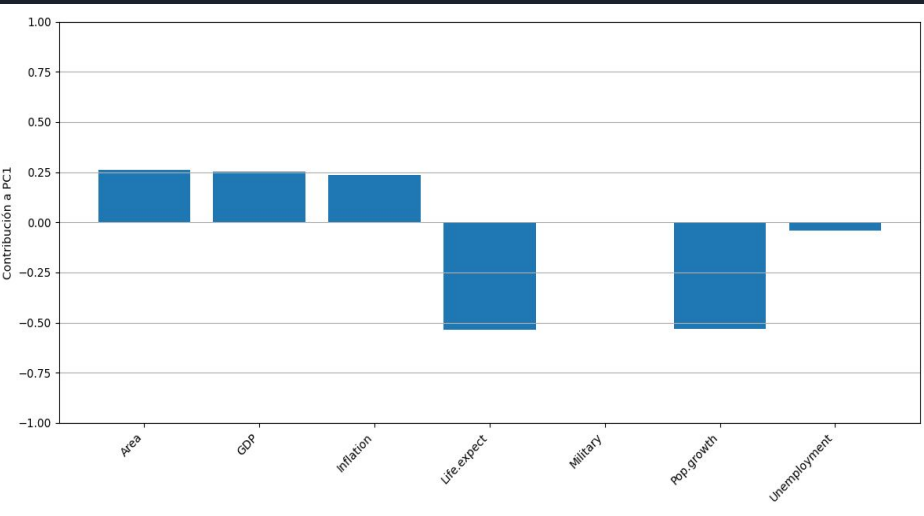
PC1 por regla de Oja



Learning rate constante
 $\eta(0) = 0.0001$

$$E = 7.57 \cdot 10^{-7}$$

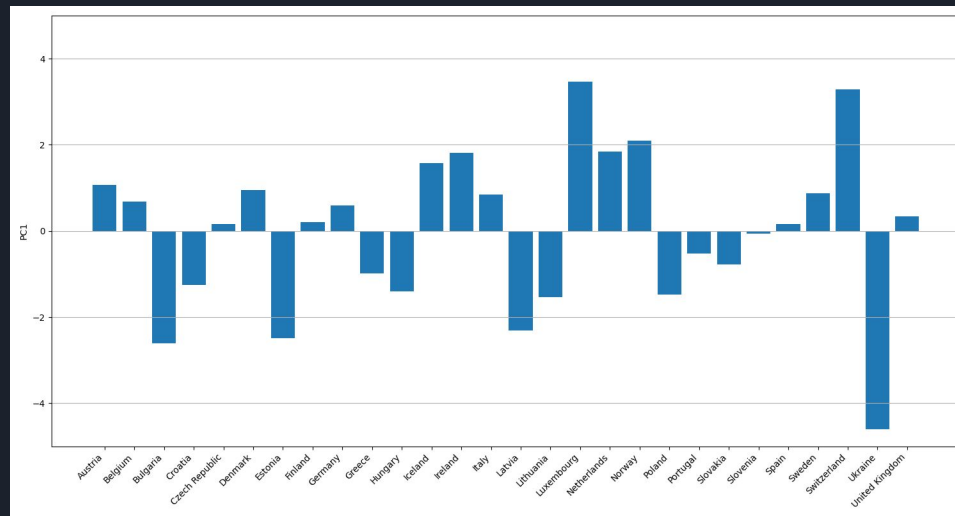
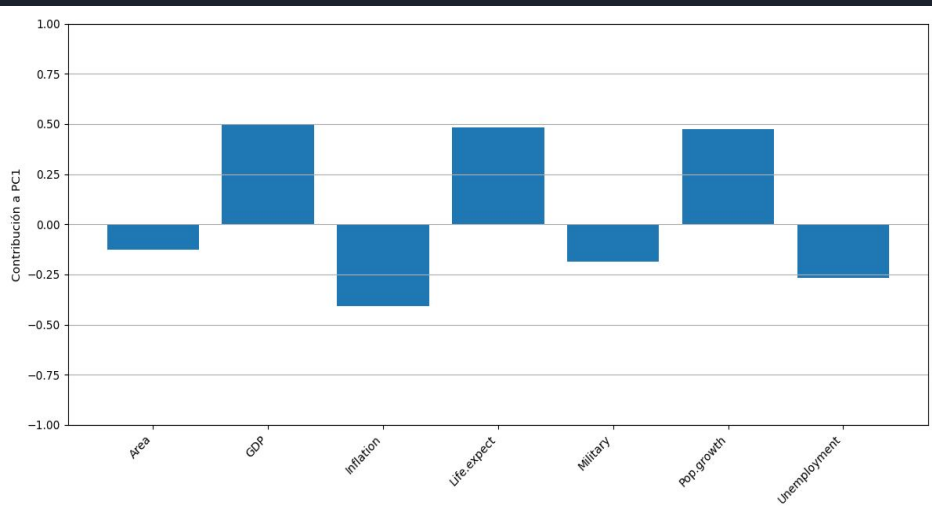
PC1 por regla de Oja



Learning rate adaptativo
 $\eta(0) = 0.001$

E = 0.423

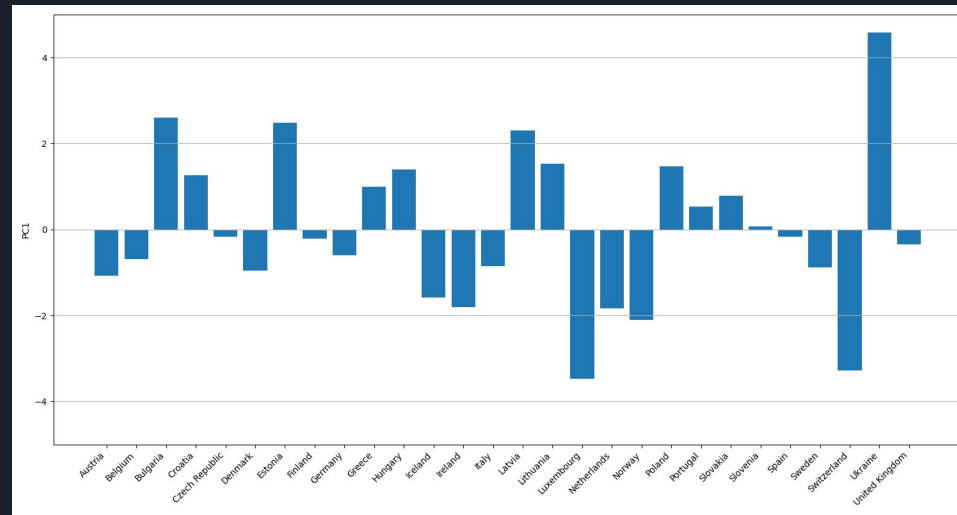
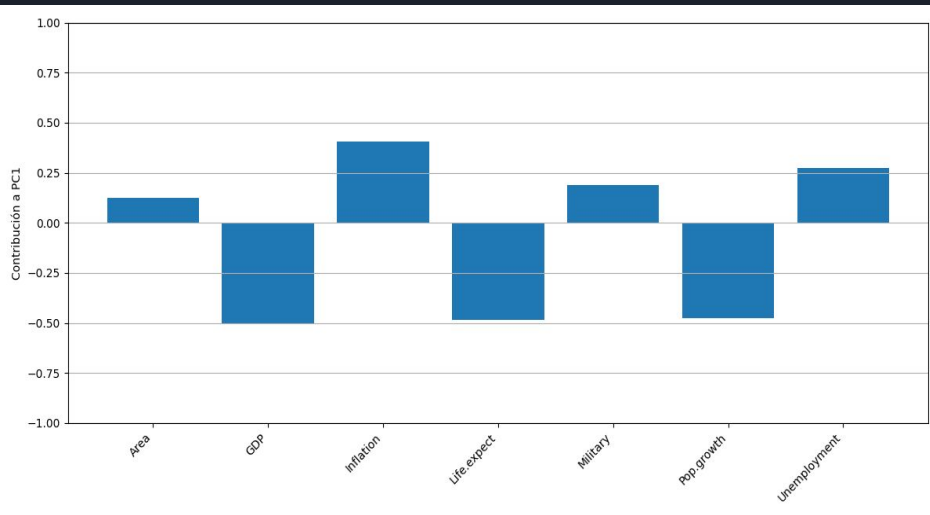
PC1 por regla de Oja



Learning rate adaptativo
 $\eta(0) = 0.01$

$$E = 1.87 \cdot 10^{-5}$$

PC1 por regla de Oja



Learning rate adaptativo
 $\eta(0) = 0.1$

$$E = 1.09 \cdot 10^{-8}$$



Resumen de los resultados

Learning Rate	$\eta(0)$	E
Constante	0.1	0.646
Constante	0.01	$8.14 \cdot 10^{-3}$
Constante	0.001	$7.63 \cdot 10^{-5}$
Constante	0.0001	$7.57 \cdot 10^{-7}$
Adaptativo	0.1	$1.09 \cdot 10^{-8}$
Adaptativo	0.01	$1.87 \cdot 10^{-5}$
Adaptativo	0.001	0.423

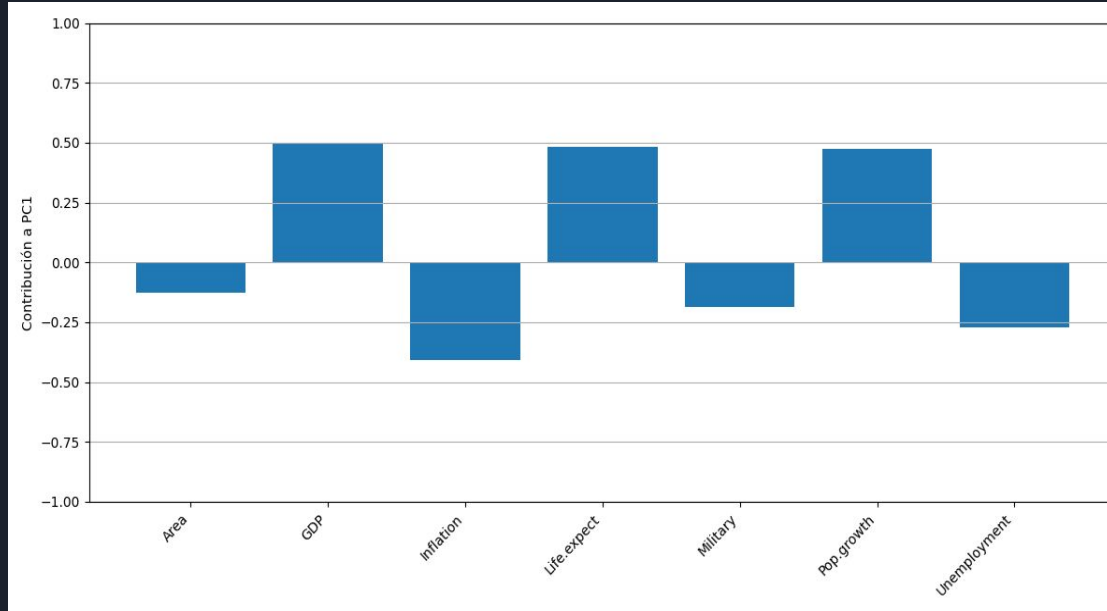


Otras posibilidades

- Un learning rate inicial más alto causa que los pesos divergan, eventualmente convirtiéndose en NaNs
- Un learning rate constante más bajo podría llevar a menor error, pero requiere más epochs para converger
- Un learning rate adaptativo más bajo causará que los pesos no convergan a la PC1, como se vio en el caso de 0.001
- Se podría probar con otras funciones para actualizar el learning rate, que quizás permitan converger a la PC1 con menor error

Interpretación de los resultados

- El GDP, la expectativa de vida y el crecimiento poblacional tienen una correlación (positiva) fuerte en el índice
- La inflación y el desempleo tienen una correlación (negativa) en el índice
- Si bien el área y la inversión militar tienen una correlación negativa, su peso en la PC1 es bajo





Conclusiones

- Se debe elegir un learning rate adecuado para que el modelo de Oja converga a la PC1
- La regla de Oja permite calcular la PC1 sin la necesidad de almacenar y operar sobre toda la matriz de datos.
 - Calcular la matriz de covarianza y sus autovectores y autovalores es muy costoso para datasets grandes
 - Oja incluso permitiría actualizar los pesos a medida que llegan nuevos datos, en caso de no tenerlos todos desde un principio

A decorative graphic on the left side of the slide. It consists of a blue parallelogram and a light green parallelogram, both tilted at an angle. The blue shape is in the foreground, and the green shape is partially behind it. They are set against a dark blue background with faint, lighter blue diagonal stripes.

Modelo de Hopfield

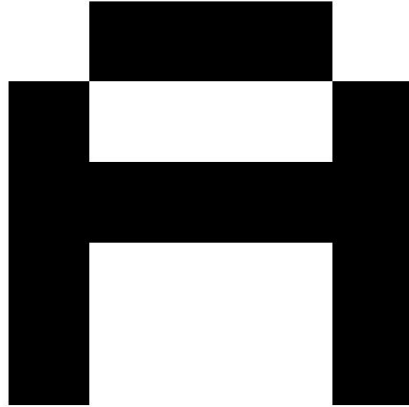


Set de 4 letras aleatorio

Letra J



Letra A

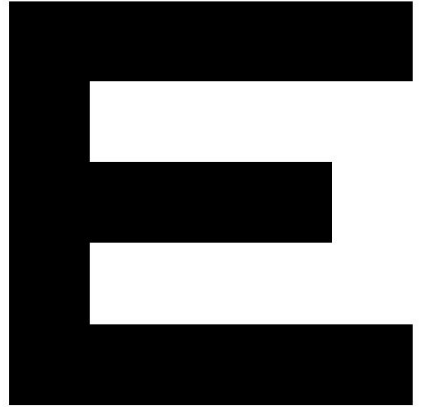


Patrones de las Letras

Letra L



Letra E



Convergencia letra A

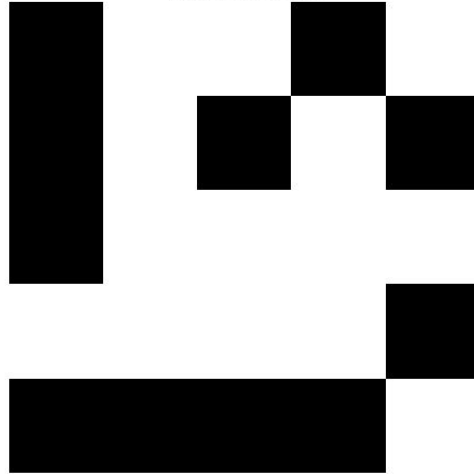


Estado espurio

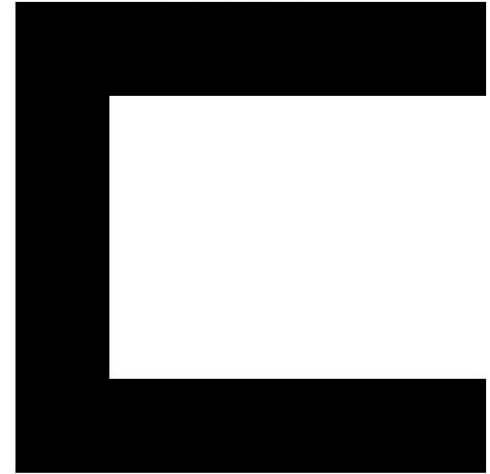
Patrón Original



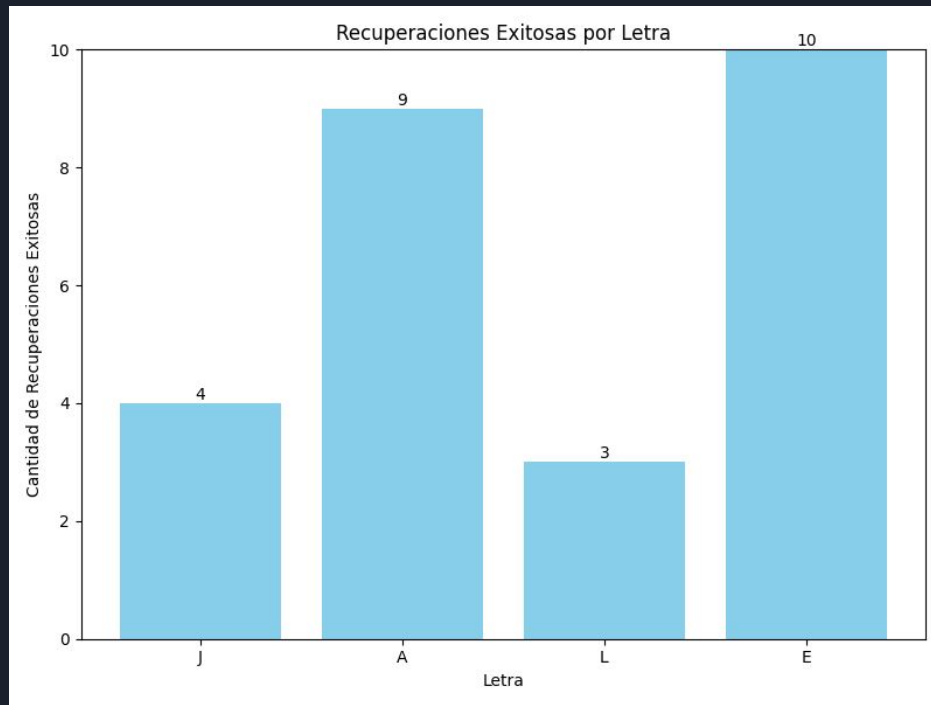
Patrones de la letra L (Ruido: 25.0%)
Patrón con Ruido



Patrón Recuperado

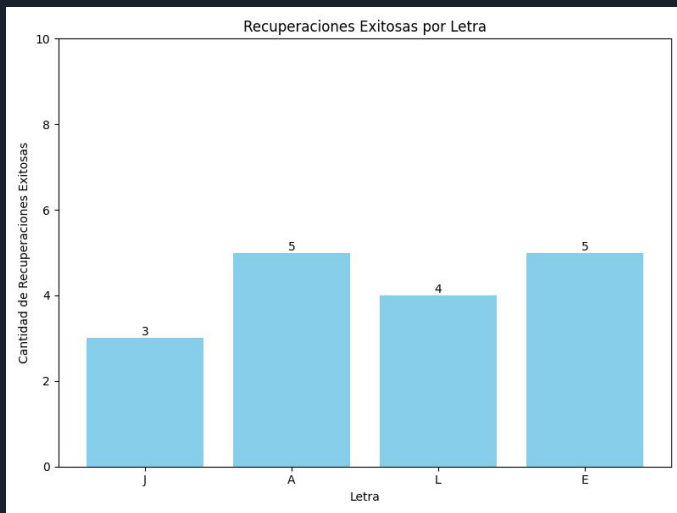


Recuperaciones exitosas

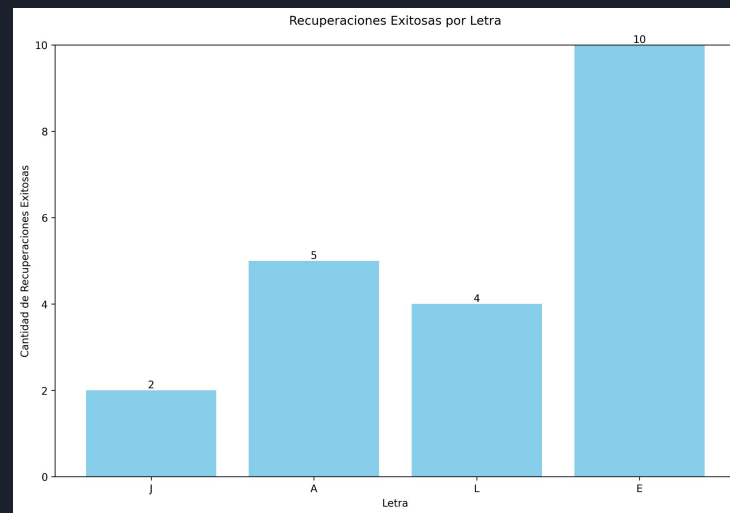


10% de ruido

Recuperaciones exitosas

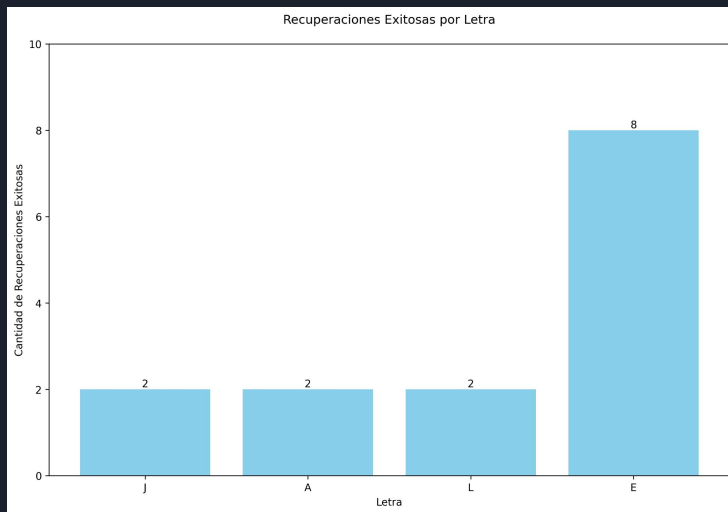


25% de ruido

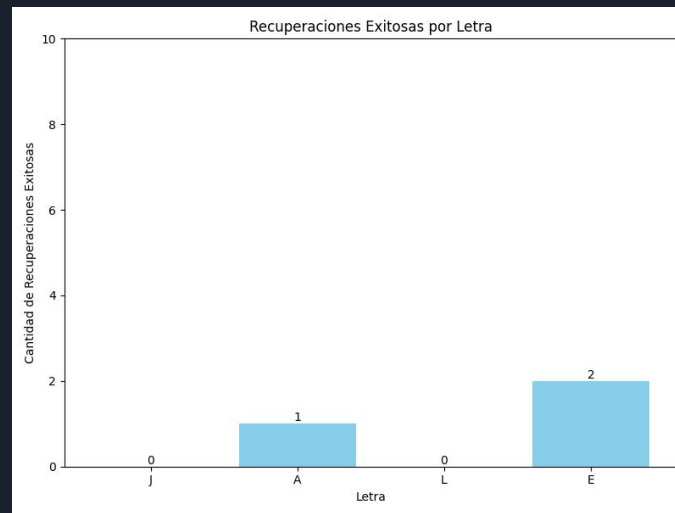


35% de ruido

Recuperaciones exitosas

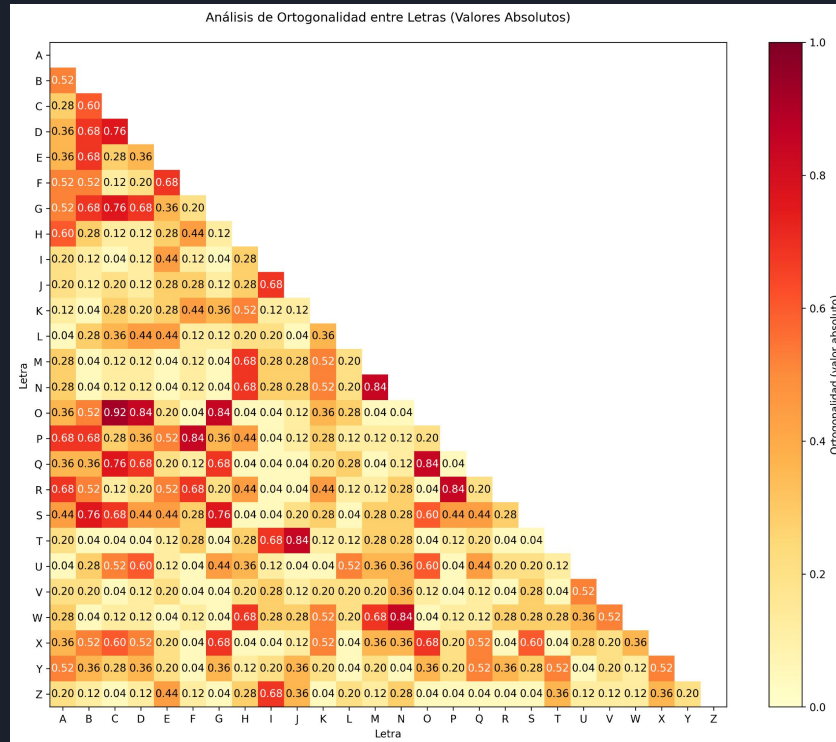


40% de ruido

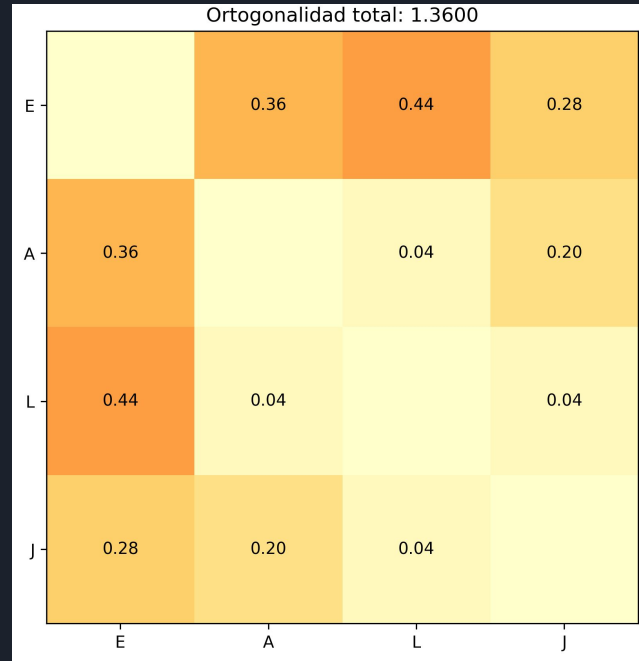


50% de ruido

Análisis de ortogonalidad

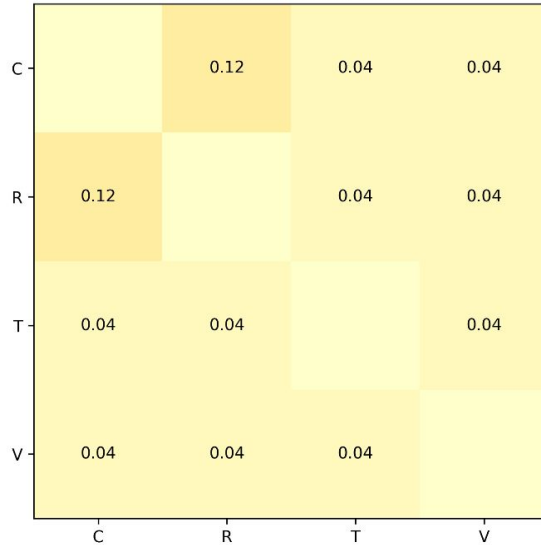


Ortogonalidad E, A, L, J

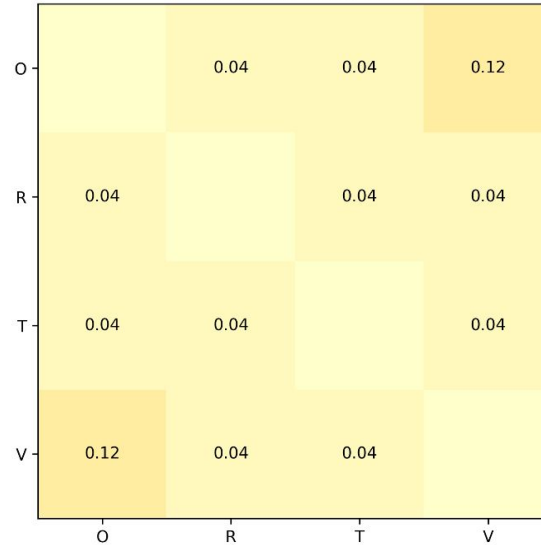


Mejores set de letras

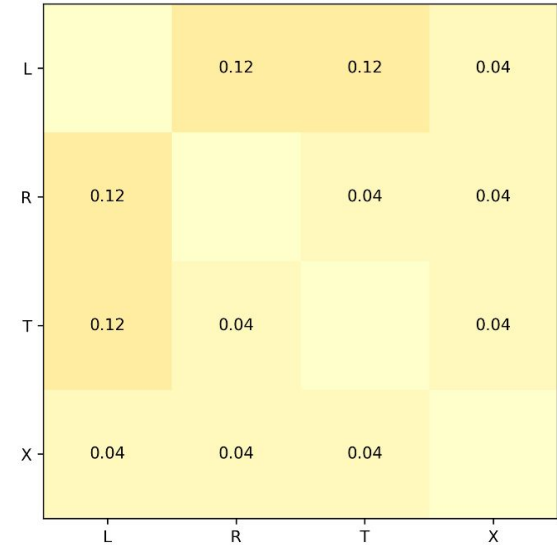
Grupo 1: C, R, T, V
Ortogonalidad total: 0.3200



Grupo 2: O, R, T, V
Ortogonalidad total: 0.3200

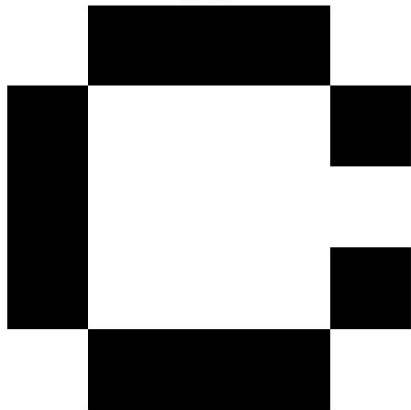


Grupo 3: L, R, T, X
Ortogonalidad total: 0.4000

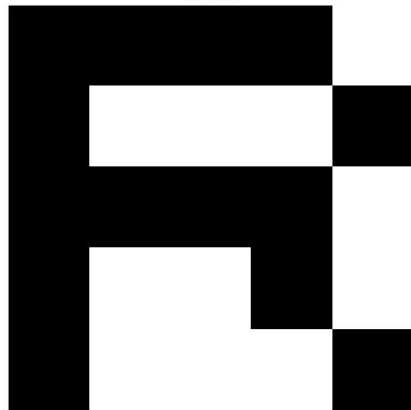


Set elegido

Letra C

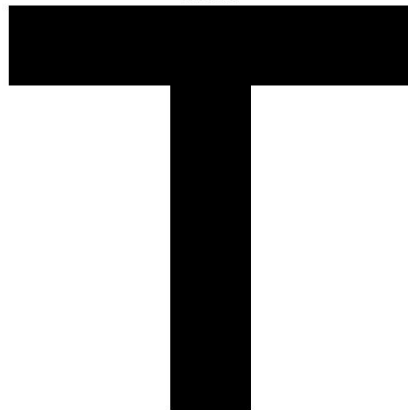


Letra R

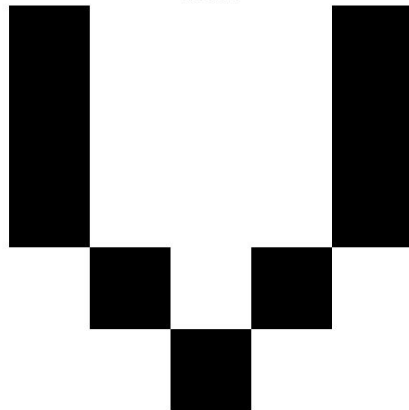


Patrones de las Letras

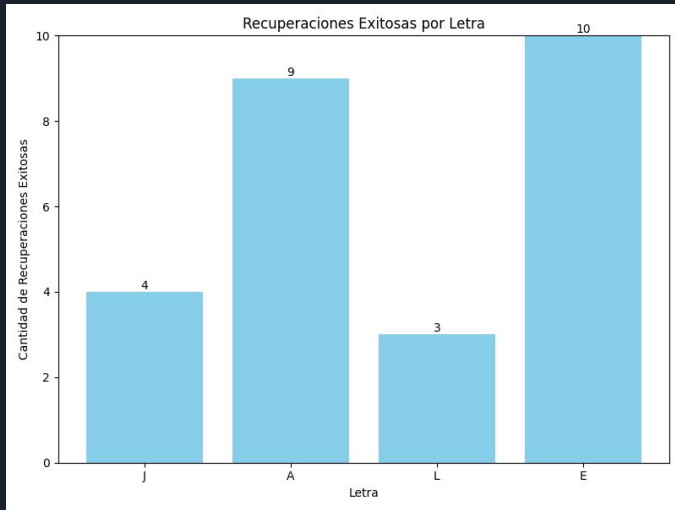
Letra T



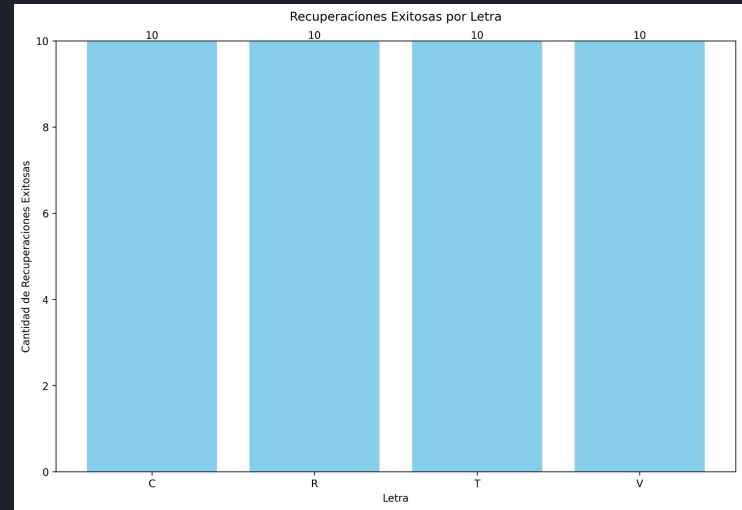
Letra V



10% de ruido

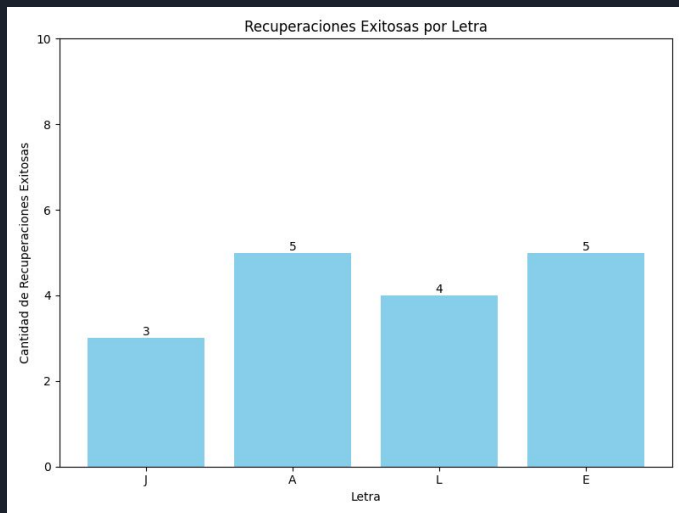


Anterior

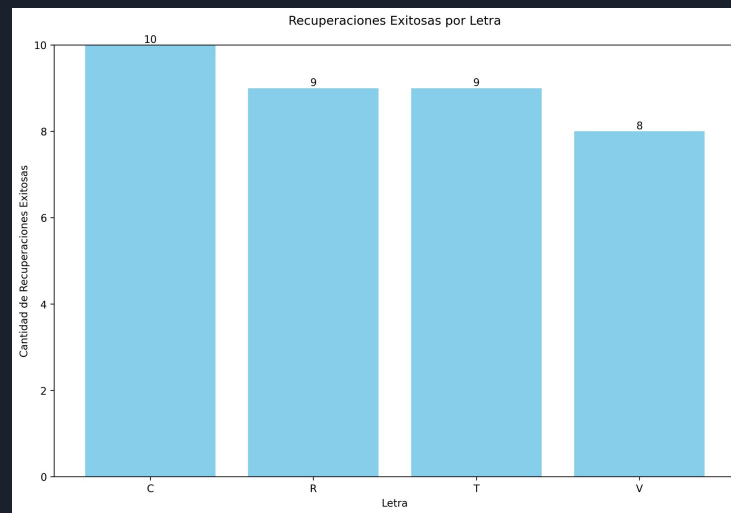


Nuevo set

25% de ruido

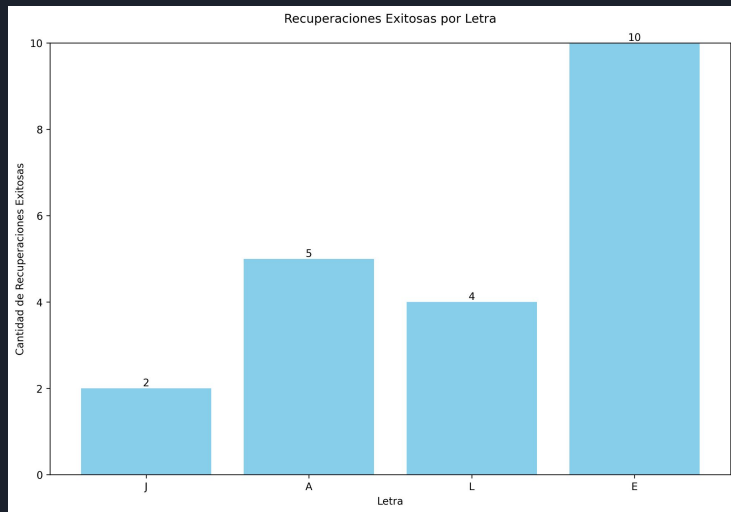


Anterior

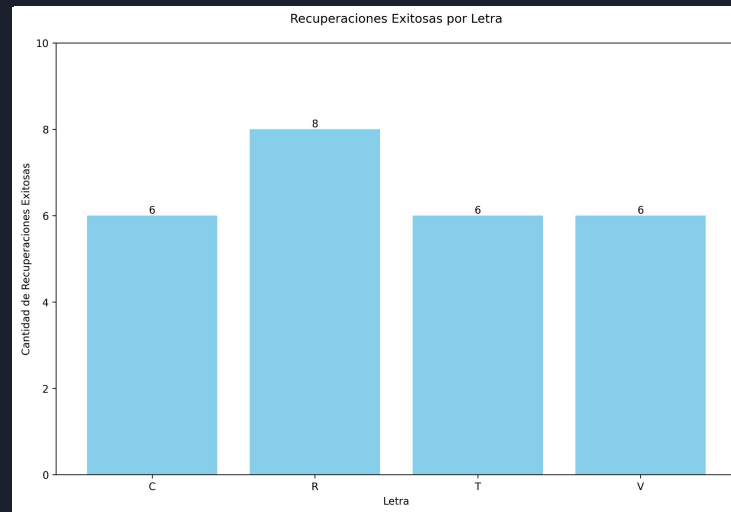


Nuevo set

35% de ruido

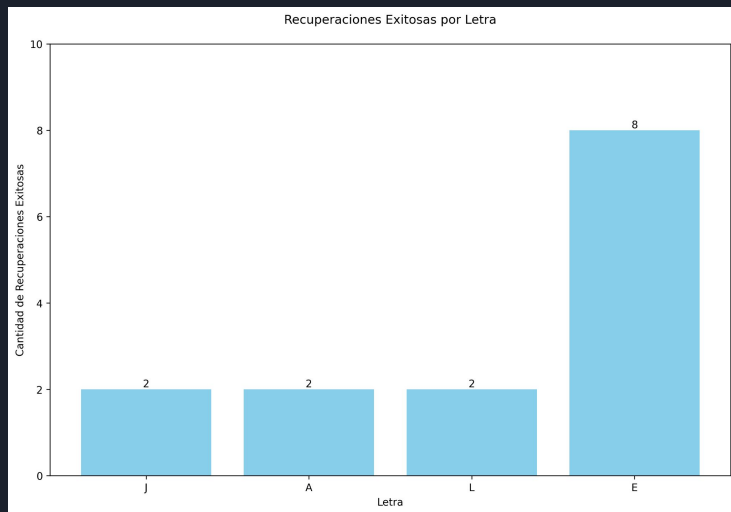


Anterior

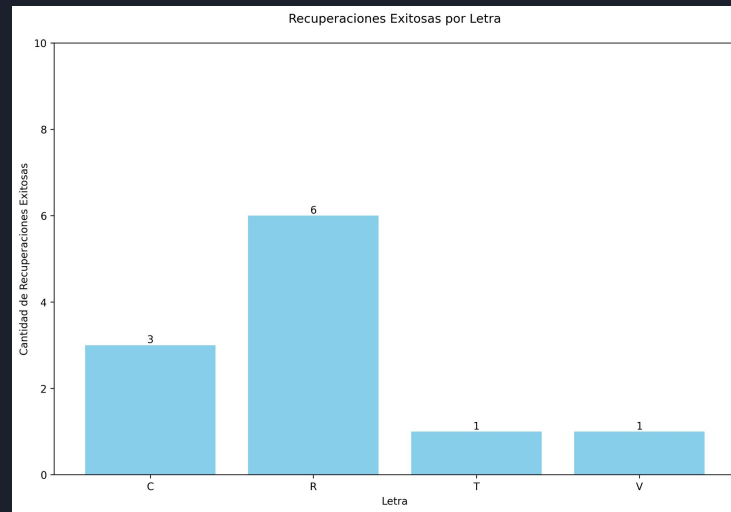


Nuevo set

40% de ruido

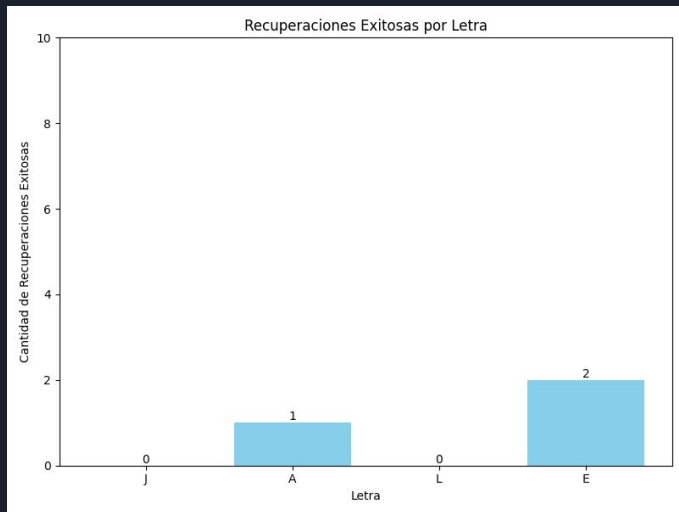


Anterior

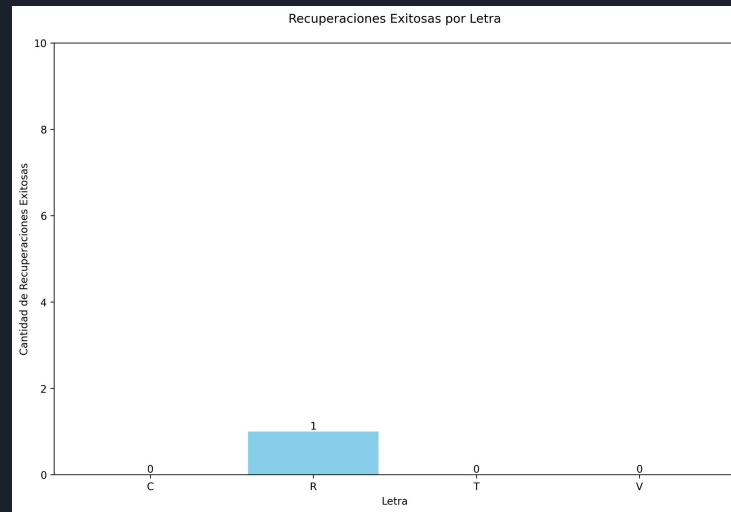


Nuevo set

50% de ruido

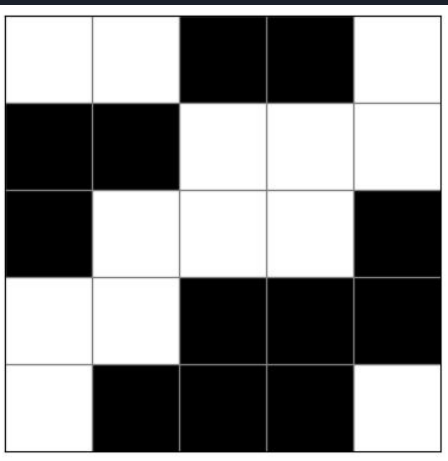


Anterior

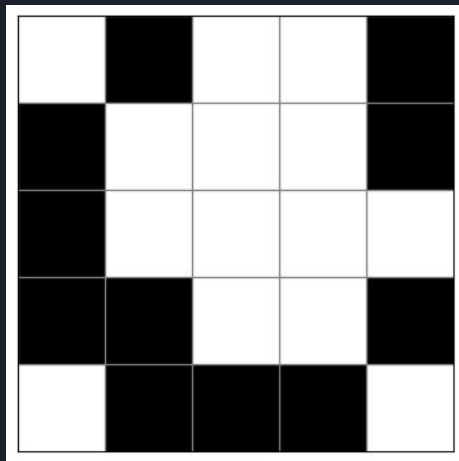


Nuevo set

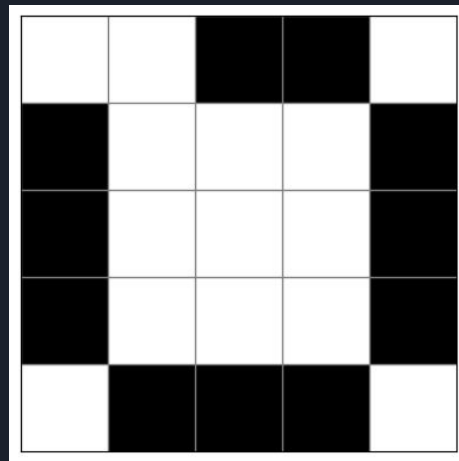
Evolución del estado



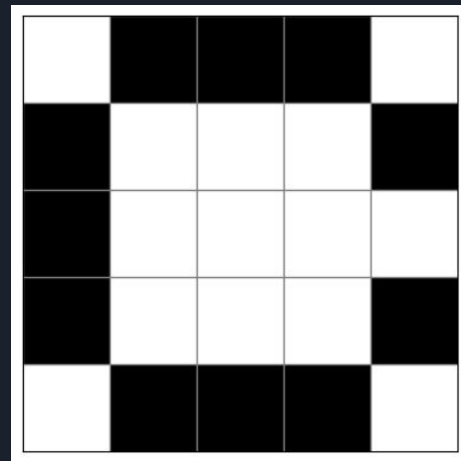
Paso 1 - Energía = -1.6



Paso 2 - Energía = -6.08



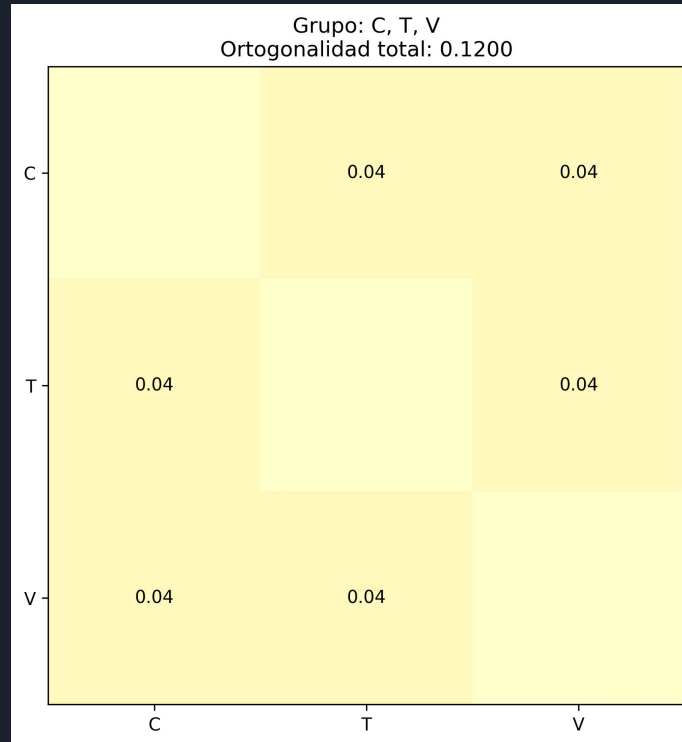
Paso 3 - Energía = -7.52



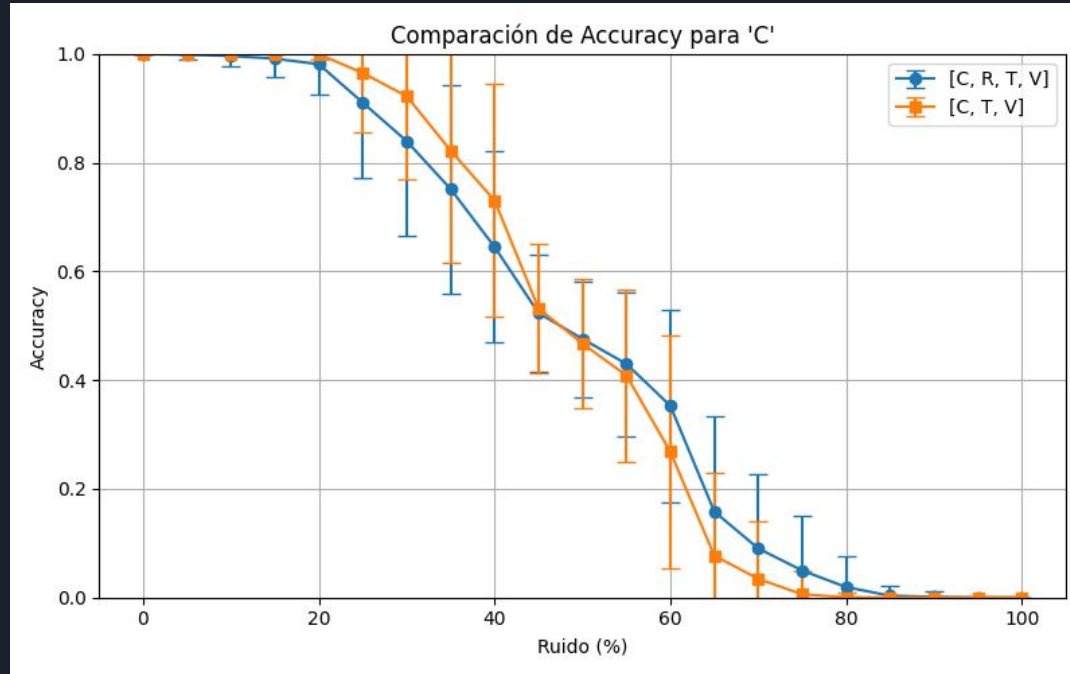
Paso 4 - Energía = -0.72

7 pixeles de ruido ~ 28%

Mejor set de 3 letras

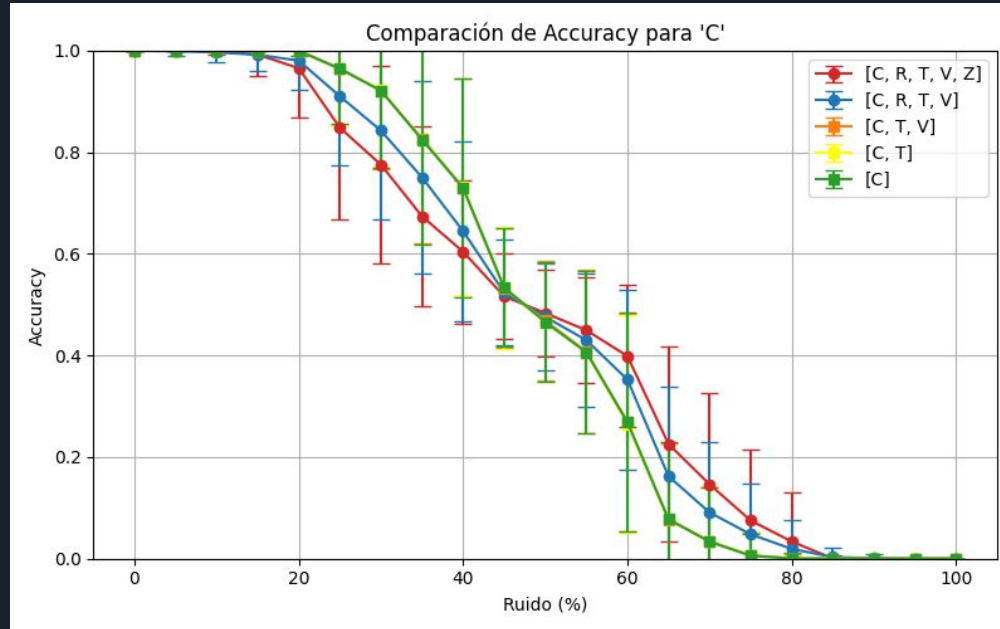


Comparación distinta cantidad de patrones almacenados



(10.000 ejec)

Comparación distinta cantidad de patrones almacenados



(10.000 ejec)



Capacidad teórica Hopfield

- Según Amit et al. (1985), el número máximo de patrones que se puede almacenar en una red según es:

$$p \leq 0,138N$$

Fuente: Amit, D.J., Gutfreund, H. & Sompolinsky, H. (1985). *Spin-glass models of neural networks*. Phys. Rev. A, 32(2), 1007-1018.



Utilizando regla de capacidad teórica

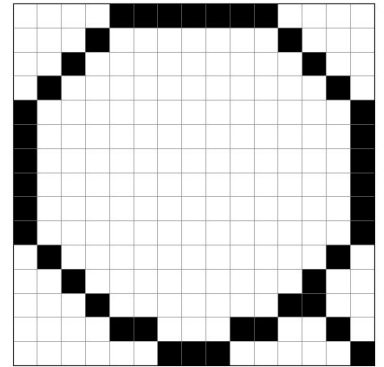
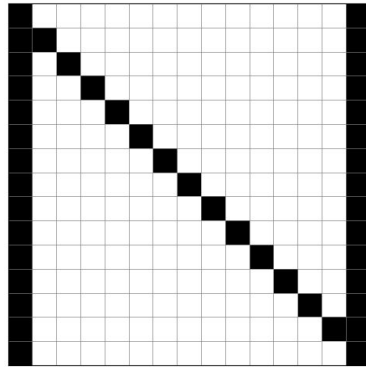
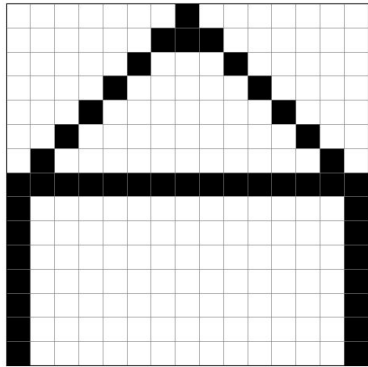
Si quisiéramos almacenar un abecedario completo de 26 letras en una red de Hopfield:

$$N \geq \sqrt{\frac{26}{0.138}}$$

N tiene que ser mayor o igual a 13,72, tomamos N= 15

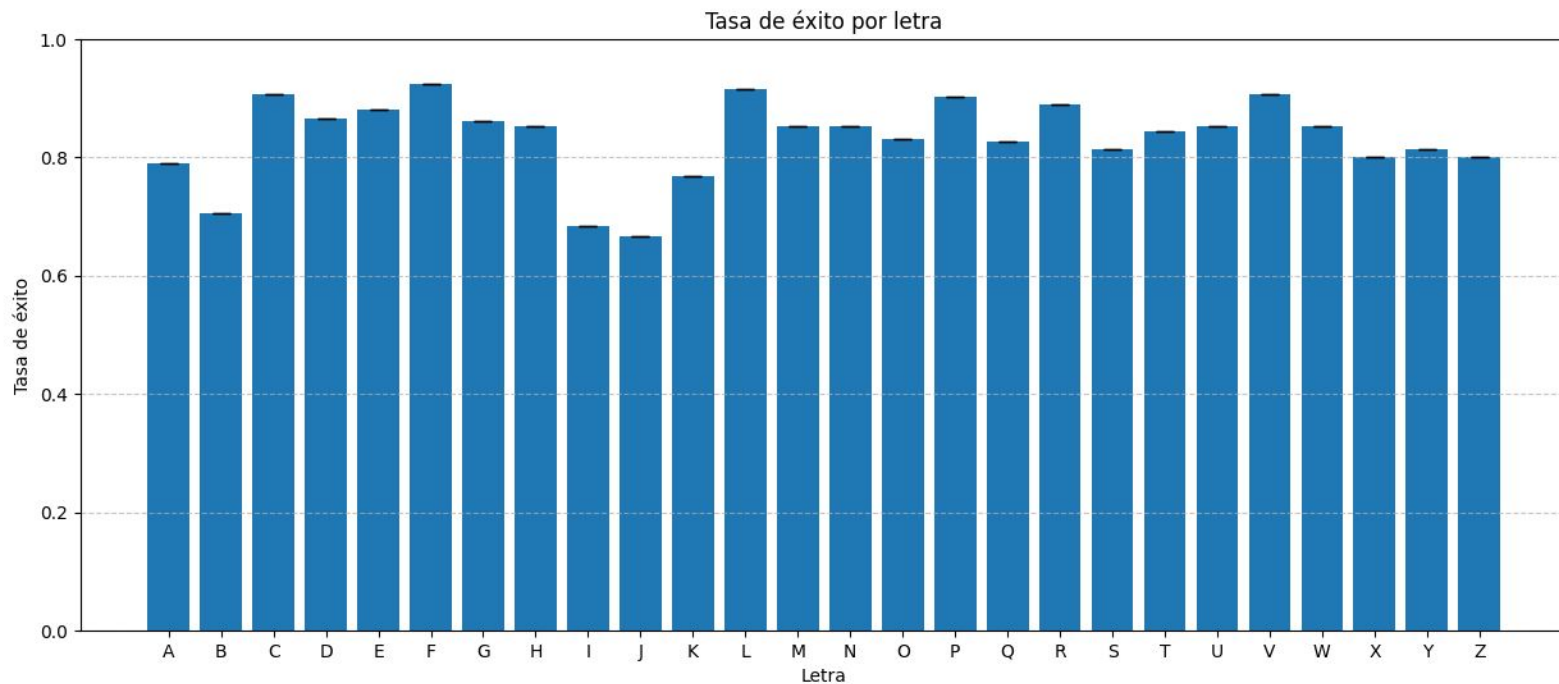
Abecedario 15x15

- Letras de 225 píxeles

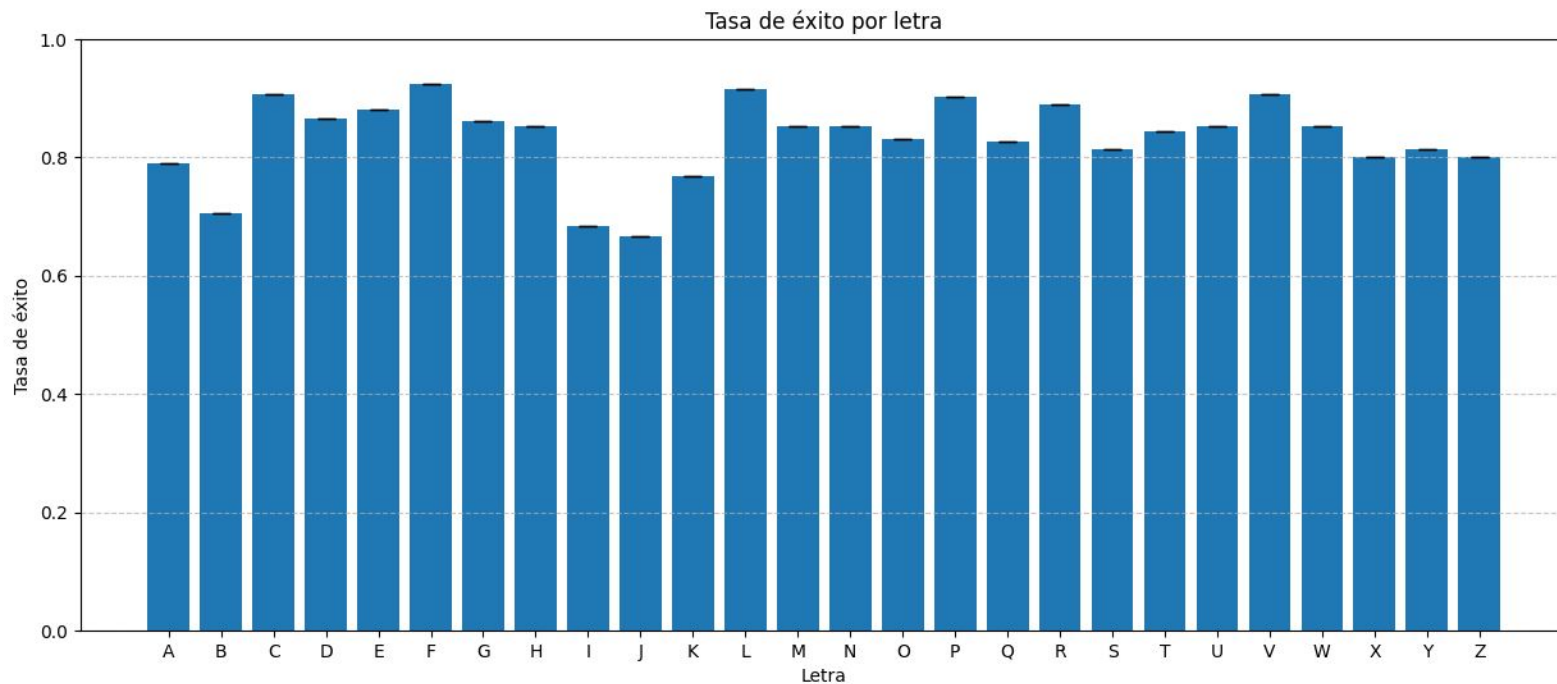




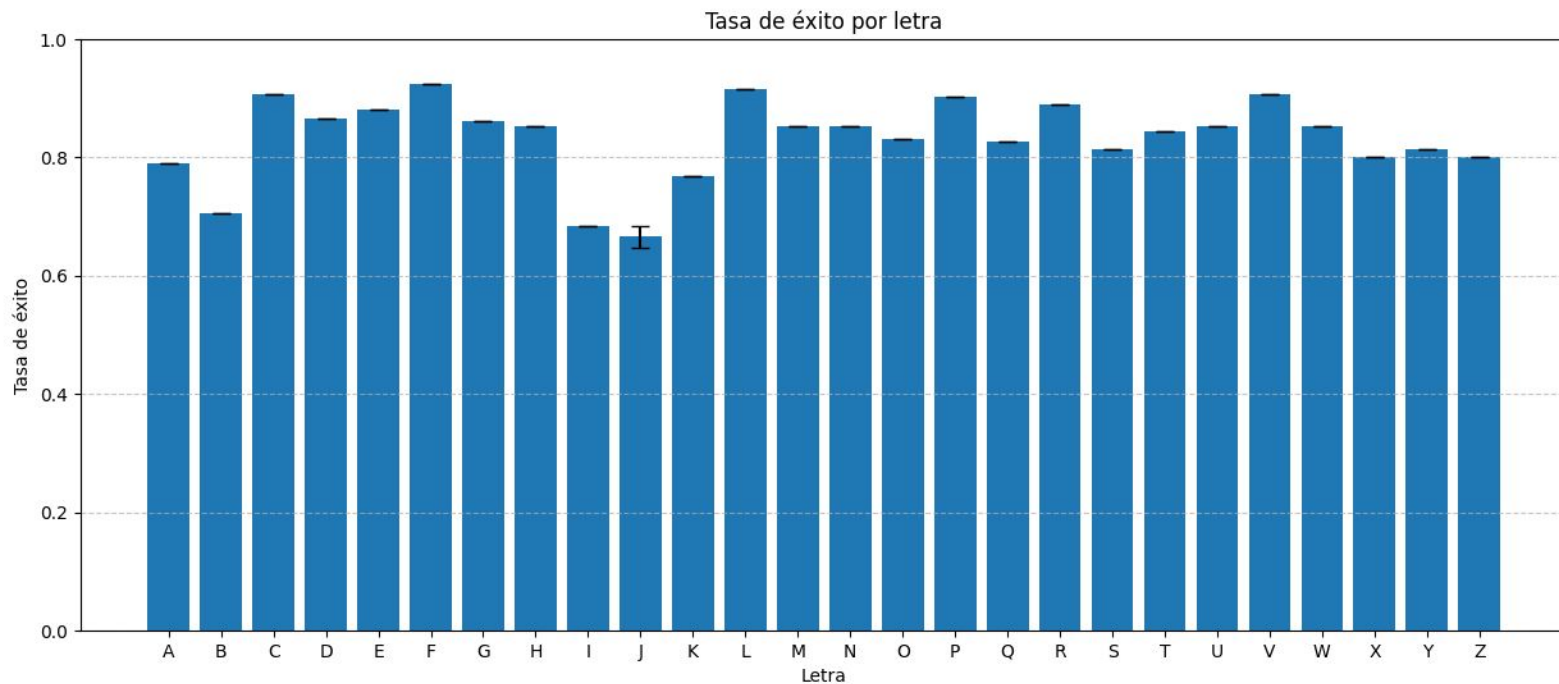
25 pixeles ~ 10% ruido



50 píxeles ~ 20% ruido

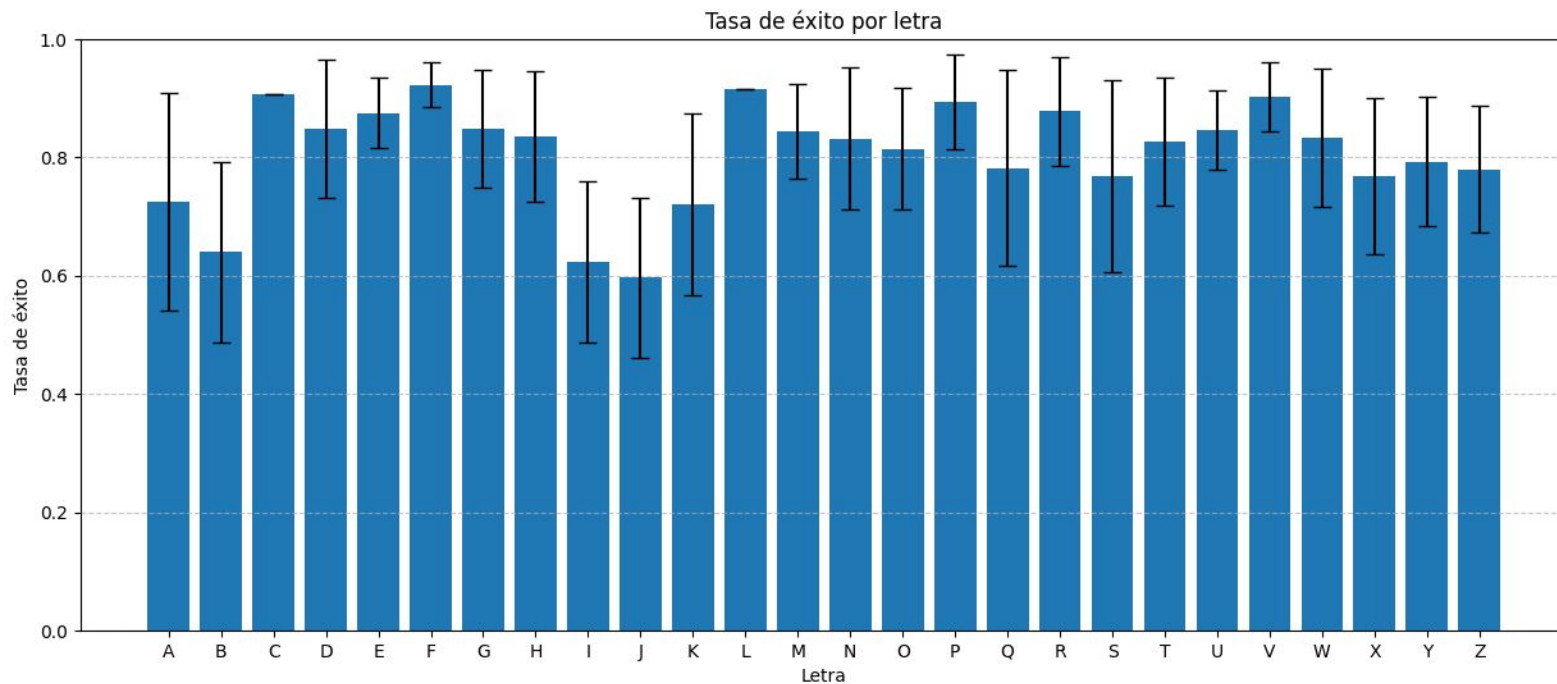


75 píxeles ~ 30% ruido

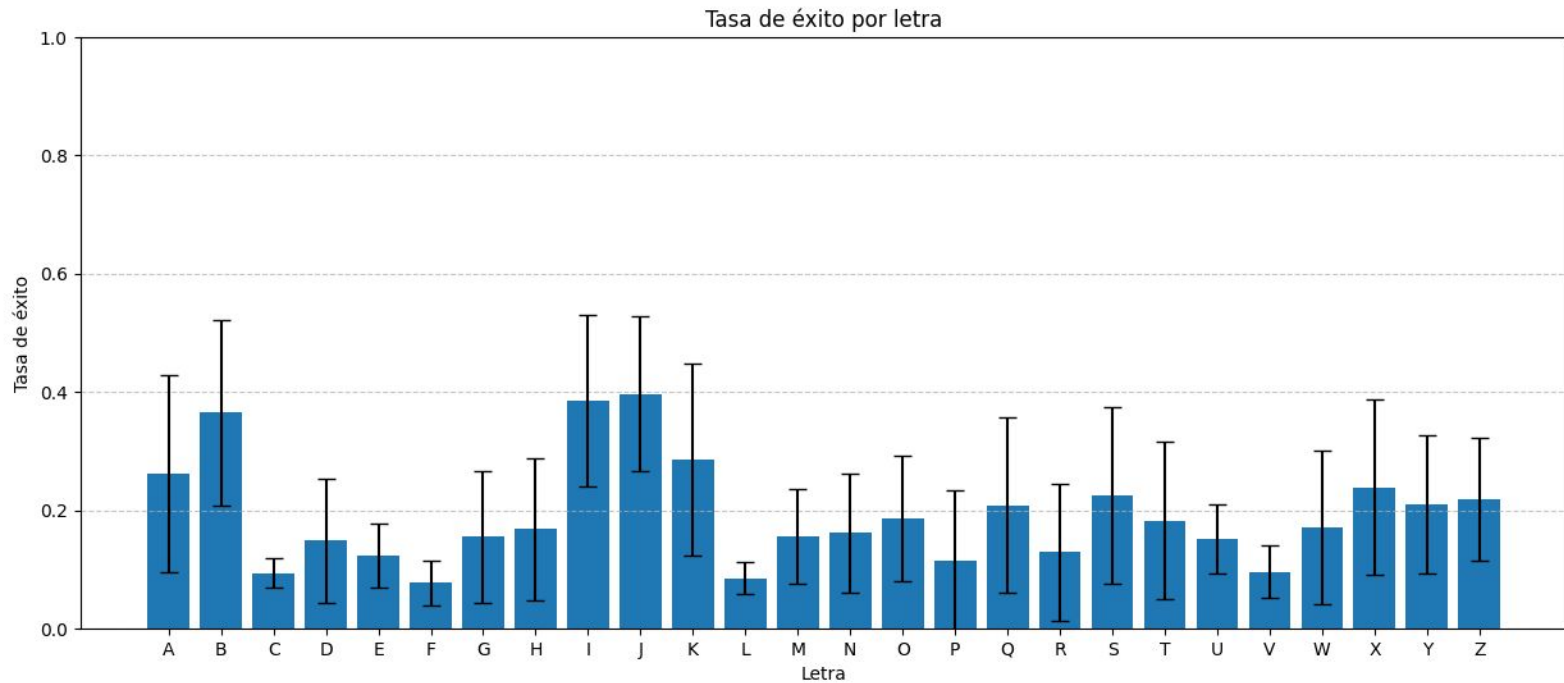




100 pixeles ~ 40% ruido



125 píxeles ~ 50% ruido





Conclusiones

- El comportamiento de la red es sensible a la correlación entre patrones almacenados
- Con bajos niveles de ruido y patrones ortogonales la red tiene un buen desempeño