

TP4: Métodos de Aprendizaje **NO** Supervisado





Grupo 7



Julián Francisco **Arce**



Ignacio Agustín **Manfredi**



Gian Luca **Pecile**



Implementación





1

Introducción



2

Kohonen



3

Oja



4

Hopfield



1

Introducción



Introducción

- El aprendizaje no supervisado es un método de aprendizaje en el cual la variable respuesta no es una información disponible.
- Se busca resolver mediante la red neuronal de **Kohonen**, el modelo de **Hopfield** y la regla de **Oja** problemas de:
 - Agrupamiento
 - Asociación
 - Reducción de Dimensionalidad

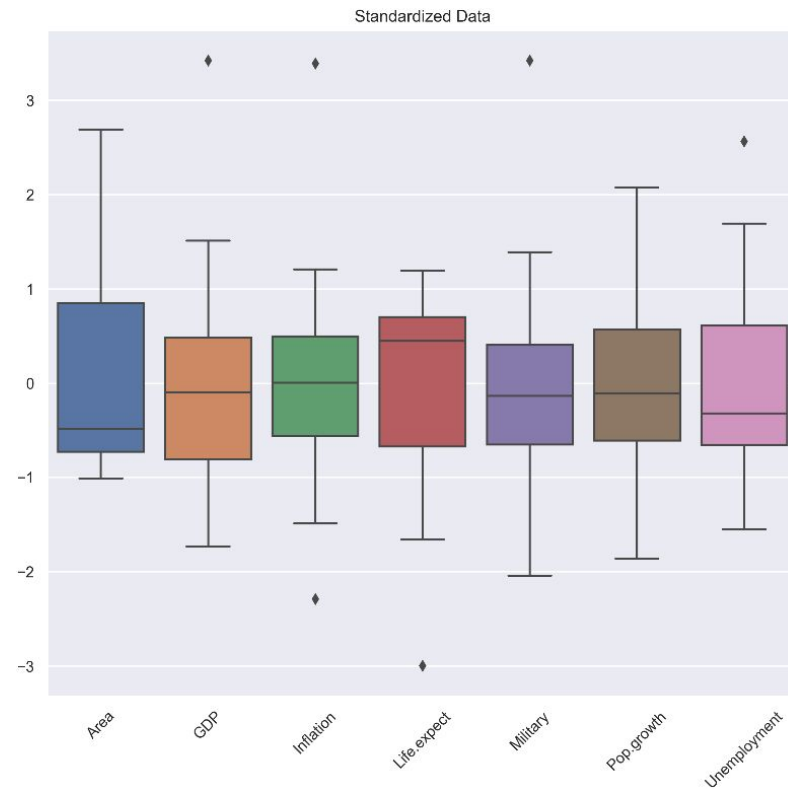
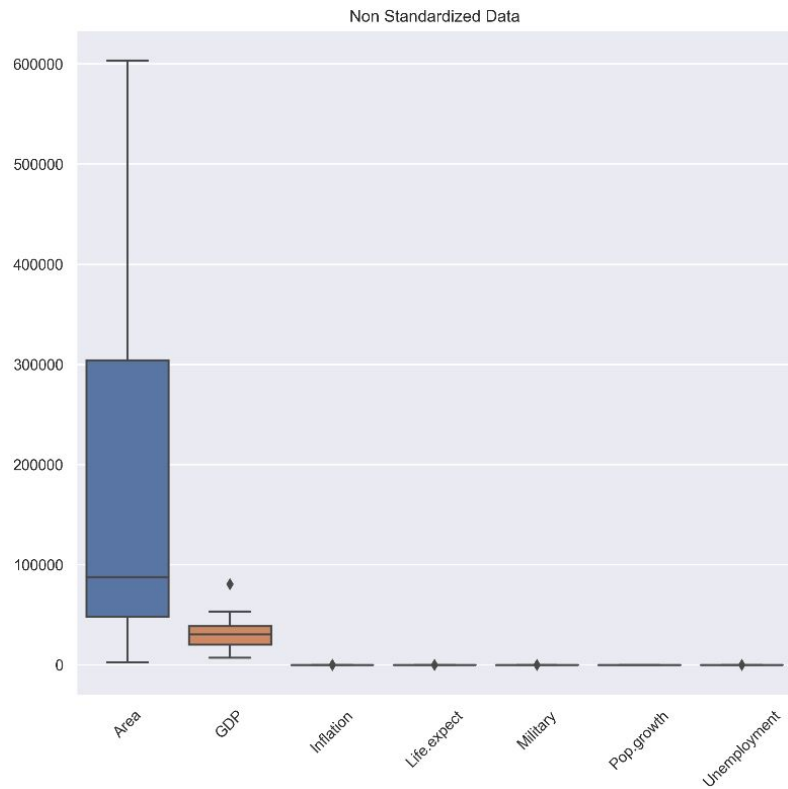


Ejercicio 1

- Se utilizó un conjunto de datos de 28 países de Europa, que detalla las características sociales, económicas y geográficas de cada uno. Las variables a considerar son:
 - *Country*: Nombre del país
 - *Area*
 - *GDP*: producto bruto interno
 - *Inflation*: inflación anual
 - *Life.expect*: Expectativa de vida media en años
 - *Military*
 - *Pop.growth*: tasa de crecimiento poblacional
 - *Unemployment*



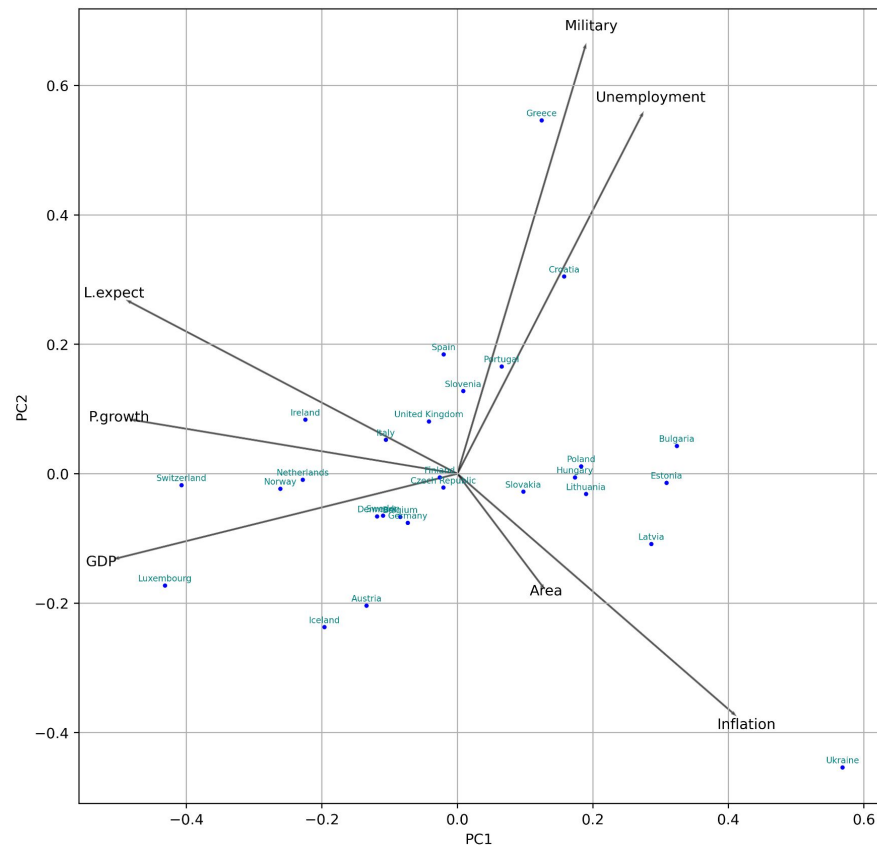
Estandarización de variables



Data no estandarizada vs estandarizada.



Primera y Segunda componentes (Ej2)



Biplot de **PC1** vs **PC2**.



2

Ejercicio 1 Parte A

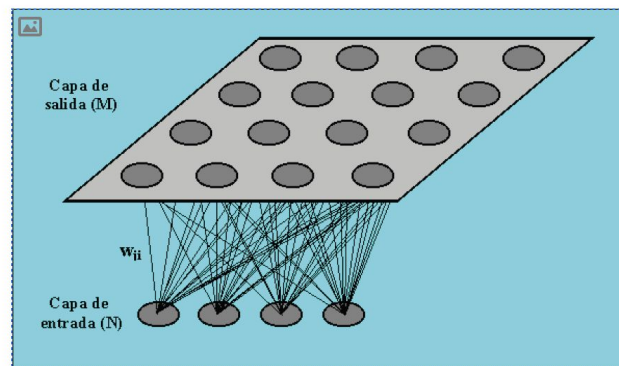
Kohonen

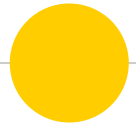
Implementar la red de Kohonen y aplicarla para poder asociar países que posean la mismas características geopolíticas, económicas y sociales



Kohonen

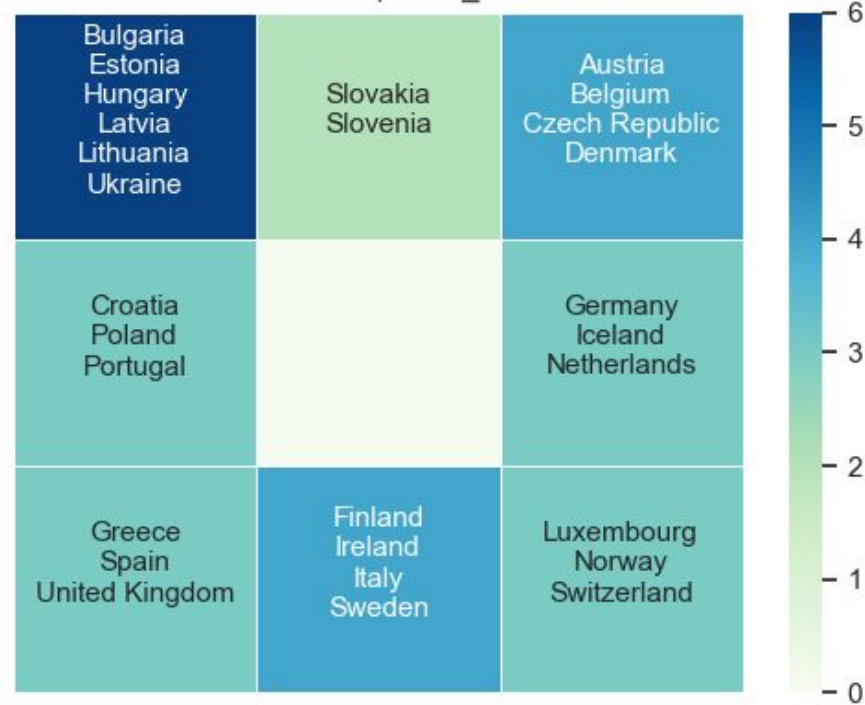
- Es una red neuronal con una única capa de entrada y de salida.
- Las capas de la neurona de salida están conectadas positivamente consigo mismas y negativamente con las demás.
- Intentan atraer a los registros de una sola neurona que van a ser aquellos que sean "parecidos" y pertenezcan a un mismo grupo.
- Cada una de las entradas se conecta a todas las neuronas de la capa de salida.
- Se trata de una red en el que cada una de las neuronas compete para ganar cada una de las entradas. La neurona ganadora será aquella que dada una entrada x , tenga el vector w más parecido a ella.



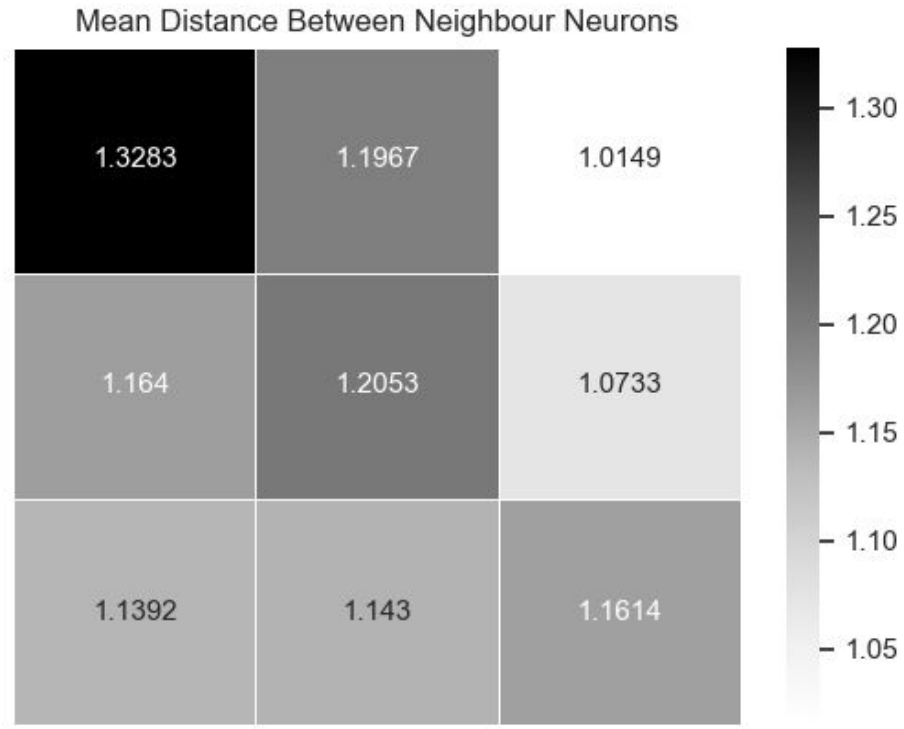


Resultados

Associated Countries per Neuron
 $\eta = 0.001$ | $\text{initial_r} = 4$

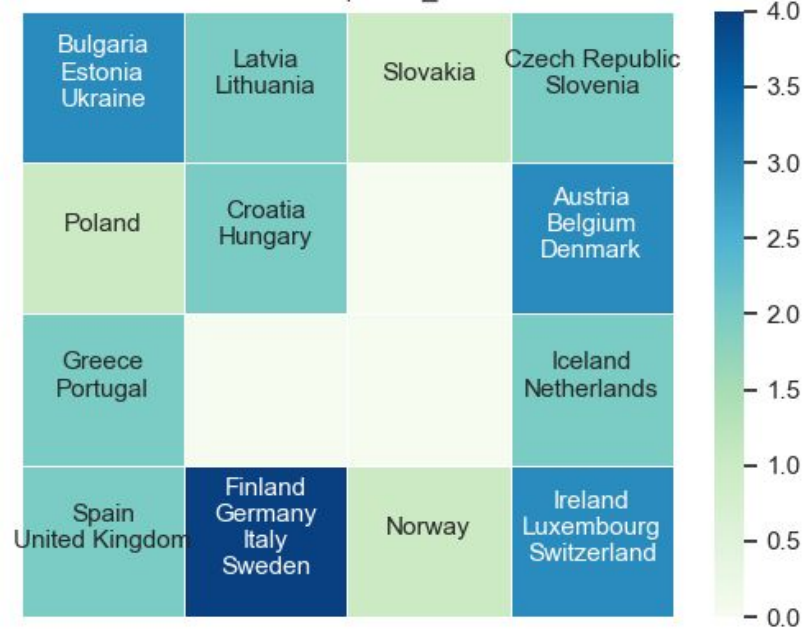


Asociación de países en matriz de 3x3.

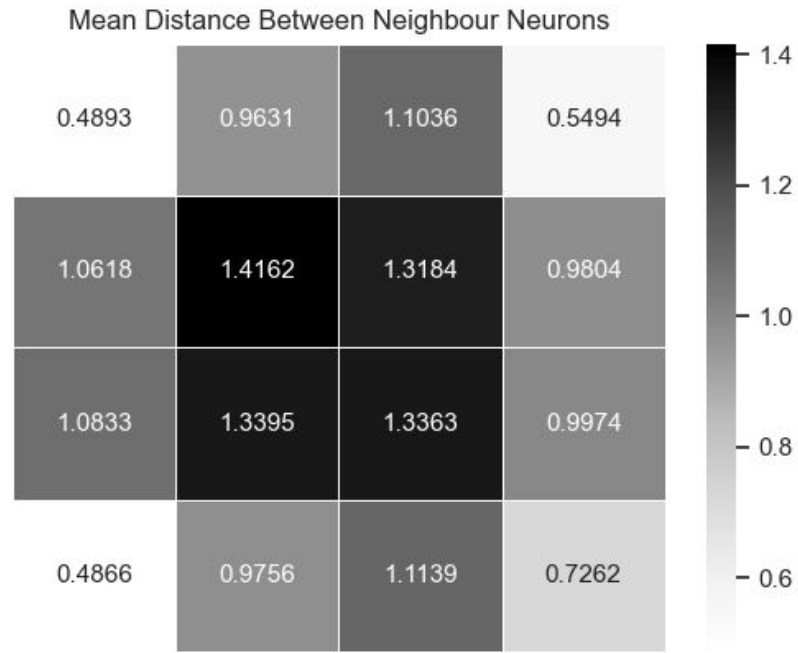


Distancia media entre neuronas en matriz de 3x3.

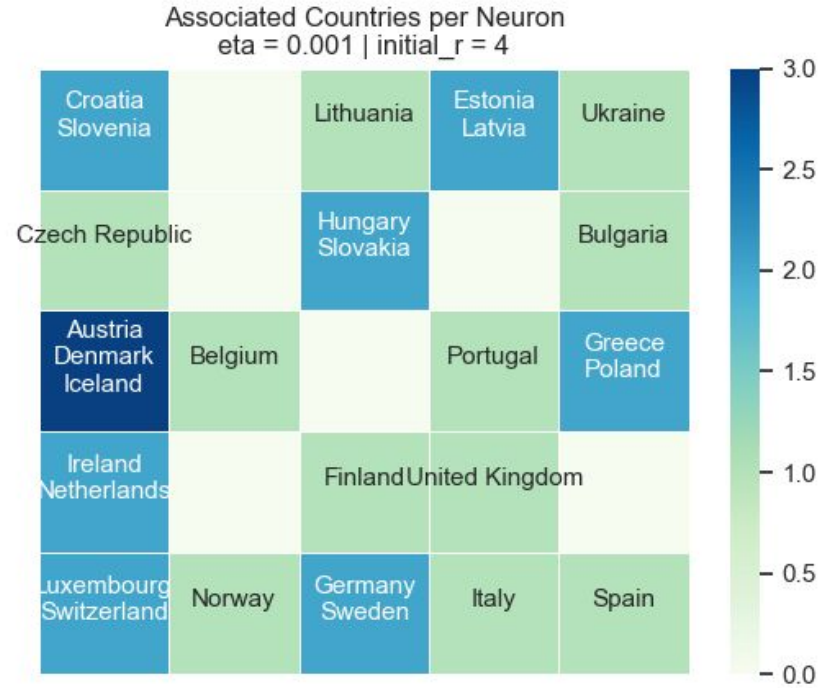
Associated Countries per Neuron
eta = 0.001 | initial_r = 4



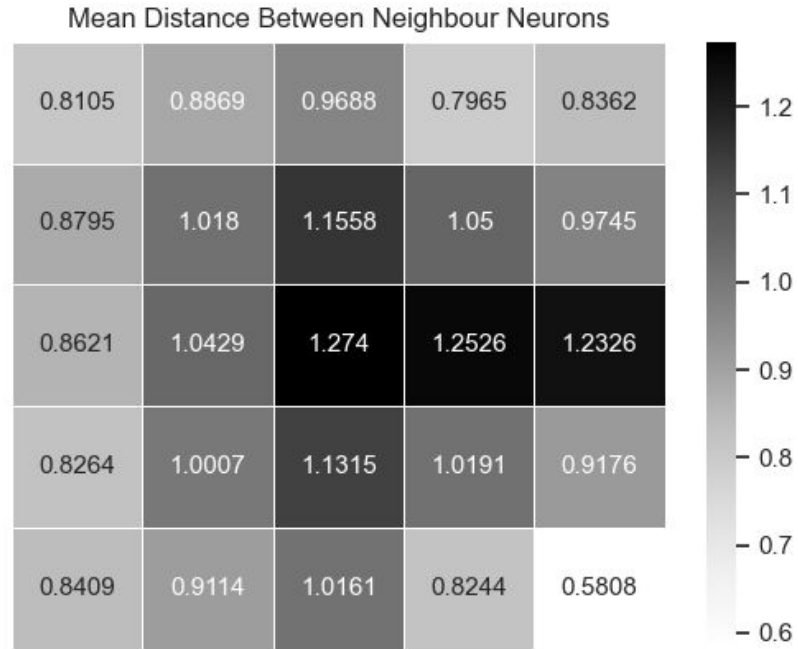
Asociación de países en matriz de 4x4.



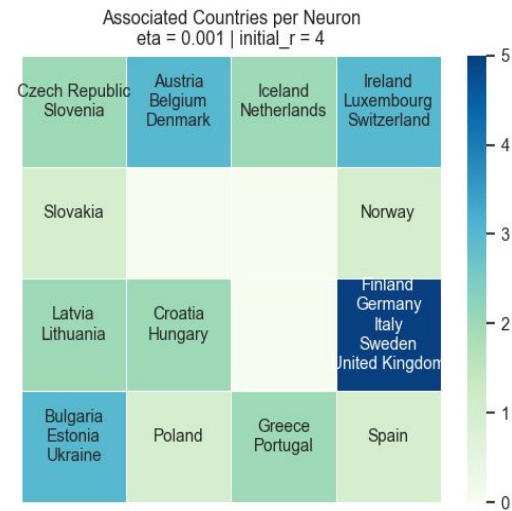
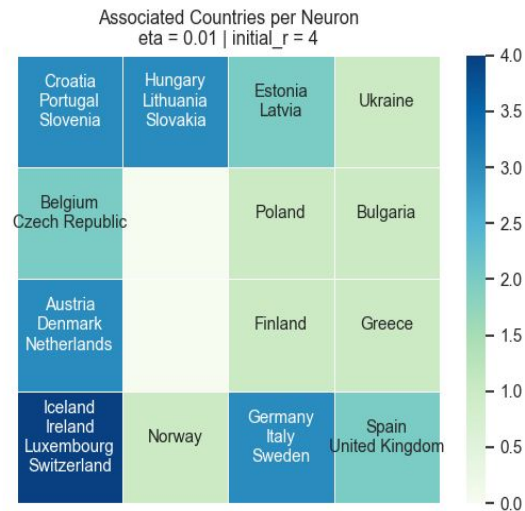
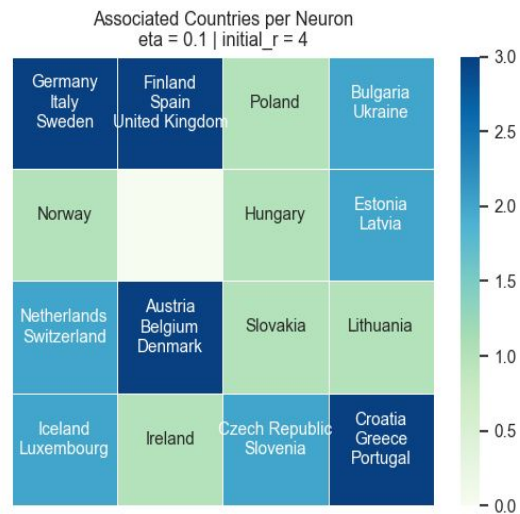
Distancia media entre neuronas en matriz de 4x4.



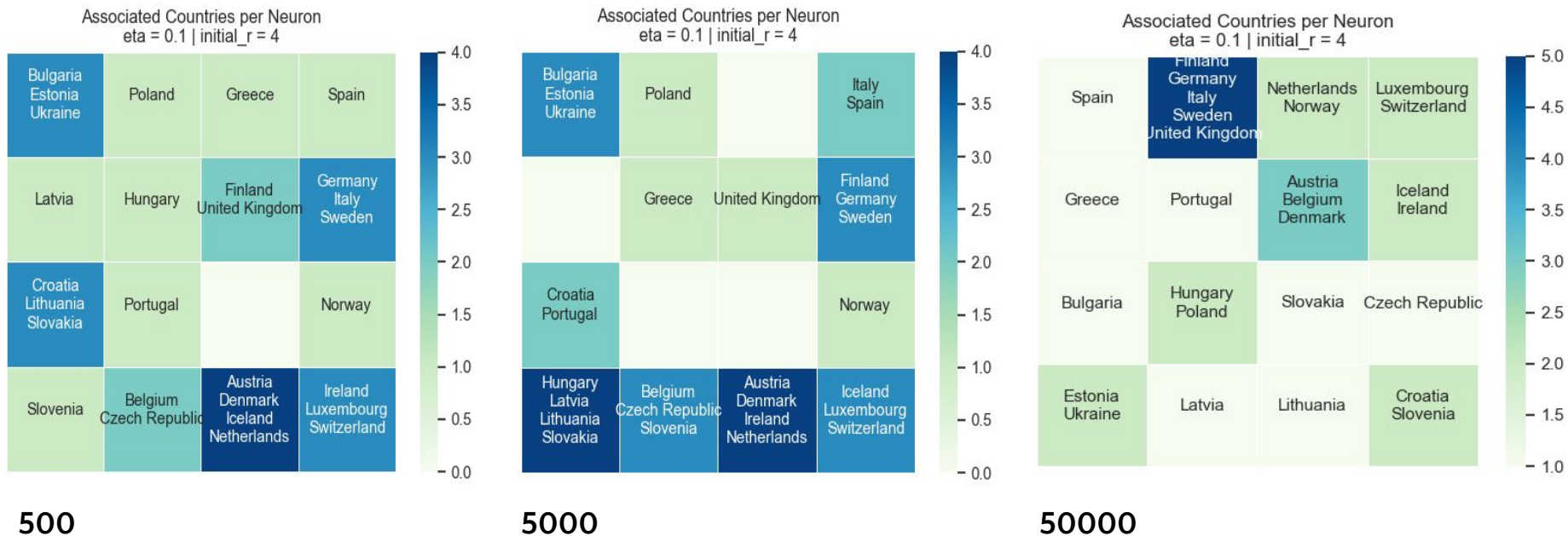
Asociación de países en matriz de 5x5.



Distancia media entre neuronas en matriz de 5x5.



Distintas **tasas de aprendizaje.**



Distinta cantidad de iteraciones.



Conclusiones

- En una matriz de 3x3 se obtiene una menor cantidad de unidades muertas. Sin embargo, esto dependerá de las entradas y del problema a resolver.
- Se consigue una buena agrupación de países en comparación al biplot (componentes principales).
- Incrementar las capas disminuye la distancia media entre neuronas vecinas.
- A menor η , los países se encuentran más agrupados.
- A mayor cantidad de iteraciones, los países se encuentran menos agrupados.
- Muy importante **estandarizar los datos**.



2

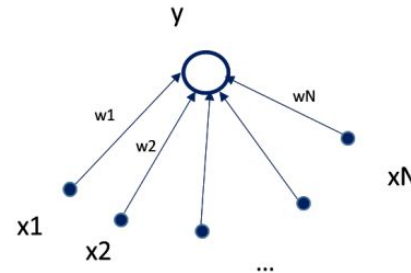
Ejercicio 1 Parte B

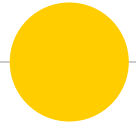
Regla de Oja

A partir de la red neuronal basada en la regla de Oja, calcular la primera componente principal para el conjunto de datos e interpretarla.

Oja

- Luego de varias iteraciones el método converge al autovector correspondiente al mayor autovalor de la matriz de correlaciones de los datos de entrada.
- Con este vector w_{final} y los datos de entrada se construye la primera componente principal.
- Las componentes principales permiten extraer características importantes de un conjunto de datos reduciendo su dimensionalidad, permitiendo realizar un análisis de forma más simple y fácil de comprender.





Resultados

Eigenvector

	Library	Oja	Error
Area	0.124874	0.125607	0.000733
GDP	-0.500506	-0.500438	0.000068
Inflation	0.406518	0.407256	0.000738
Life.expect	-0.482873	-0.483004	0.000131
Military	0.188112	0.187564	0.000548
Pop.growth	-0.475704	-0.475528	0.000176
Unemployment	0.271656	0.271297	0.000359

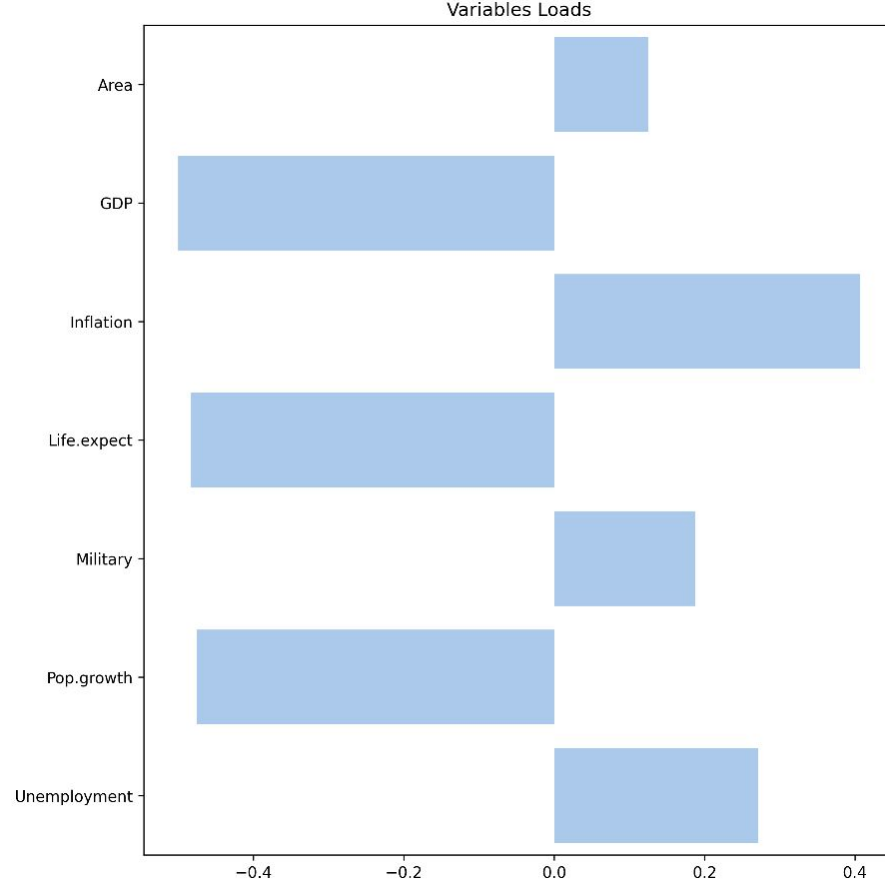
Comparación del autovector **Implementación de Oja vs Librería.**

First Component

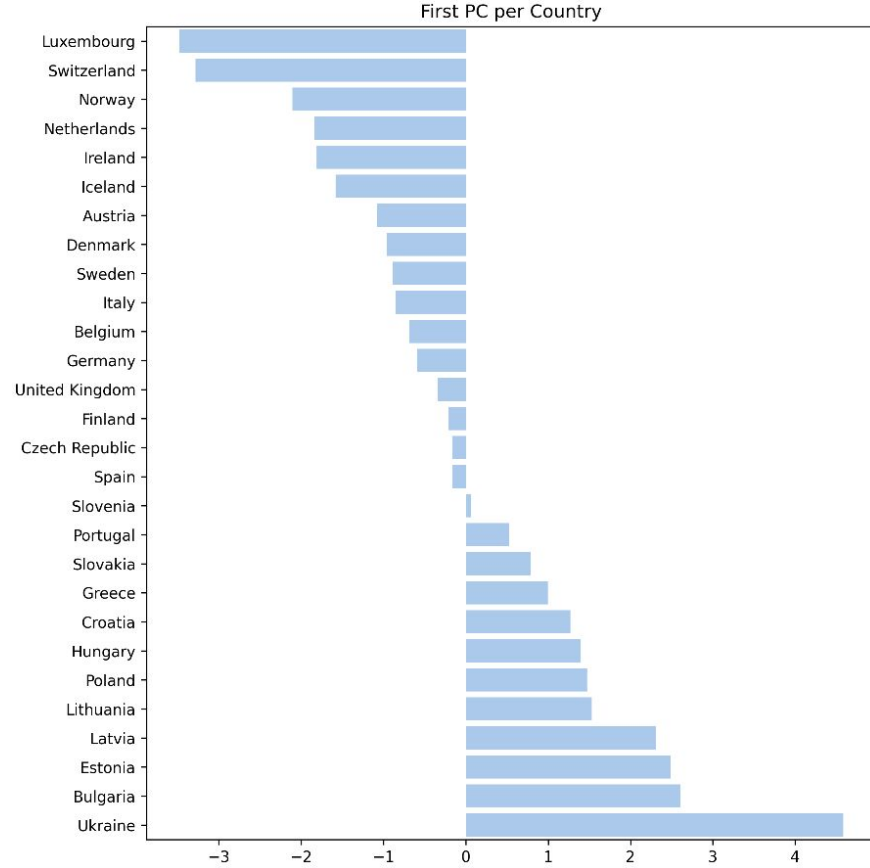
	Library	Oja	Error
Luxembourg	-3.478435	-3.477793	0.000642
Switzerland	-3.281586	-3.282642	0.001056
Norway	-2.106511	-2.106499	0.000012
Netherlands	-1.840053	-1.840691	0.000637
Ireland	-1.808918	-1.809303	0.000386
Iceland	-1.583720	-1.582231	0.001488
Austria	-1.081748	-1.081087	0.000661
Denmark	-0.955191	-0.955486	0.000296
Sweden	-0.885105	-0.883813	0.001292
Italy	-0.853224	-0.852938	0.000286
Belgium	-0.681094	-0.681253	0.000159
Germany	-0.592394	-0.591777	0.000617
United Kingdom	-0.340819	-0.340378	0.000442
Finland	-0.210563	-0.209971	0.000592

	Library	Oja	Error
Czech Republic	-0.167209	-0.168031	0.000821
Spain	-0.163767	-0.162953	0.000814
Slovenia	0.067543	0.065746	0.001797
Portugal	0.526493	0.525608	0.000885
Slovakia	0.782966	0.782886	0.000080
Greece	1.000472	0.997708	0.002764
Croatia	1.270149	1.267890	0.002259
Hungary	1.396898	1.396655	0.000243
Poland	1.471774	1.472448	0.000674
Lithuania	1.530100	1.530019	0.000081
Latvia	2.306059	2.306185	0.000125
Estonia	2.487735	2.487465	0.000270
Bulgaria	2.609879	2.609188	0.000691
Ukraine	4.580268	4.585050	0.004782

Primer componente de cada país **Implementación de Oja** vs **Liberia**.



Cargas de cada una de las variables ($\eta = 0.0001$).



Primer componente de cada país ($\eta^2 = 0.0001$).



Conclusiones

- GDP, esperanza de vida y crecimiento poblacional aportan una carga **negativa**.
- Area, inflación, desempleo y tropas militares aportan carga **positiva**.
- La diferencia (**error**) entre implementación y la librería es:
 - **Menor** a 10^{-3} para los **autovectores**.
 - **Menor** a 10^{-2} para la **primera componente**.
- Es muy importante estandarizar los datos.
- La tasa de aprendizaje debe ser **baja**.



4

Ejercicio 2

Hopfield

Almacenar 4 patrones de letras y realizar un programa que aplique el modelo de Hopfield para asociar patrones ruidosos con los patrones de letras almacenados.



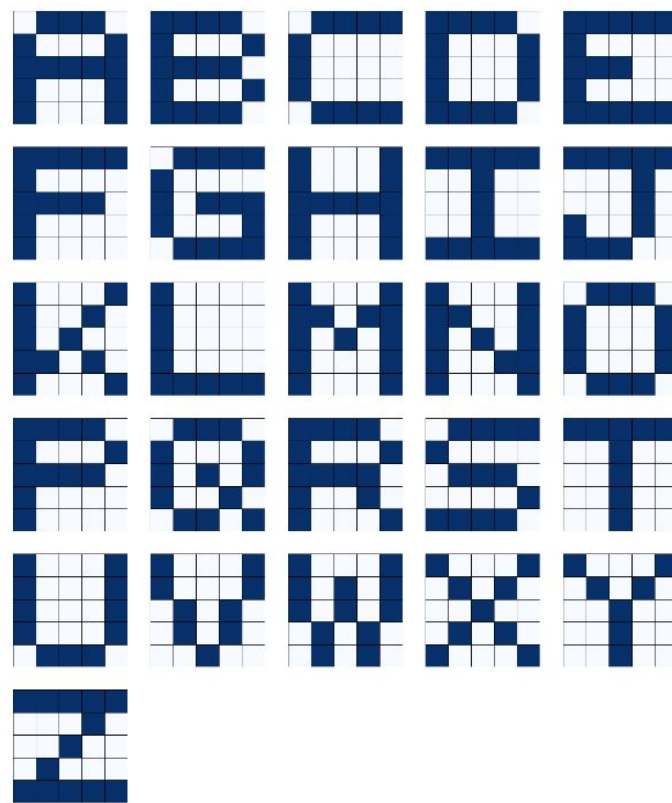
Hopfield

- Todas las neuronas están conectadas entre sí.
- Ninguna neurona está conectada con sí misma.
- El conjunto permitido de valores de entrada y salida es $\{-1, 1\}$.
- Todas las neuronas están en una sola capa de entrada y salida.
- La red está asociada a una función de energía, que siempre decrece (o se mantiene constante) cuando el sistema evoluciona.



Características

- Los mínimos locales de la función de energía son los patrones almacenados.
- Existen mínimos locales que no son los patrones almacenados, se llaman **estados espúreos**.
- Tiene capacidad de almacenamiento limitada. En este caso, se utiliza una red de 25 neuronas (el tamaño de cada patrón). Solo se puede aproximadamente el 15% del tamaño del patrón. Es por esa razón que se almacenan solamente 4 patrones.



Representación de Cada Letra

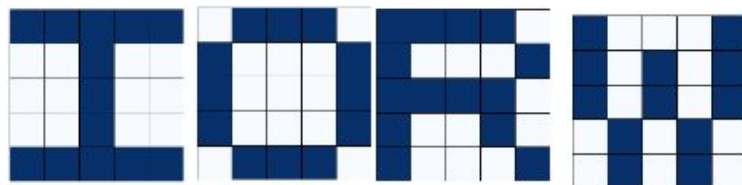


Ortogonalidad de los patrones

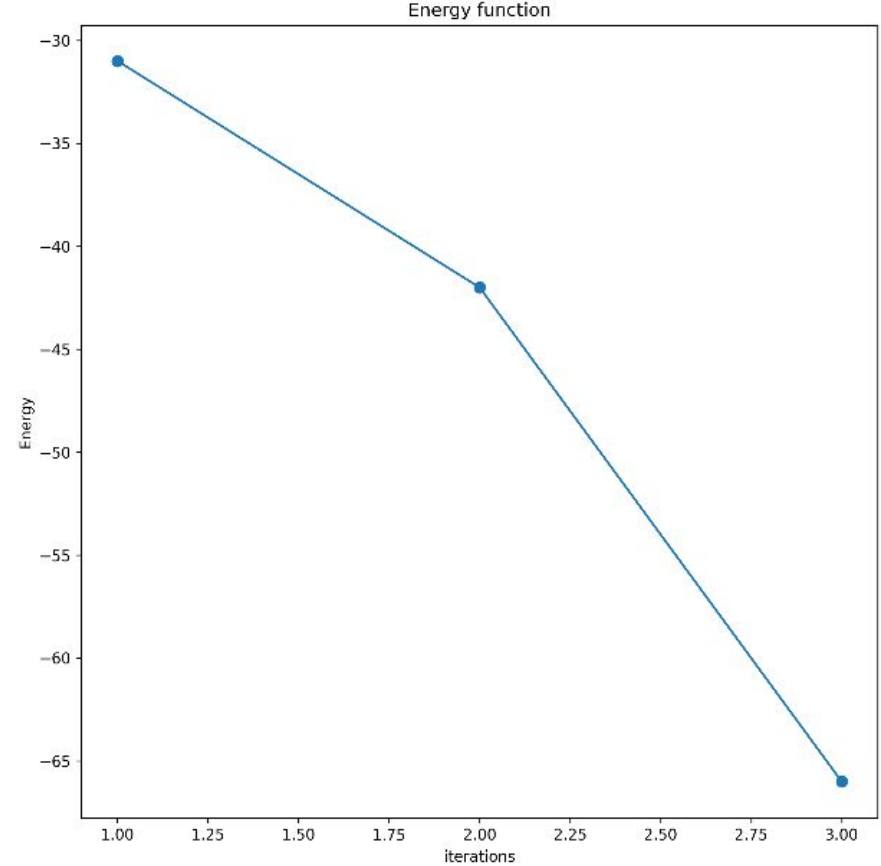
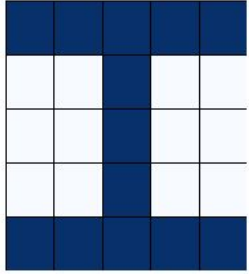
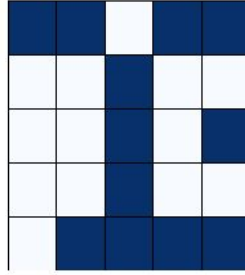
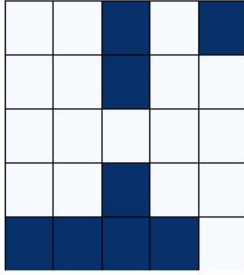
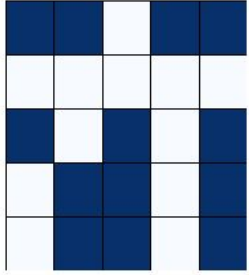
- $\{I, O, R, W\} \rightarrow 1.33$
- $\{O, R, T, W\} \rightarrow 1.33$
- $\{O, R, W, Z\} \rightarrow 1.33$
- $\{A, J, K, U\} \rightarrow 1.66$
- ...
- $\{A, F, P, R\} \rightarrow 17$
- $\{A, F, P, R\} \rightarrow 17.66$



Entrenamiento con patrón ortogonal

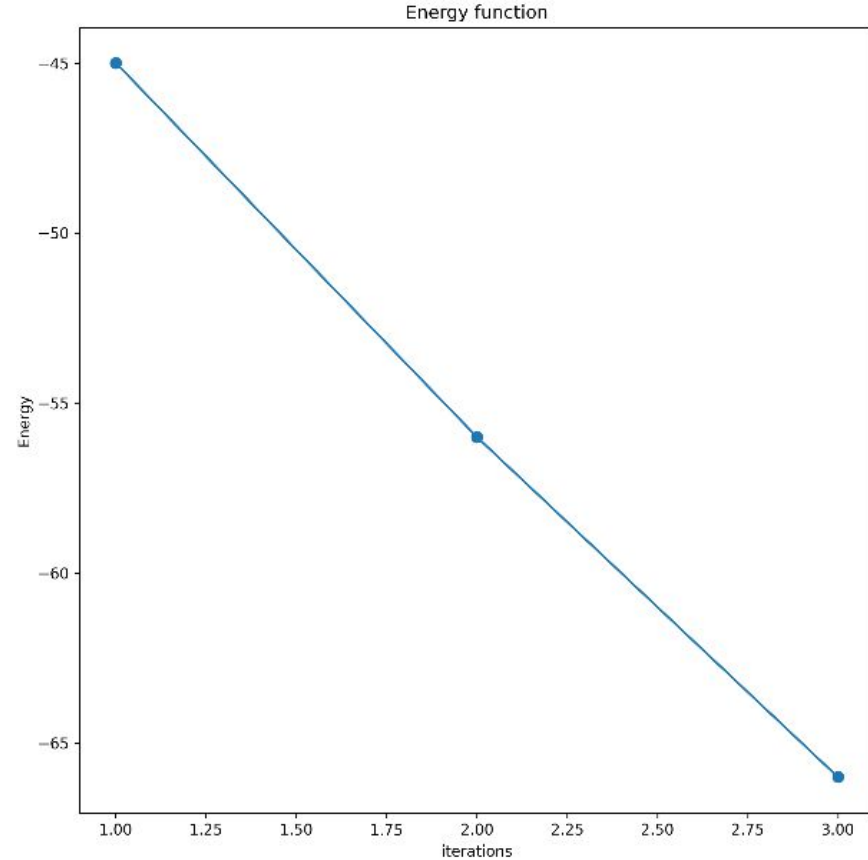
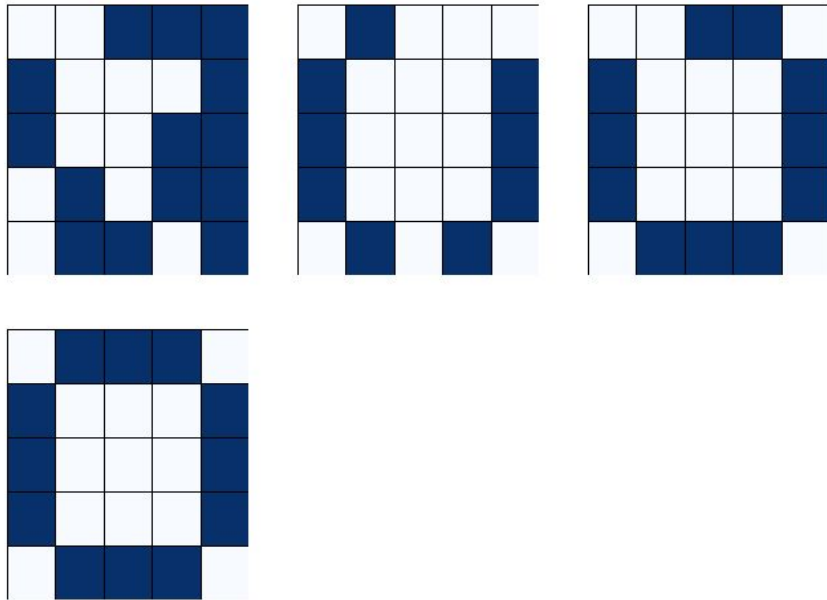


Letter: I | Noise: 8



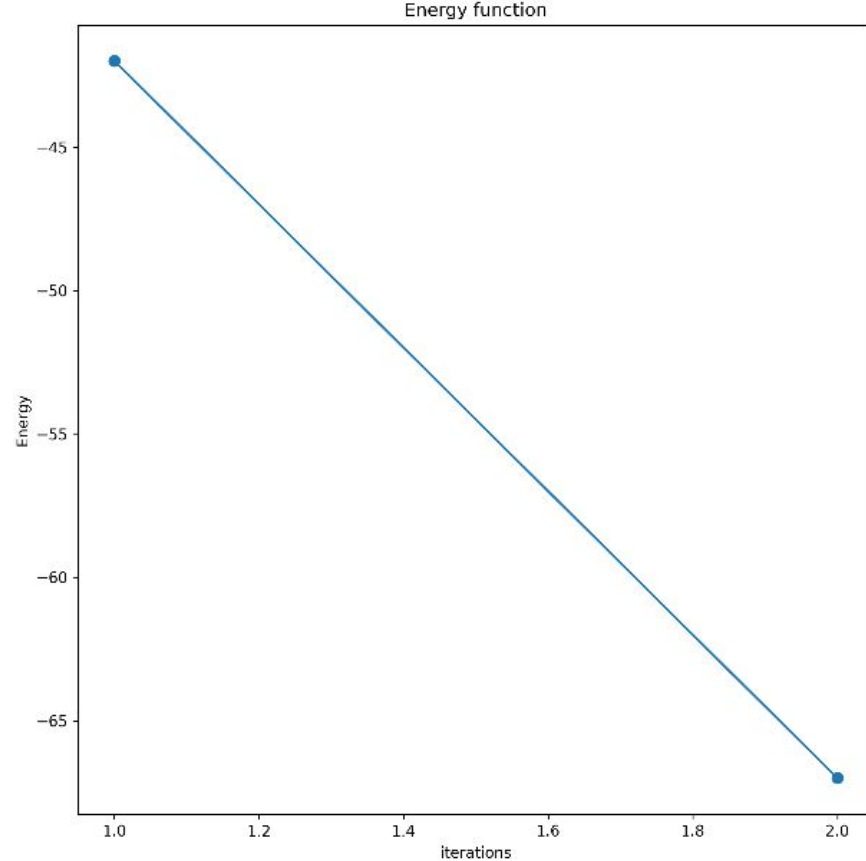
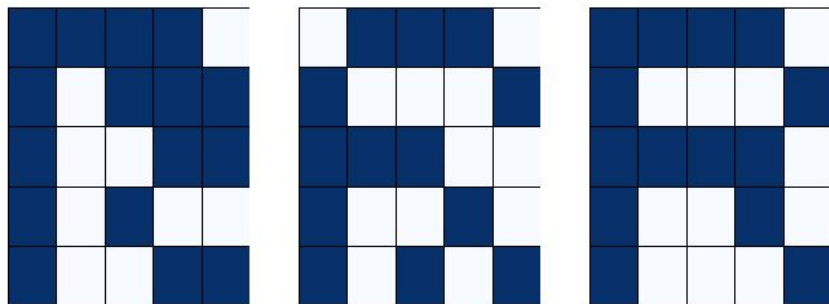
Evolución de la letra I con ruido (8 variaciones) y **Función de Energía**

Letter: O | Noise: 8



Evolución de la letra O con ruido (8 variaciones) y **Función de Energía.**

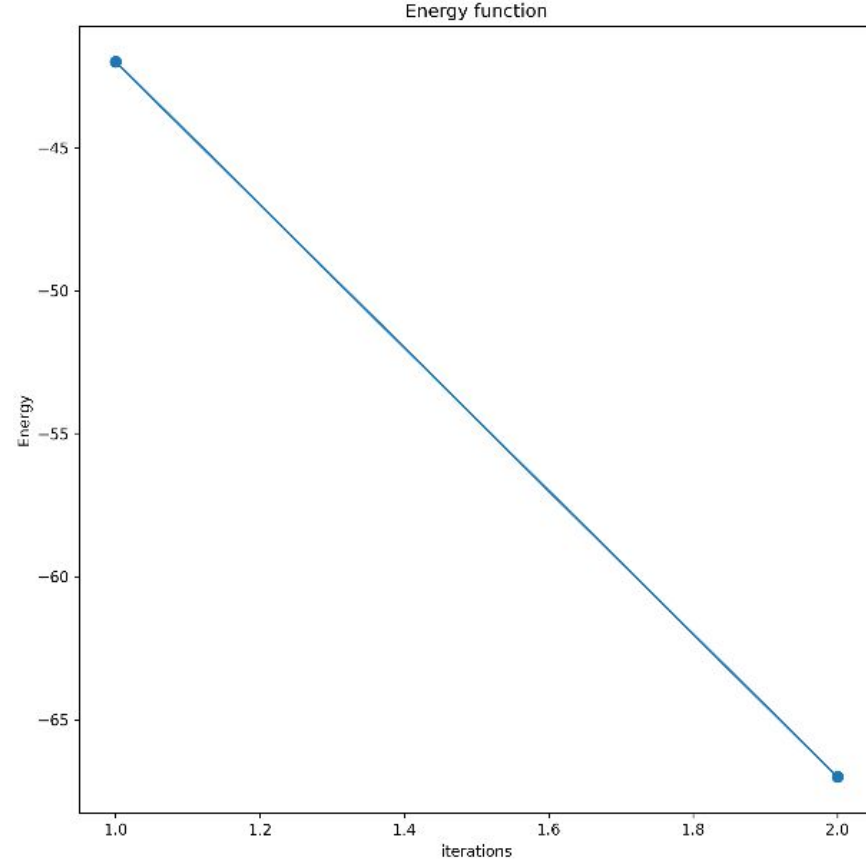
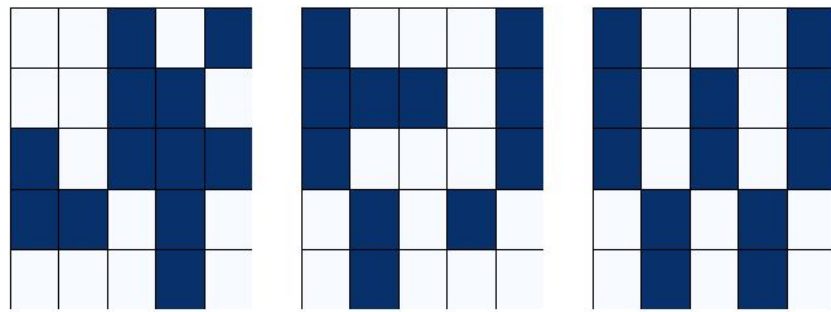
Letter: R | Noise: 8



Evolución de la letra R con ruido (8 variaciones) y **Función de Energía.**

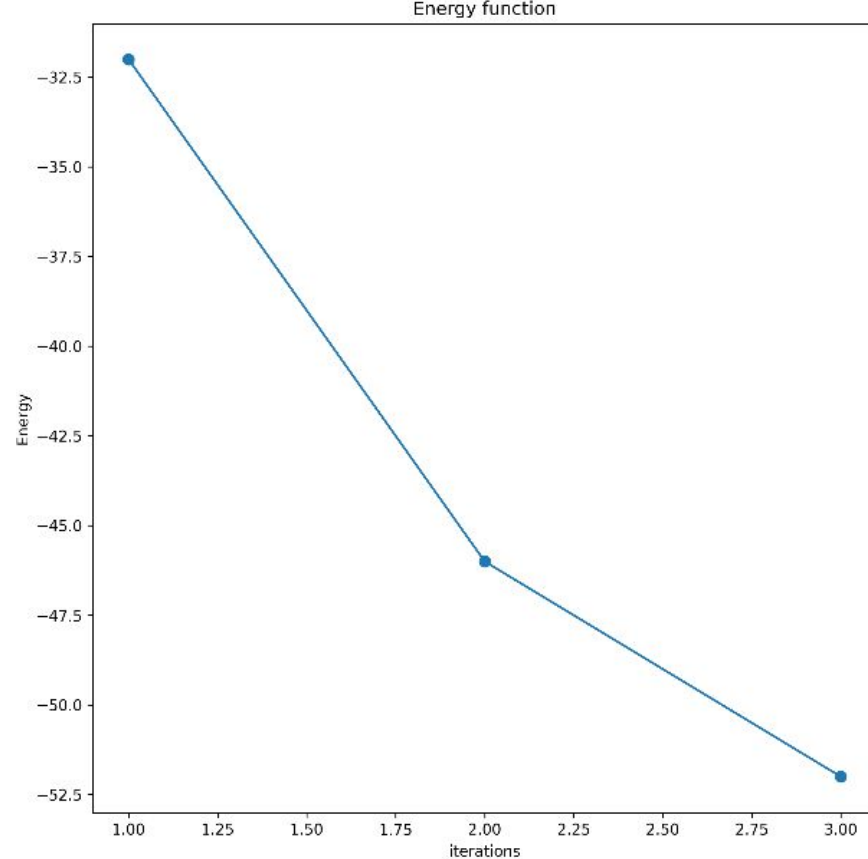
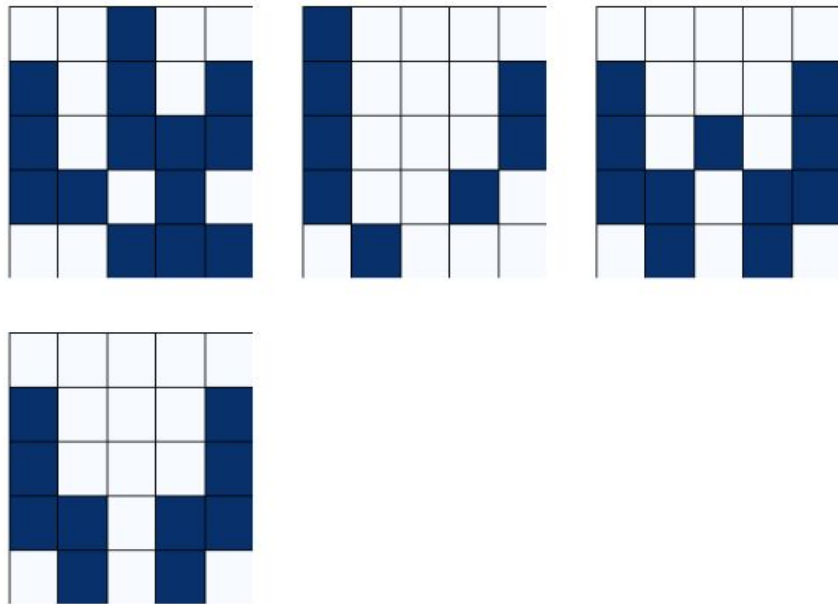


Letter: W | Noise: 8

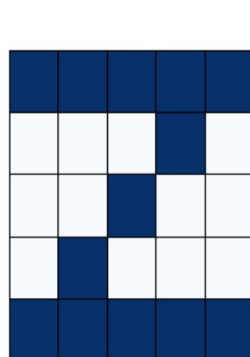


Evolución de la letra W con ruido (8 variaciones) y **Función de Energía.**

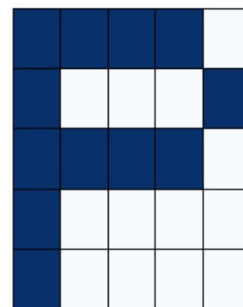
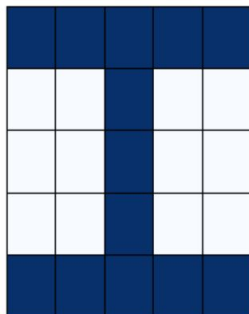
Letter: O | Noise: 10



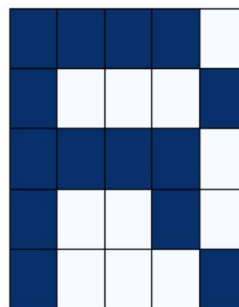
Estado espúreo **partiendo de la letra O con 10 variaciones.**



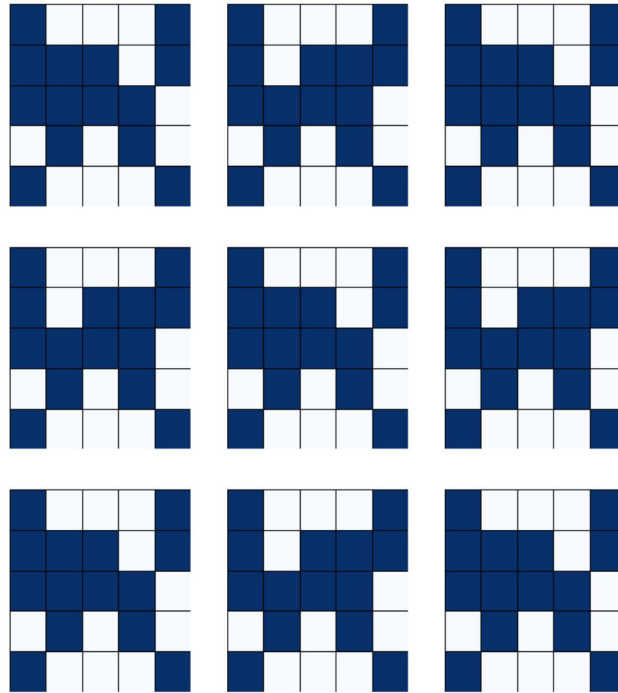
Letter: Z | Noise: 0



Letter: P | Noise: 0



Letras fuera del conjunto de 4 patrones de entrenamiento.

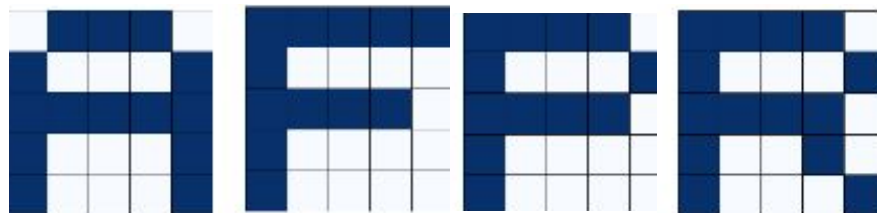


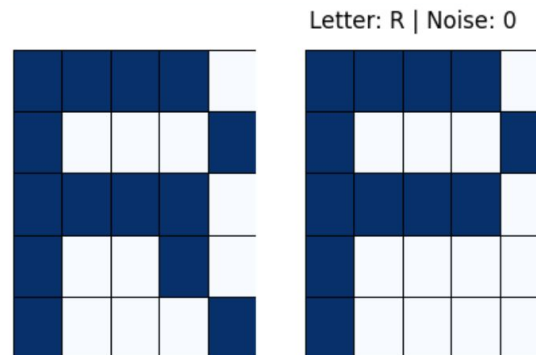
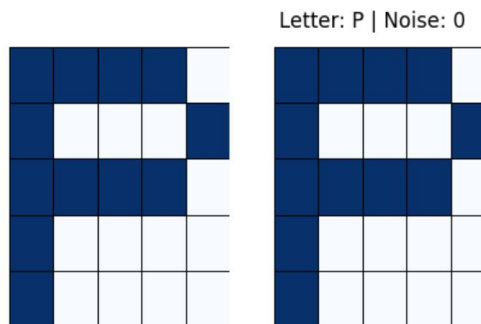
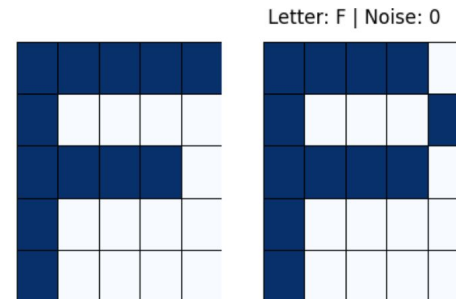
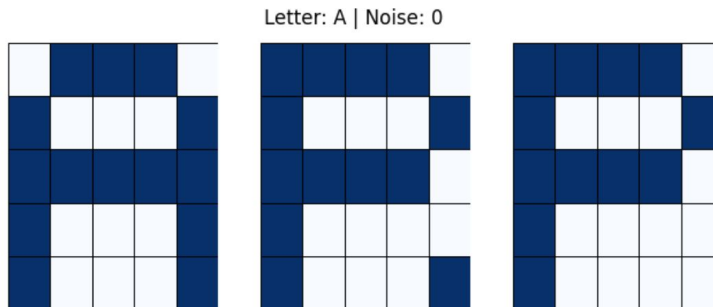
Estado espúreo (últimos 9 estados de 5000)





Entrenamiento con patrones no ortogonales





Resultados de evaluar no ortogonales **sin aplicar ruido**



Conclusiones

- Si los patrones son más ortogonales, mejor será la asociación de los patrones.
 - Para los patrones ortogonales se resuelven bien hasta con 8 valores de ruido.
 - Letras poco ortogonales no otorga buenos resultados.
- Si usamos letras fuera del conjunto entrenado, resuelve al patrón más parecido.
- Si el patrón a asociar tiene demasiado ruido es más probable que se llegue a un estado espúreo



¡Gracias!

¿Alguna pregunta ?

- imanfredi@itba.edu.ar
- juarce@itba.edu.ar
- gpecile@itba.edu.ar