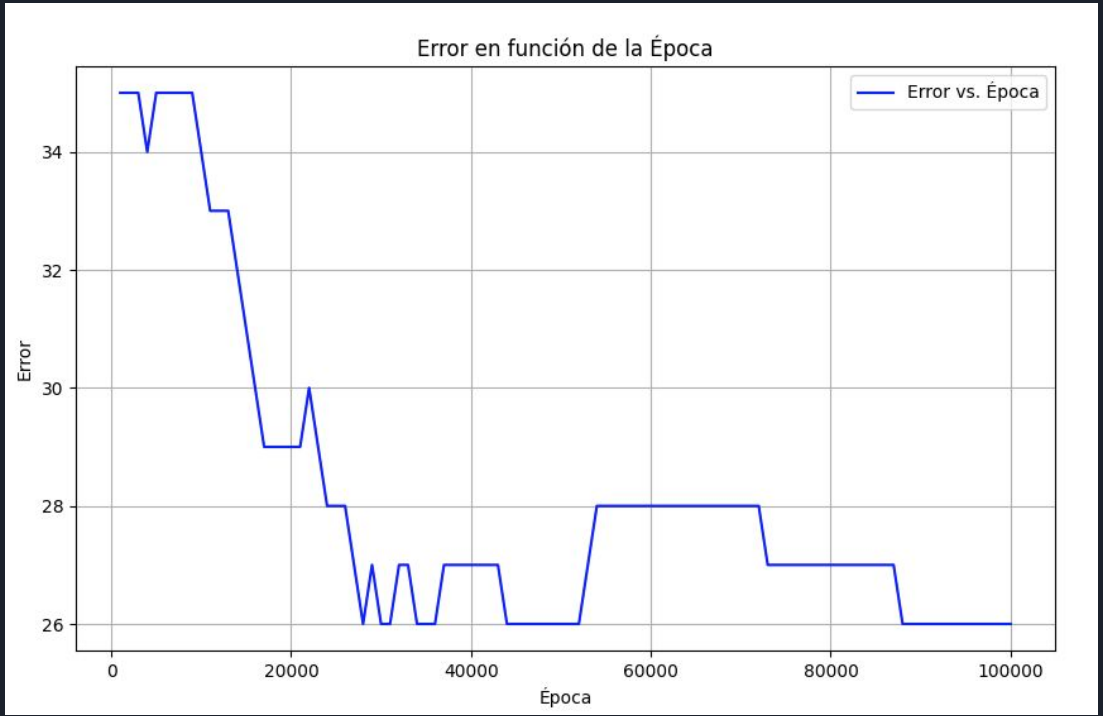
TP5 - Deep learning

Grupo 1
Santiago José Hirsch
Matías Ignacio Luchetti
Santiago Tomás Medin
Mariano Agopian

Ej 1: Autoencoder

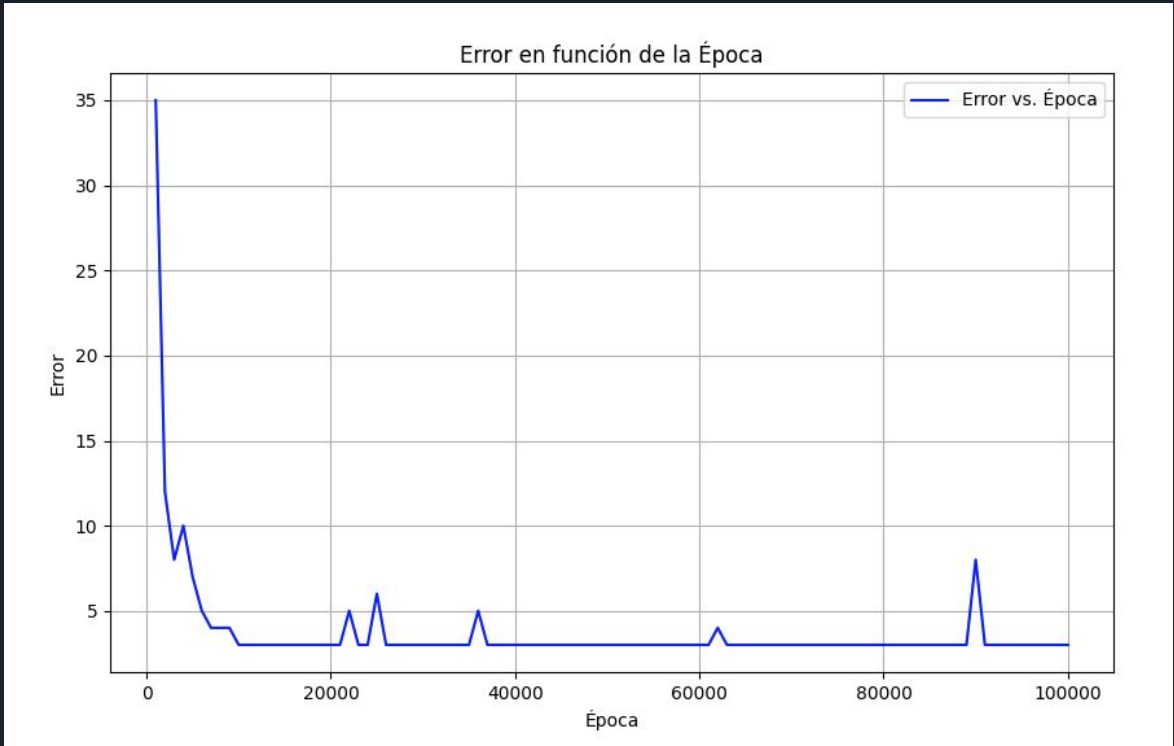
Arrancamos mal...

- Estructura: [35, 10, 2, 10, 35]
- Optimizer: Adam
- Función de activación: TANH
- beta 0,5
- Épocas: 100000
- Learning rate: 0.0001



Pero no tan mal!

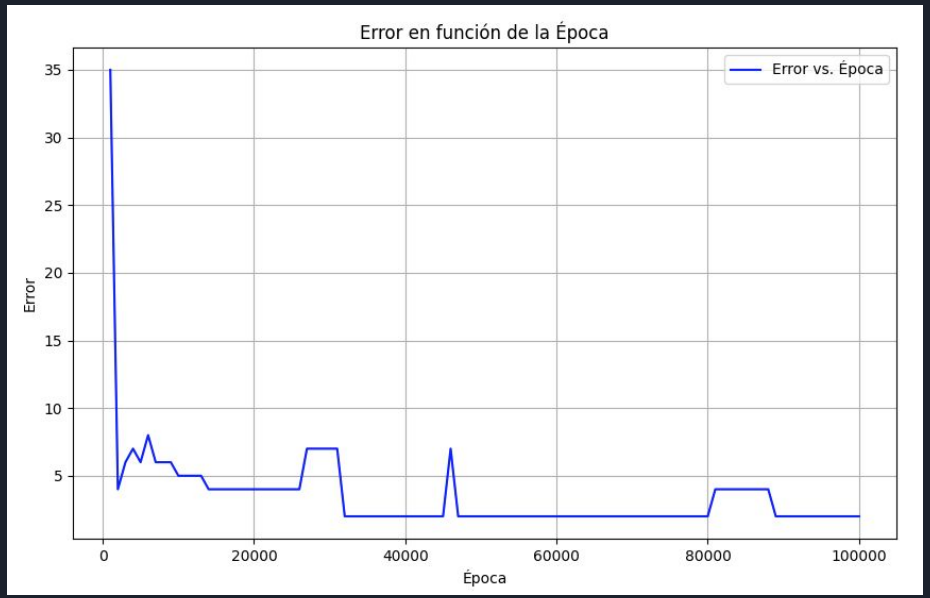
- Estructura: [35, 10, 2, 10, 35]
- Optimizer: Adam
- Función de activación: TANH
- beta 0,5
- Épocas: 100000
- Learning rate: 0.01



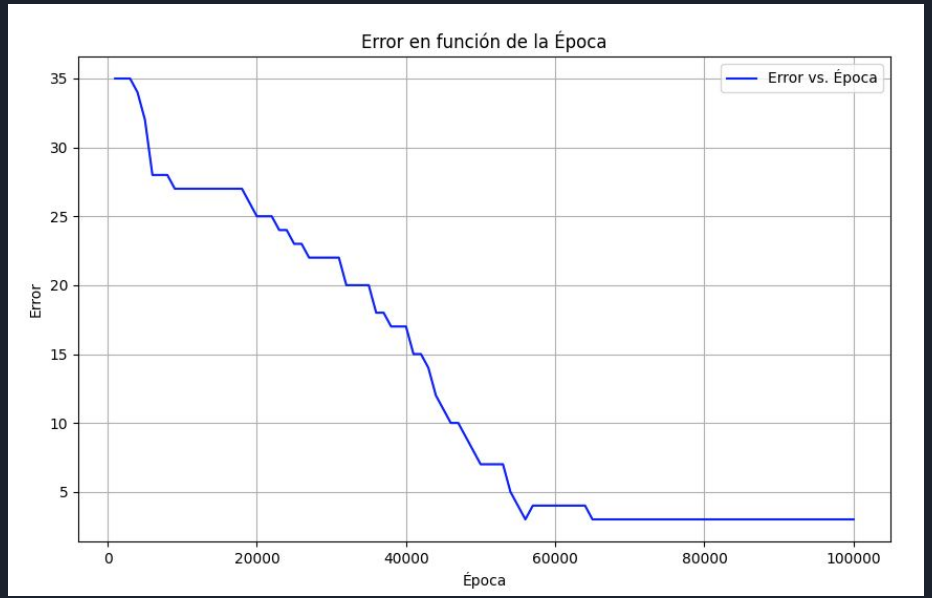
Elección arquitectura

- Estructura: [35, 20, 10, 2, 10, 20, 35]
- Optimizer: Adam
- Función de activación: TANH
- beta 0,5
- Épocas: 100000

Learning rate: 0.01



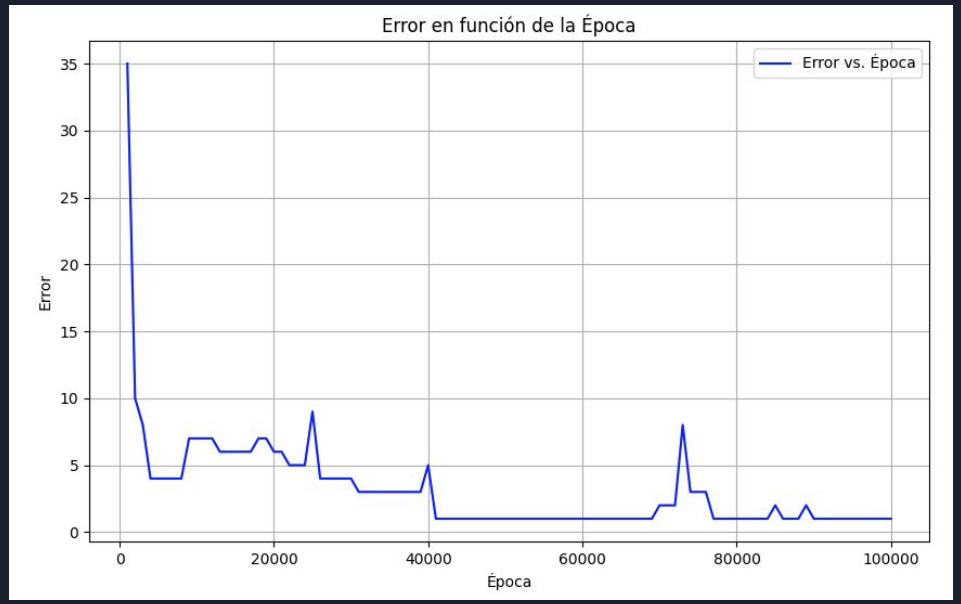
Learning rate: 0.0001



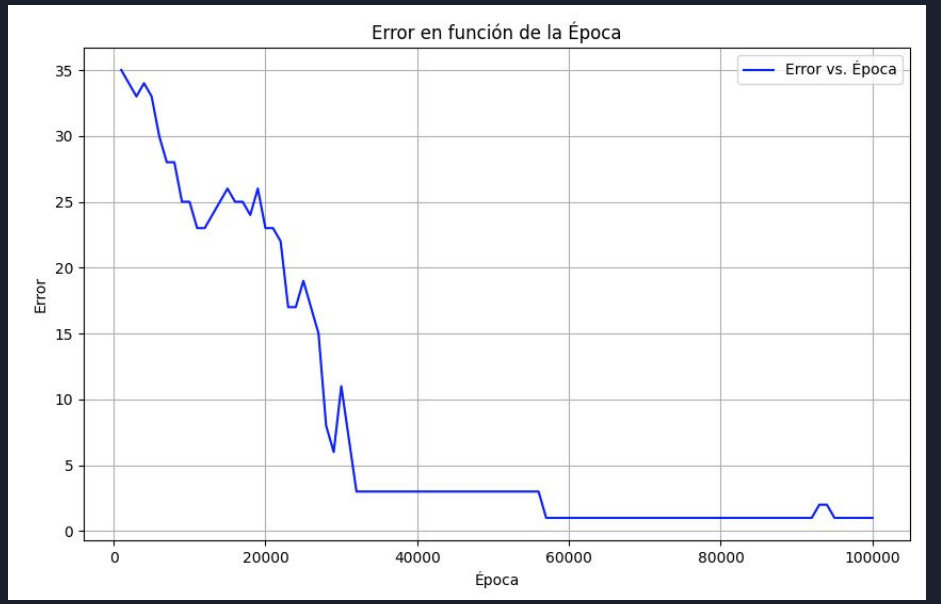
Elección arquitectura

- Estructura: [35, 30, 20, 10, 5, 2, 5, 10, 20, 30, 35]
- Optimizer: Adam
- Función de activación: TANH
- beta 0,5
- Épocas: 100000

Learning rate: 0.01



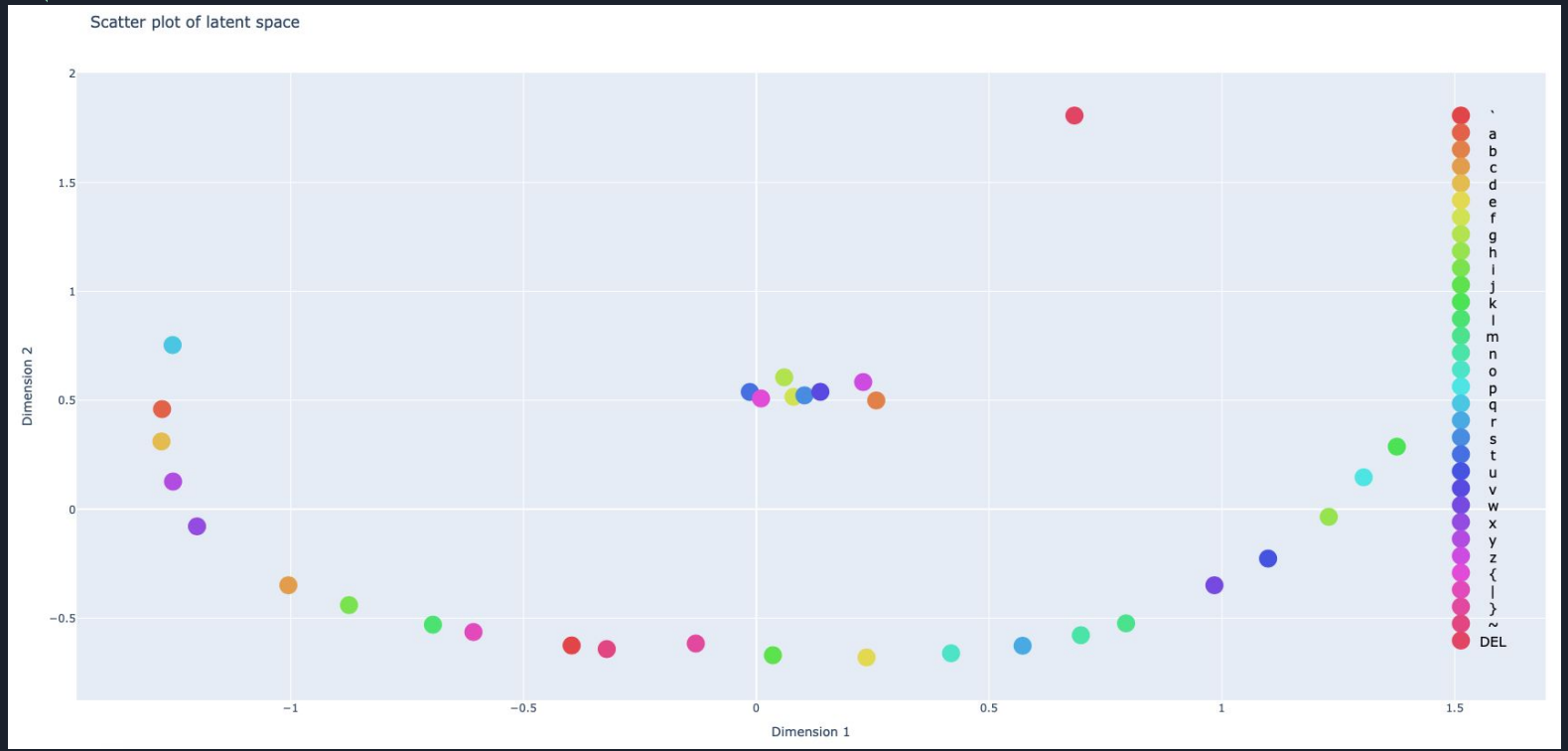
Learning rate: 0.0001



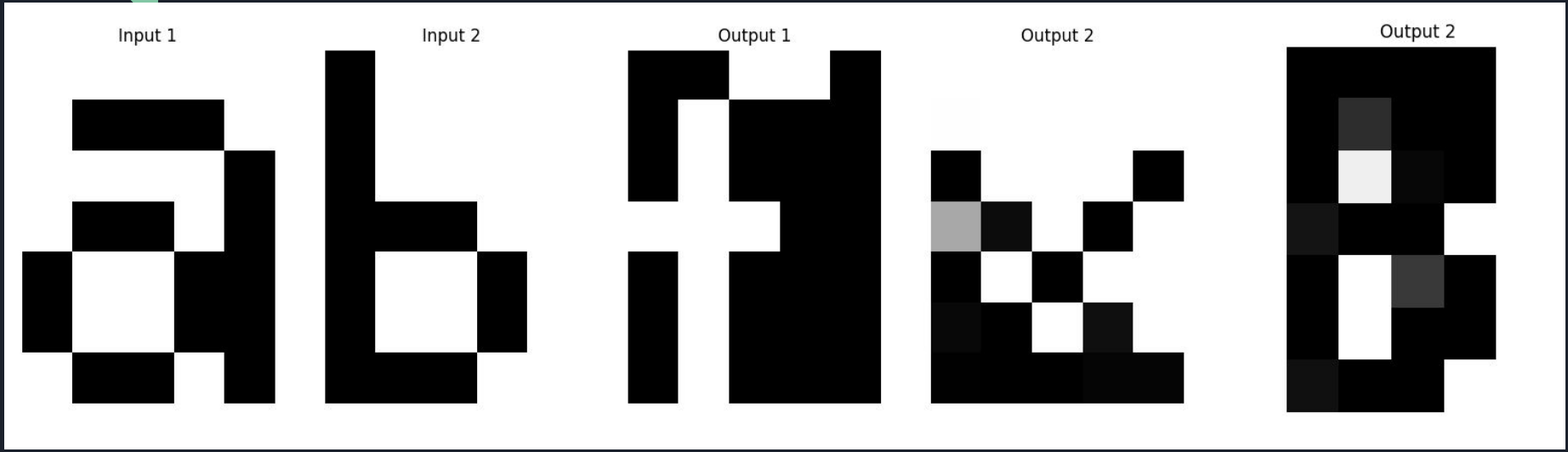
Entrada al espacio latente

Parámetros:

- Learning Rate = 0.0001
- Optimizador = Adam
- Función de Activación = TANH
- Arquitectura = [35, 30, 20, 15, 10, 5, 2, 5, 10, 15, 20, 30, 35]
- Épocas = 100000



Nuevas letras agregando 0,1 en espacio latente

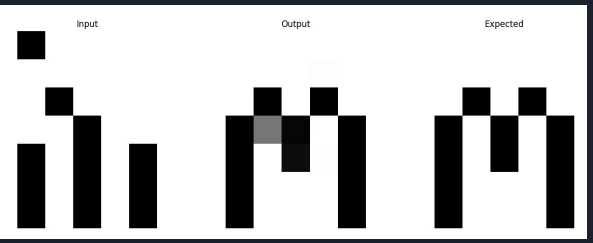


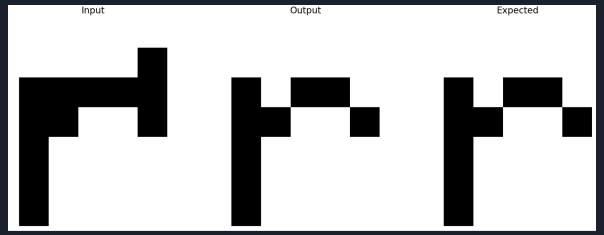
Denoising autoencoder

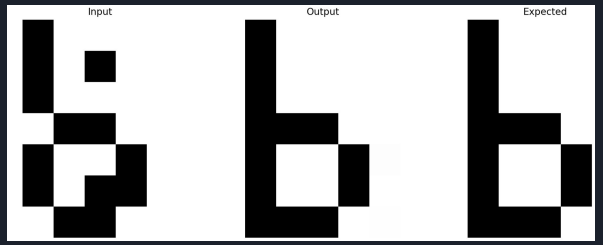
Denoising - Aciertos

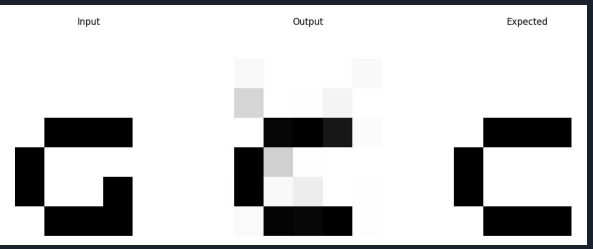
Parámetros:

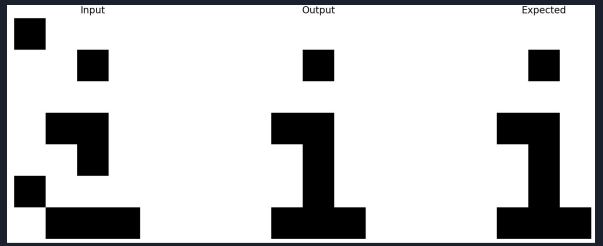
- Learning Rate = 0.0001
- Optimizador = Adam
- Función de Activación = TANH
- Ruido = 0.1
- Épocas = 100000
- Arquitectura = [35, 30, 20, 15, 10, 5, 2, 5, 10, 15, 20, 30, 35]







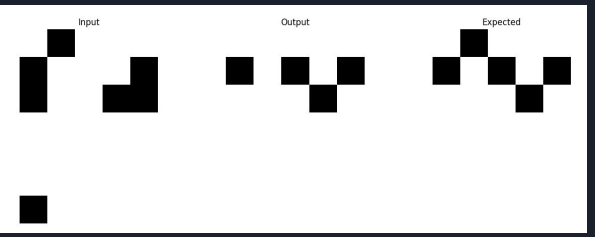


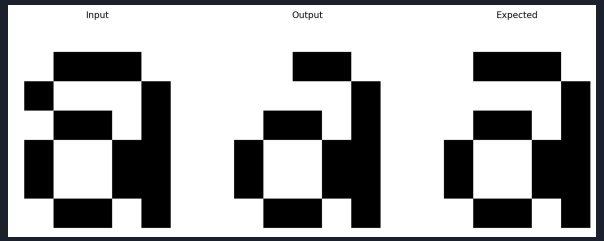


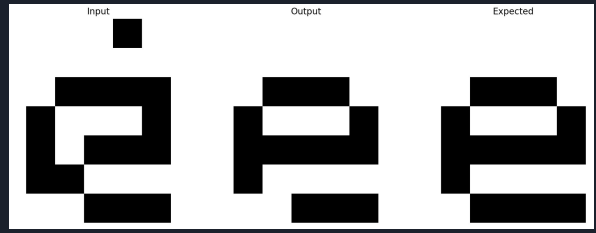
Denoising - Cercanos

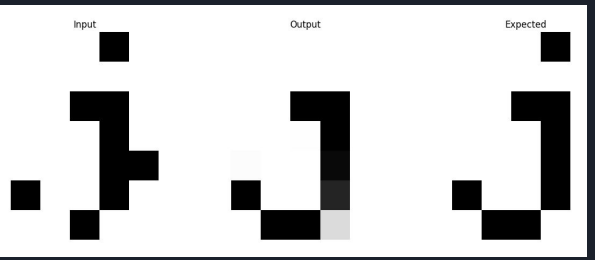
Parámetros:

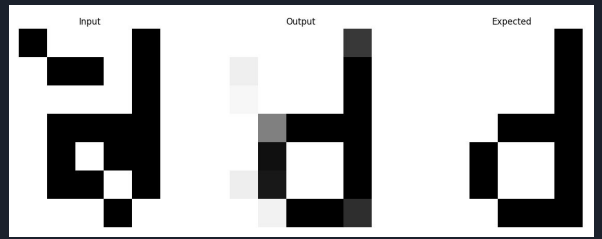
- Learning Rate = 0.0001
- Optimizador = Adam
- Función de Activación = TANH
- Ruido = 0.1
- Épocas = 100000
- Arquitectura = [35, 30, 20, 15, 10, 5, 2, 5, 10, 15, 20, 30, 35]

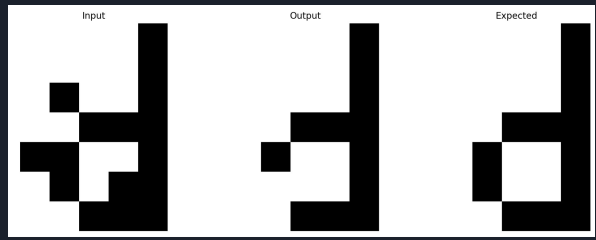








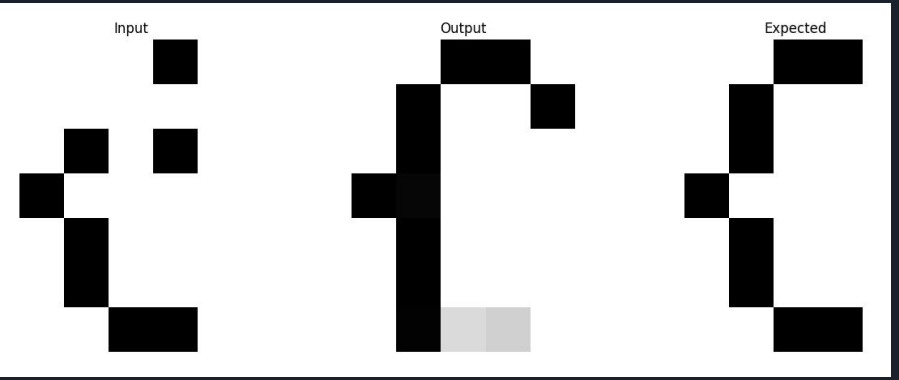


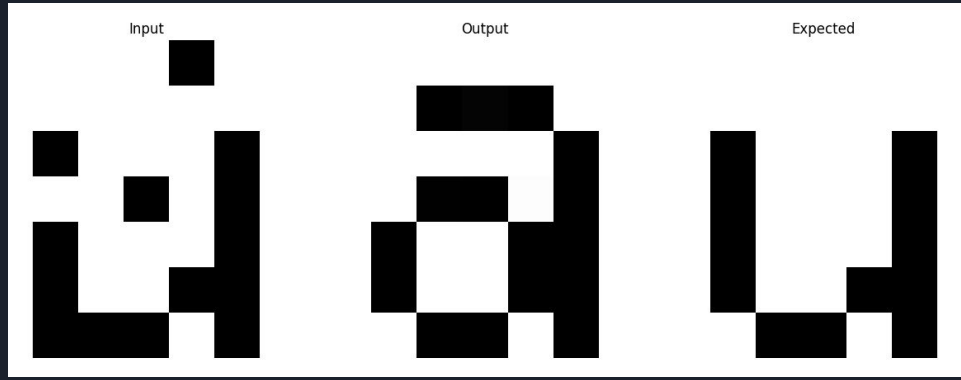


Denoising - Errores

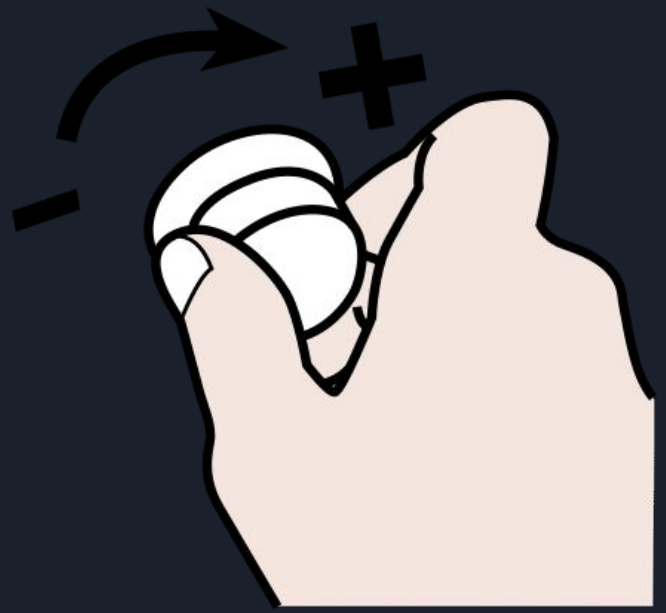
Parámetros:

- Learning Rate = 0.0001
- Optimizador = Adam
- Función de Activación = TANH
- Ruido = 0.1
- Épocas = 100000
- Arquitectura = [35, 30, 20, 15, 10, 5, 2, 5, 10, 15, 20, 30, 35]





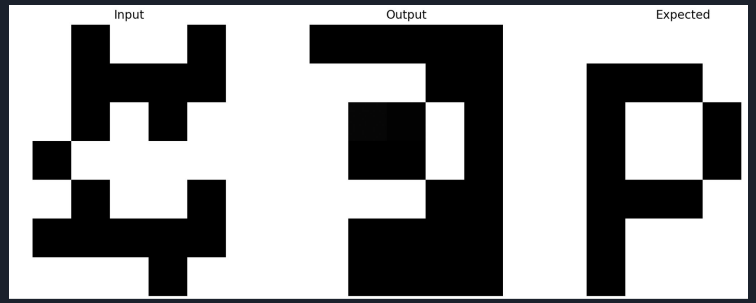
Si subimos un poco el ruido...

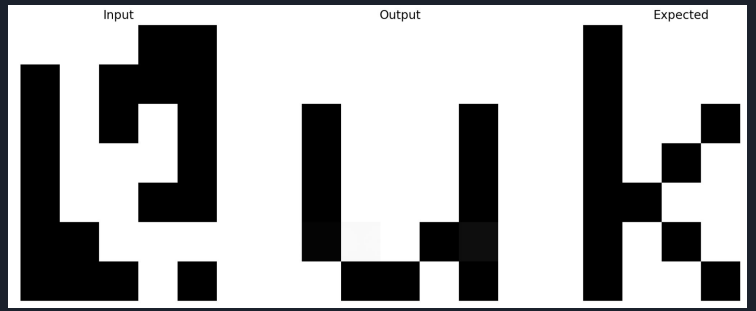


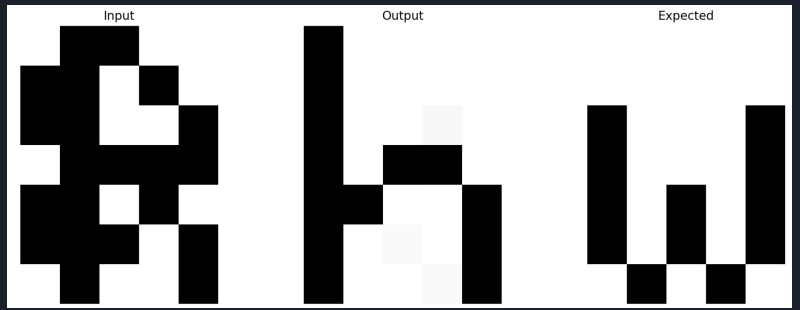
Denoising - Mucho Ruido

Parámetros:

- Learning Rate = 0.0001
- Optimizador = Adam
- Función de Activación = TANH
- Ruido = 0.5
- Épocas = 100000
- Arquitectura = [35, 30, 20, 15, 10, 5, 2, 5, 10, 15, 20, 30, 35]



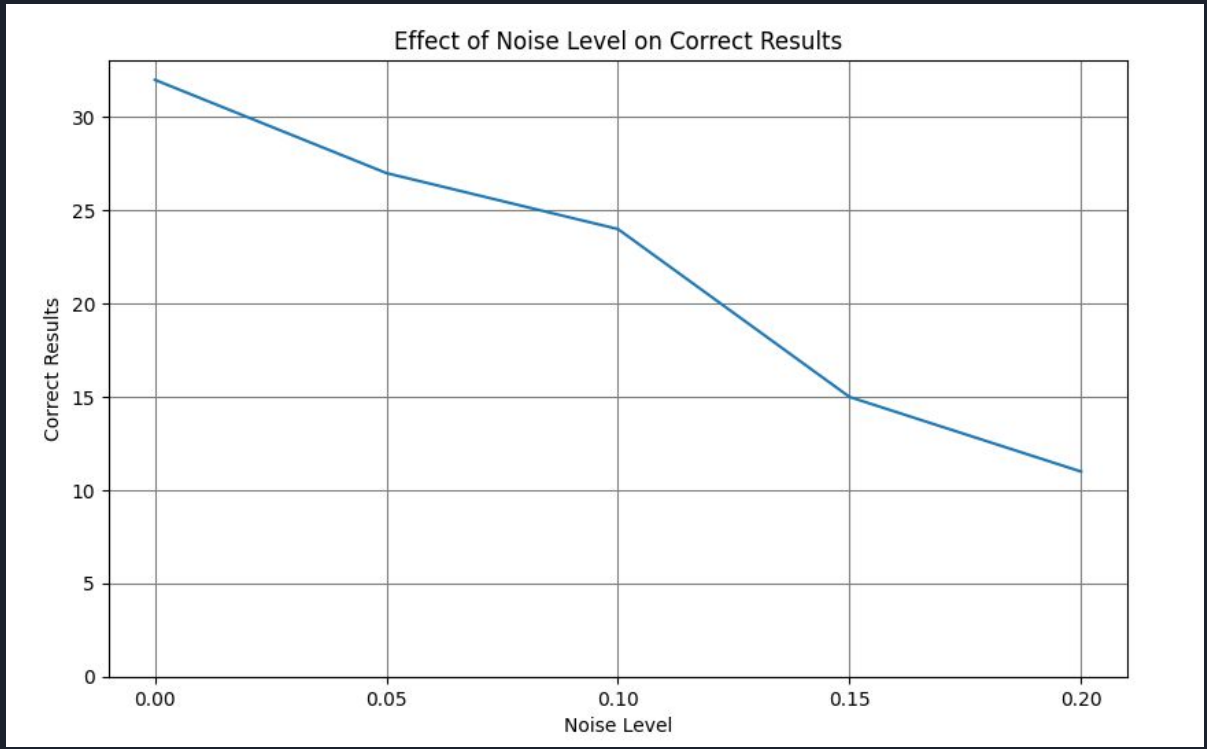




Resultados correctos sobre nivel de ruido

Parámetros:

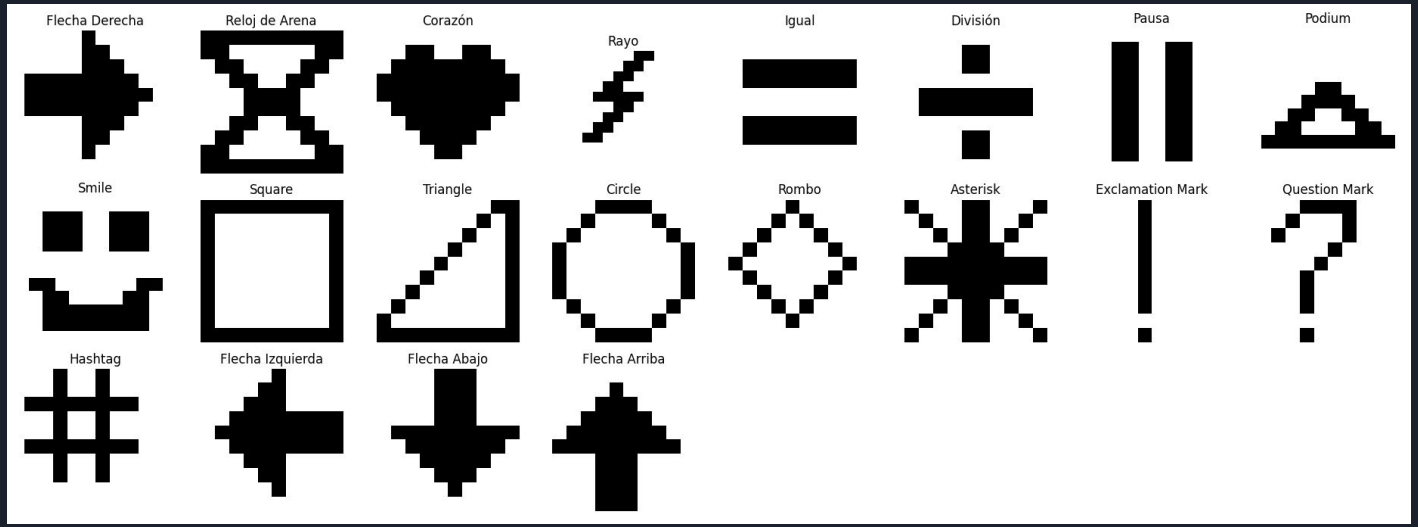
- Learning Rate = 0.001
- Optimizador = Adam
- Función de Activación = TANH
- Ruido = 0.1
- Épocas = 100000
- Arquitectura = [35, 30, 20, 15, 10, 5, 2, 5, 10, 15, 20, 30, 35]



Teniendo en cuenta un resultado correcto como un máximo de dos píxeles erróneos

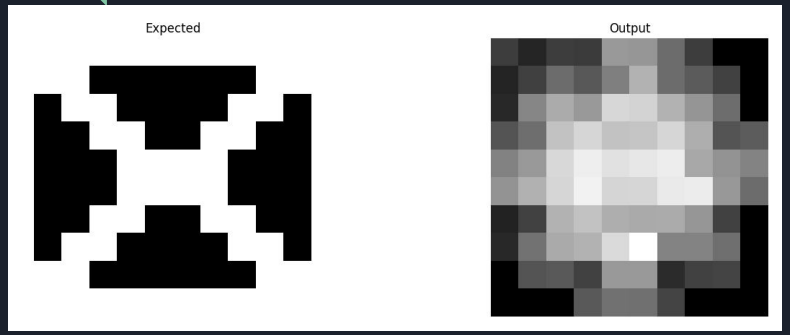
Ejercicio 2 - VAE

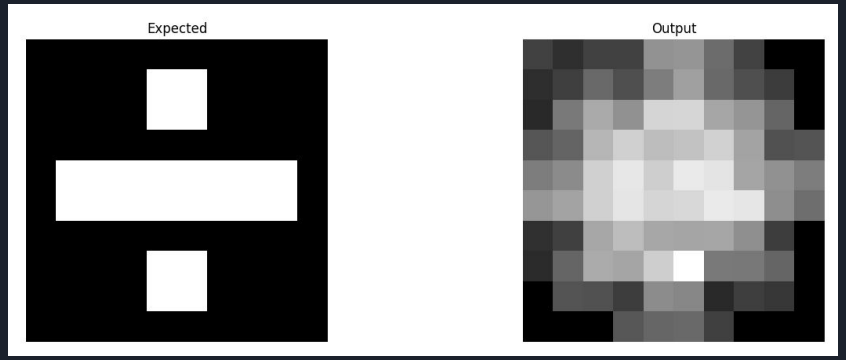
Nueva muestra

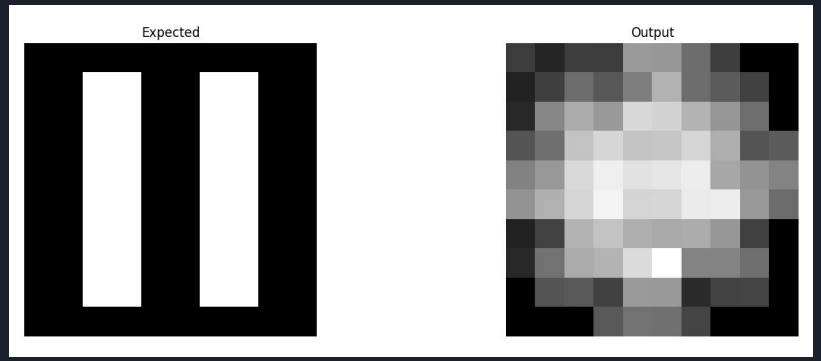


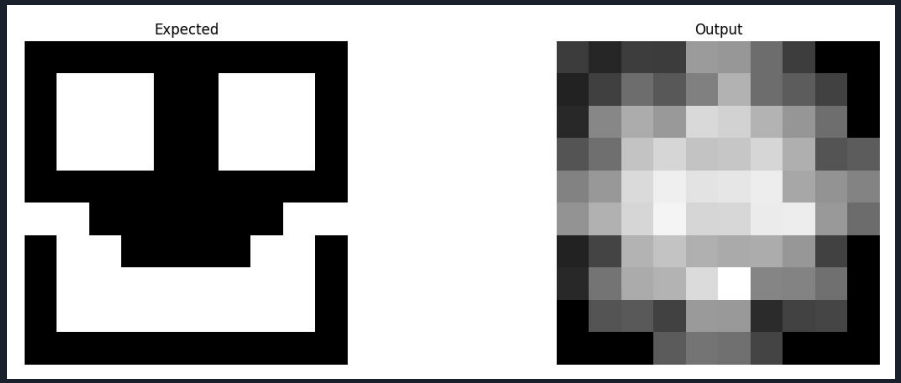
Espacio latente chico

- Arquitectura = [100, 60, 20, 10, 4, 2, 10, 20, 60, 100]
- Learning Rate = 0.001
- Optimizador = Adam
- Función de Activación = TANH
- Épocas = 10000



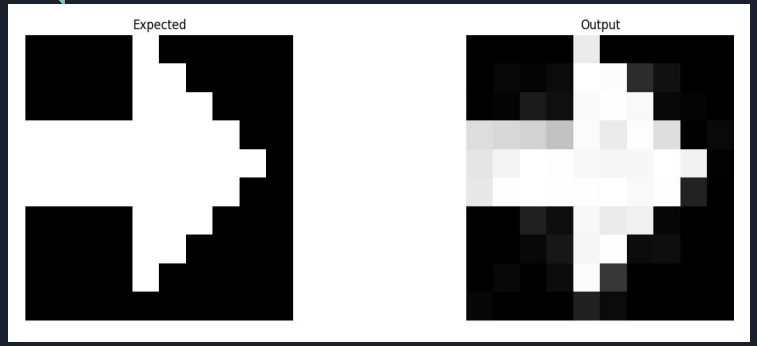


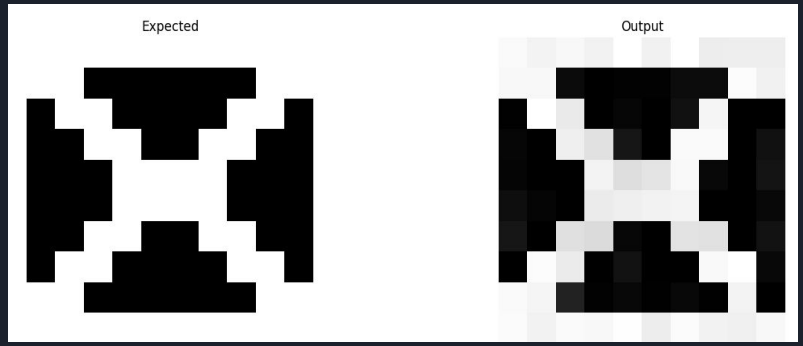


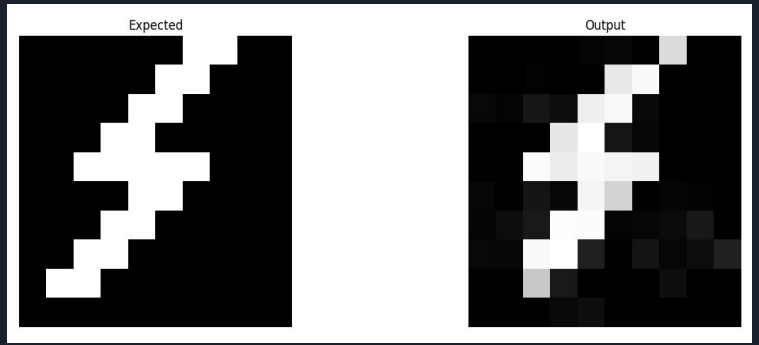


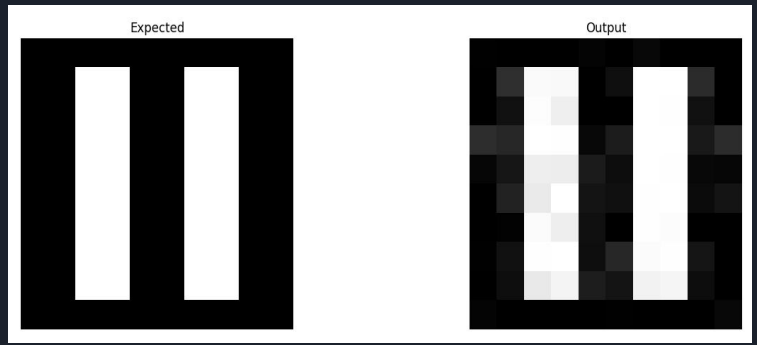
Espacio latente más amplio

- Arquitectura = [100,80,60,40,20,10,20,40,60,80,100]
- Learning Rate = 0.001
- Optimizador = Adam
- Función de Activación = TANH
- Épocas = 10000

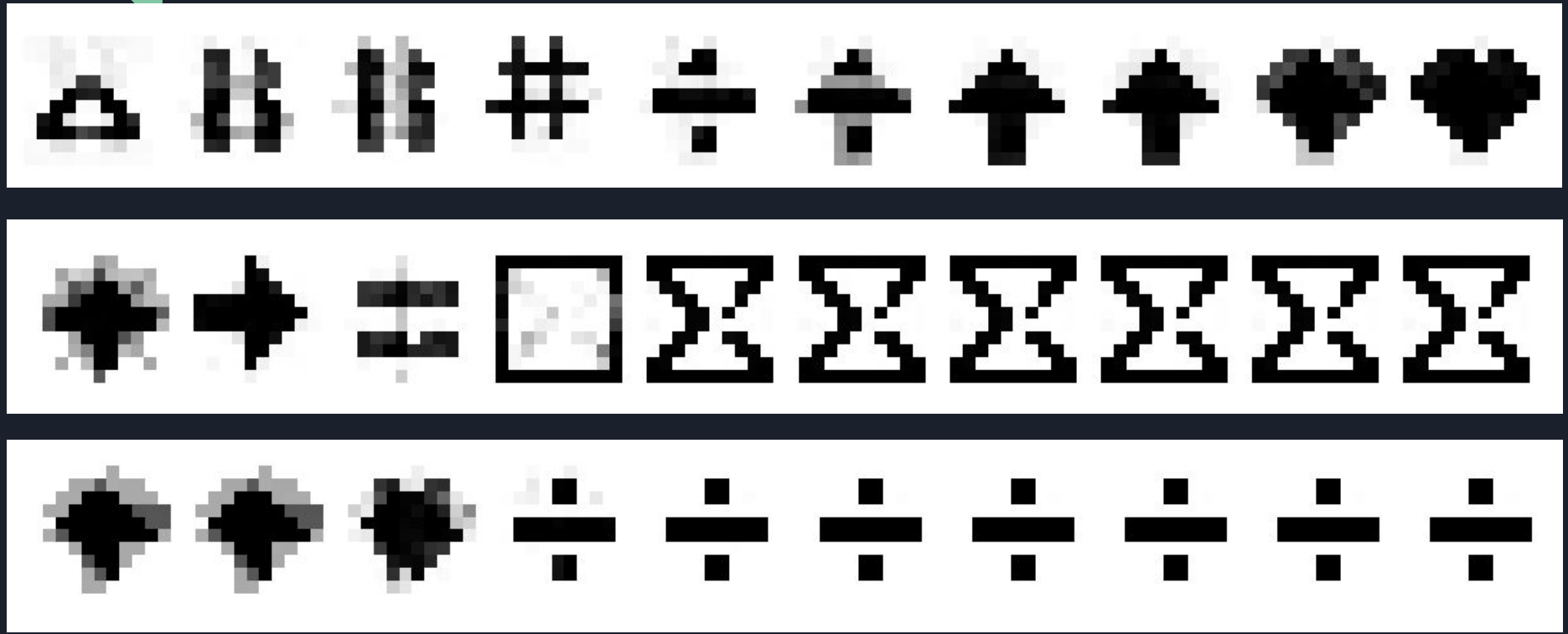








VAE Interpolación



Conclusiones

- Encontramos una estructura que reconozca los caracteres con un error máximo de 1 pixel
- Como venimos viendo en todos los tps, todos los parámetros influyen en los resultados que se obtienen (método Adam y learning rate por ejemplo)
- Vimos la capacidad del autoencoder para crear nuevos caracteres que no forman parte del data set de entrenamiento
- Vimos como el autoencoder puede eliminar el ruido de una imagen, devolviendo la imagen “original”, con buenos rendimientos hasta un ruido de 0.1 (método Salt and Pepper)
- El método Salt and Pepper destruye por completo las imágenes si se usa un ruido de 0.5 (probabilísticamente, debería cambiar casi la mitad de la imagen)
- Es más difícil de lo que pensamos dibujar con pixeles en una matriz de 10x10!!
- Si el espacio latente es muy chico en el VAE, puede significar una incapacidad del autoencoder