

# Procesamiento de Lenguaje Natural

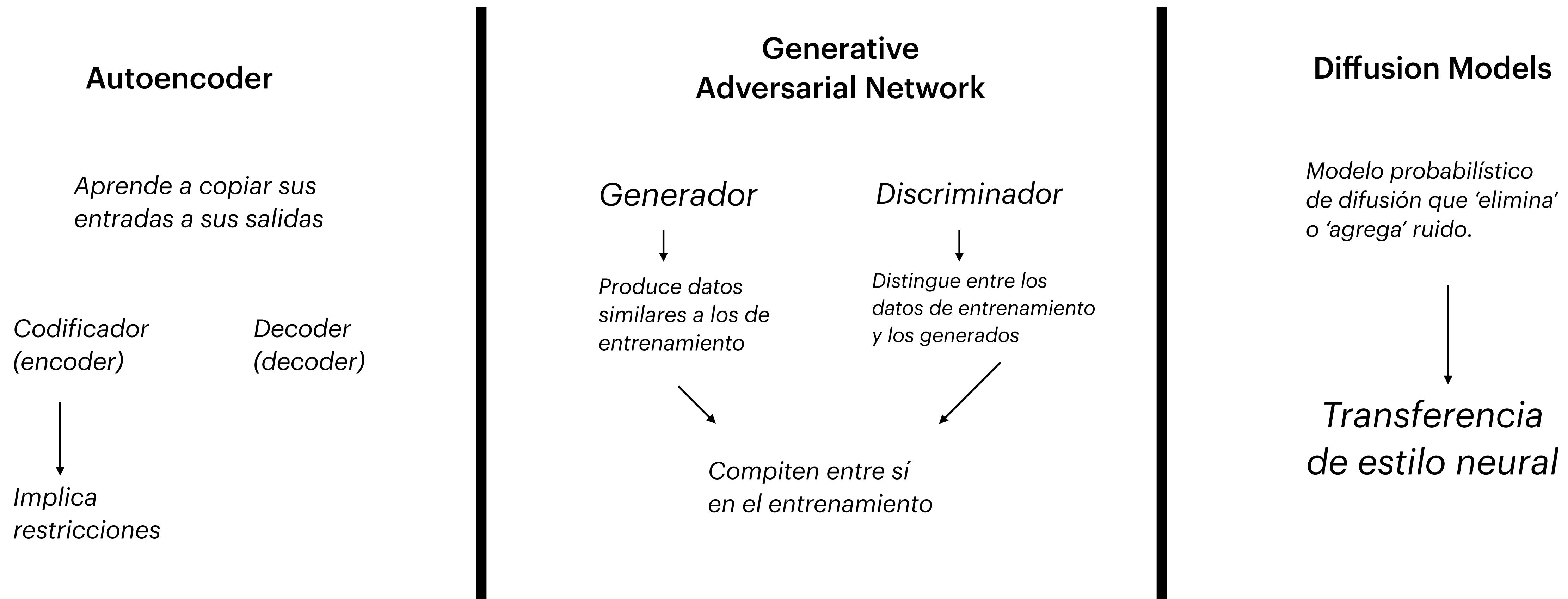
Tópicos Avanzados en Analítica  
**Maestría en Analítica para la Inteligencia de Negocios**

Sergio Alberto Mora Pardo - H2 2023

# Encoder-Decoder

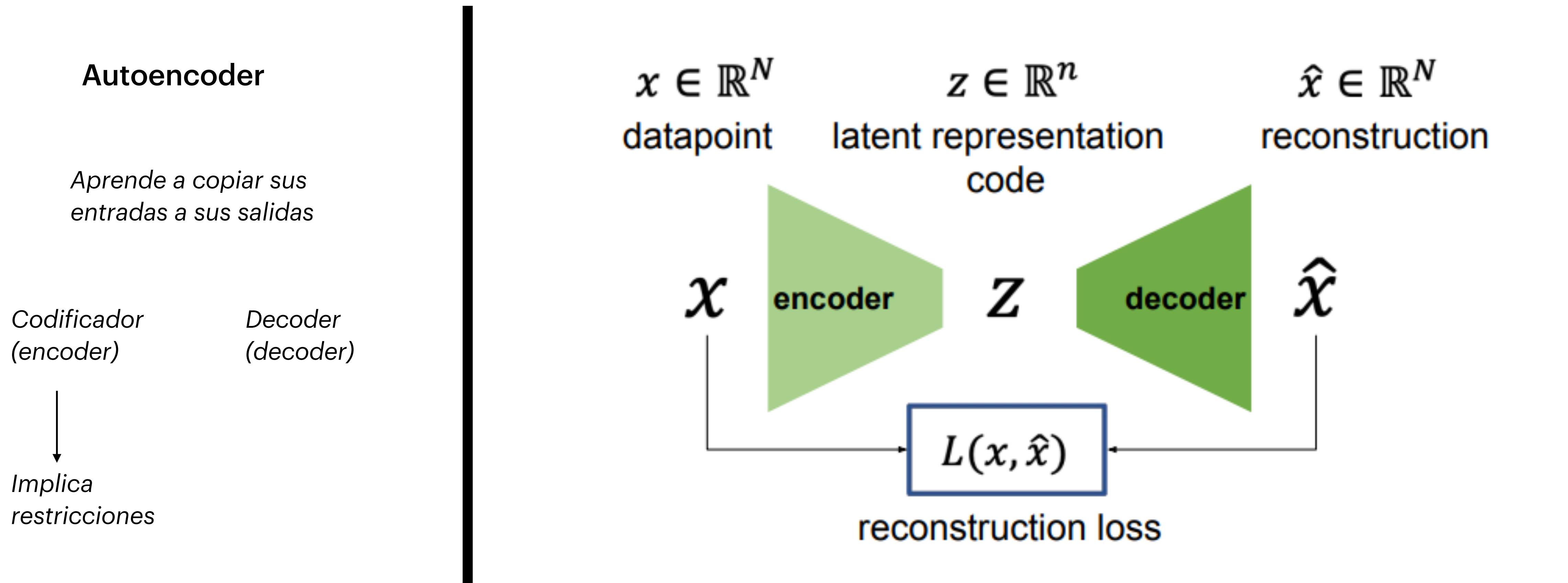
# Encoder-Decoder

## Autoencoders, GANs, Diffusion Models



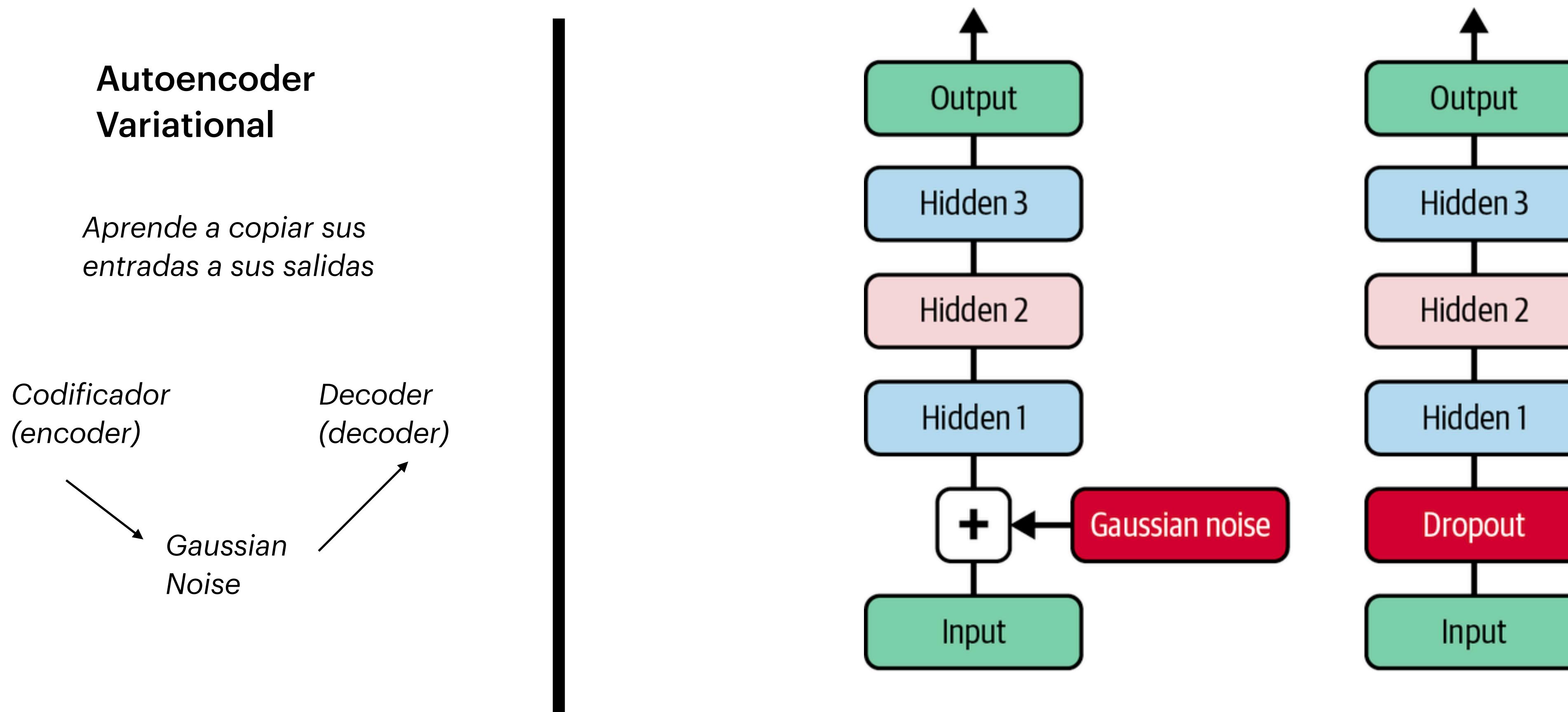
# Encoder-Decoder

## Autoencoders, GANs, Diffusion Models



# Encoder-Decoder

## Autoencoders, GANs, Diffusion Models



# Encoder-Decoder

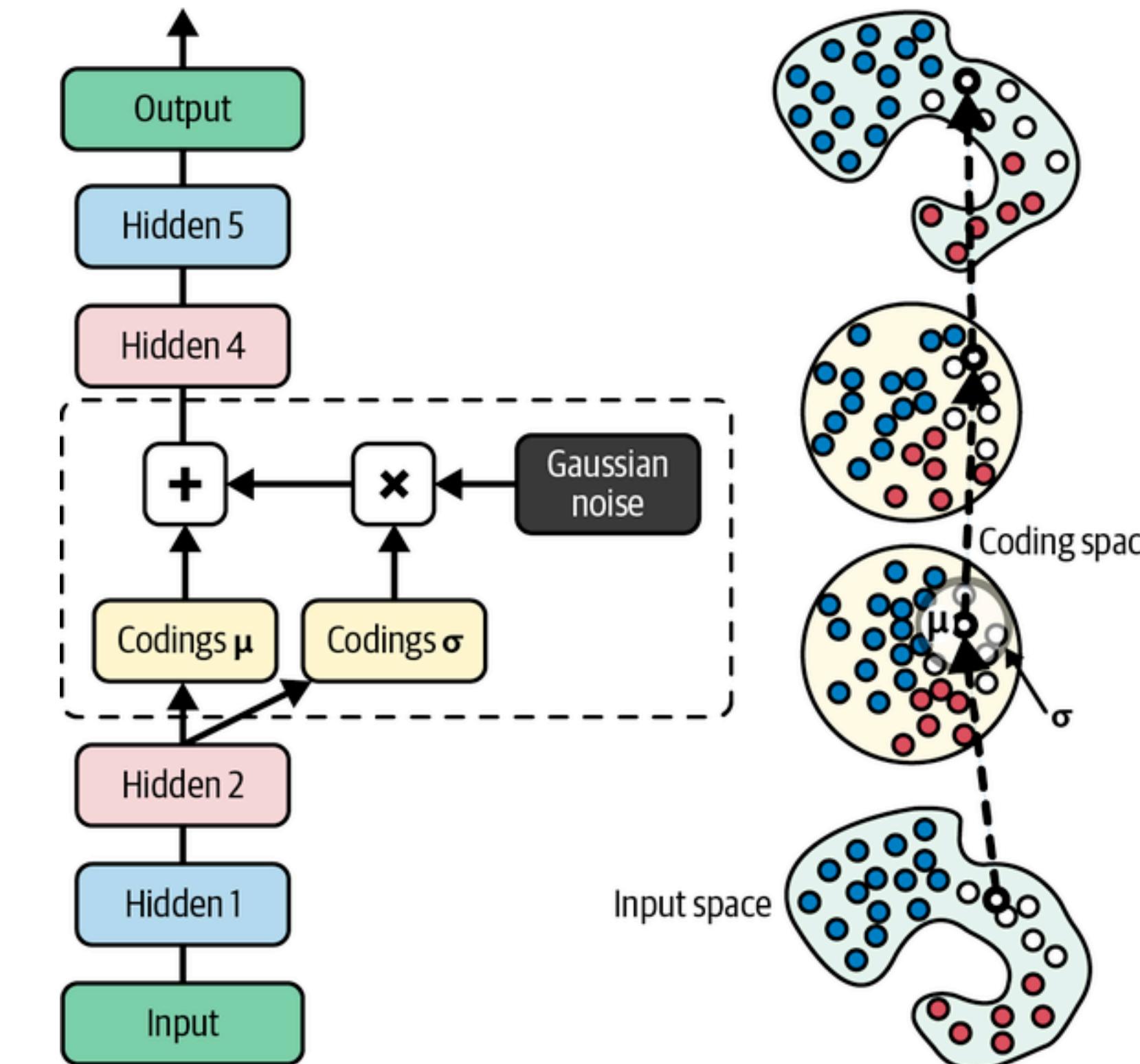
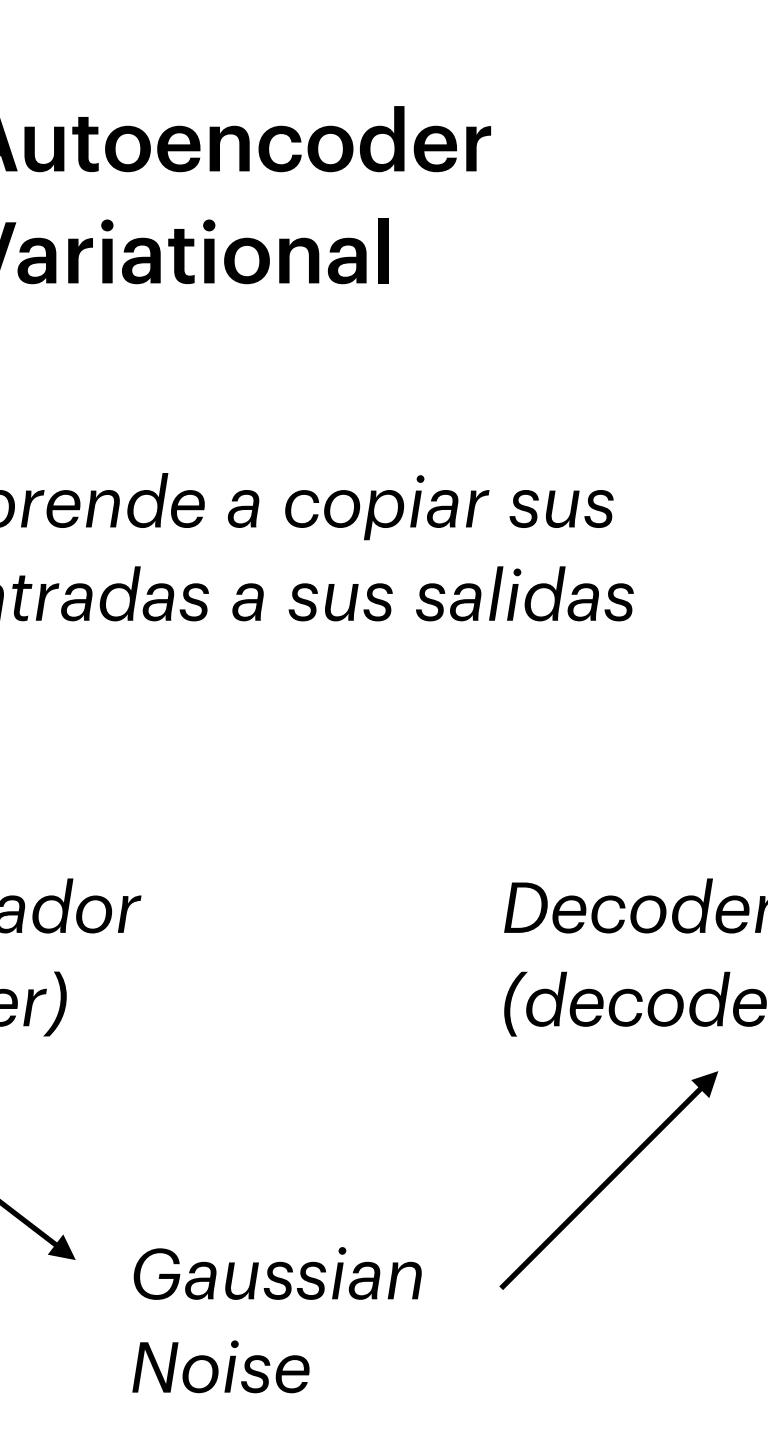
## Autoencoders, GANs, Diffusion Models

**Autoencoder Variational**

Aprende a copiar sus entradas a sus salidas

Codificador (encoder)      Decoder (decoder)

Gaussian Noise



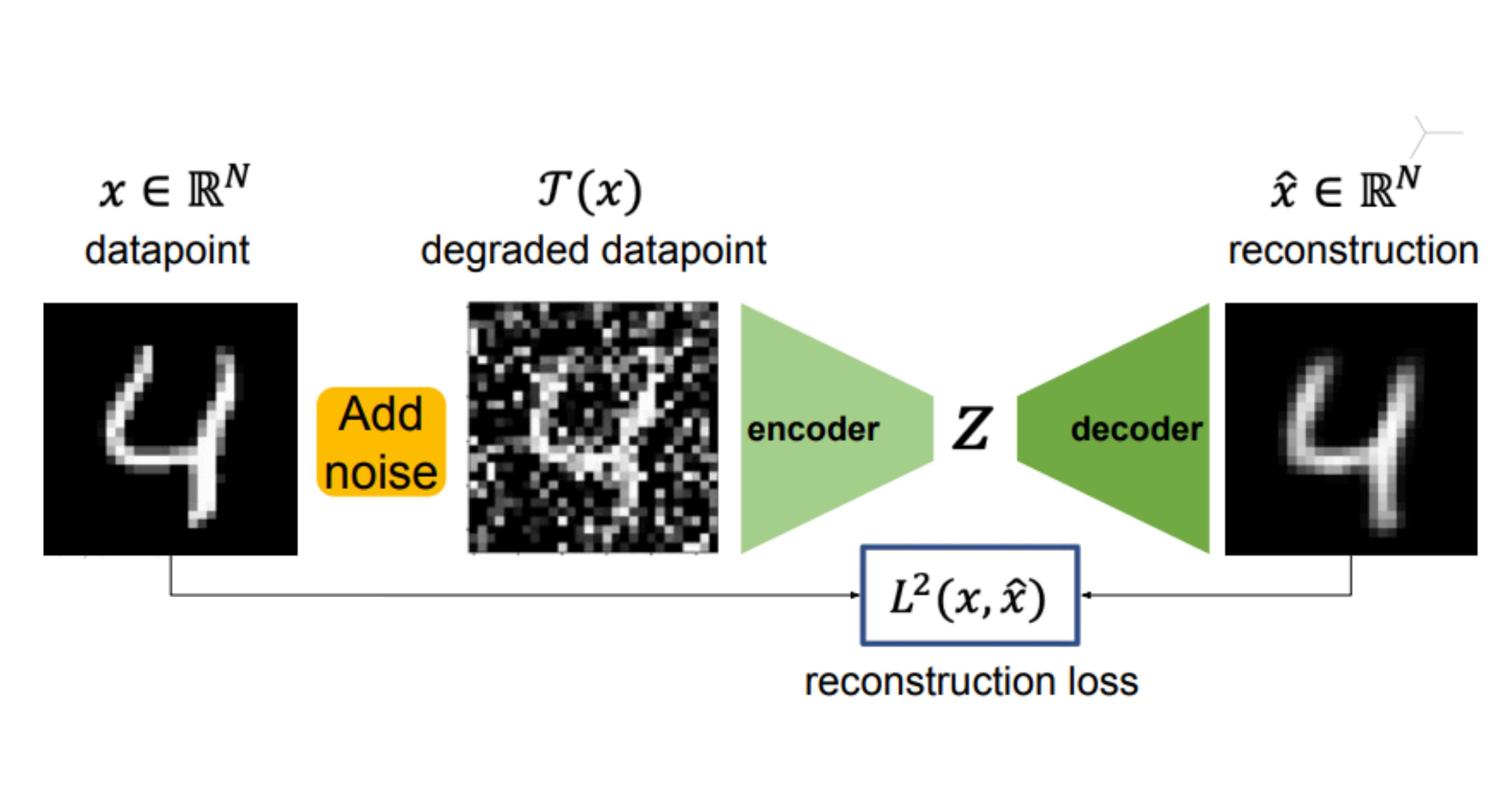
# Encoder-Decoder

## Autoencoders, GANs, Diffusion Models

### Diffusion Models

Modelo probabilístico de difusión que ‘elimina’ o ‘agrega’ ruido.

Transferencia de estilo neural



# Encoder-Decoder

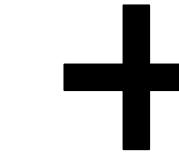
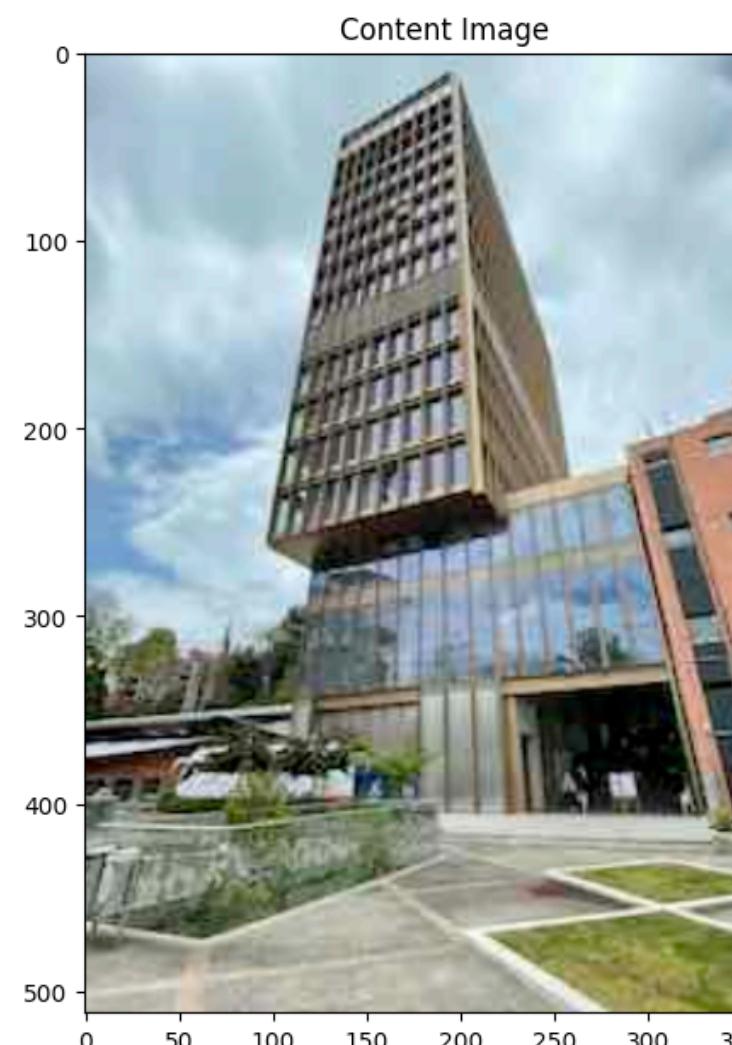
## Autoencoders, GANs, Diffusion Models

### Diffusion Models

*Modelo probabilístico de difusión que 'elimina' o 'agrega' ruido.*



*Transferencia de estilo neural*



# Encoder-Decoder

## Autoencoders, GANs, Diffusion Models

### Diffusion Models

*Modelo probabilístico de difusión que 'elimina' o 'agrega' ruido.*



*Transferencia de estilo neural*



# Encoder-Decoder

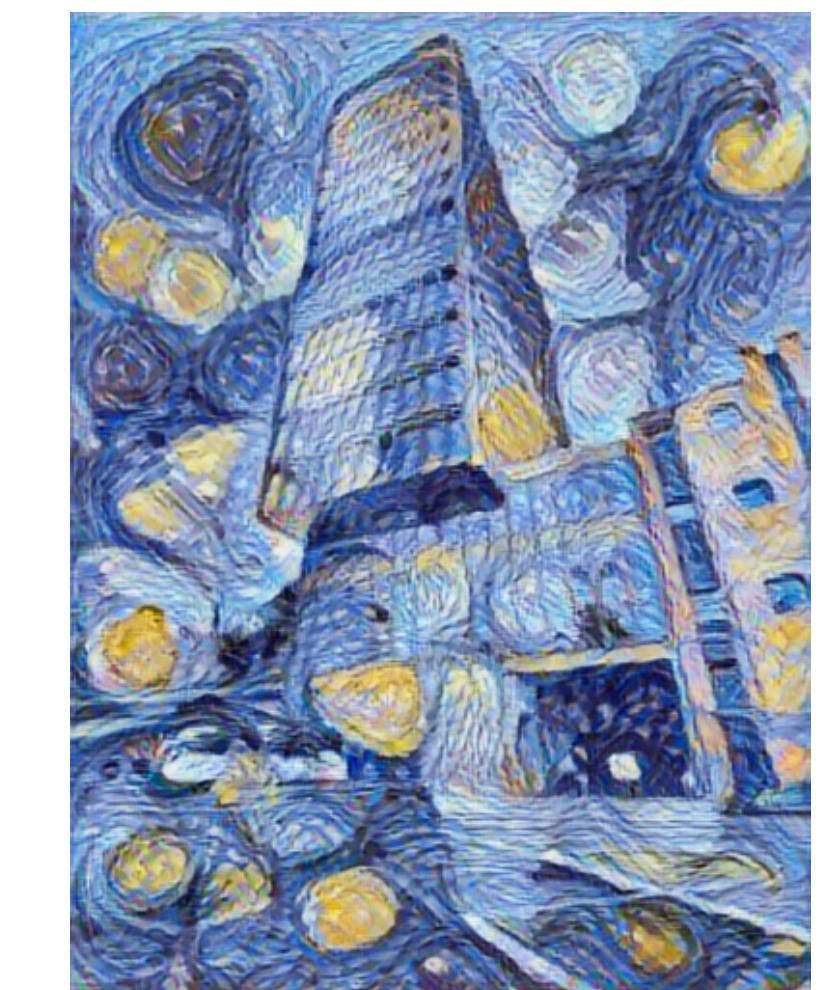
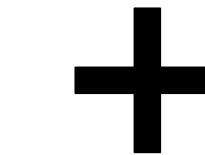
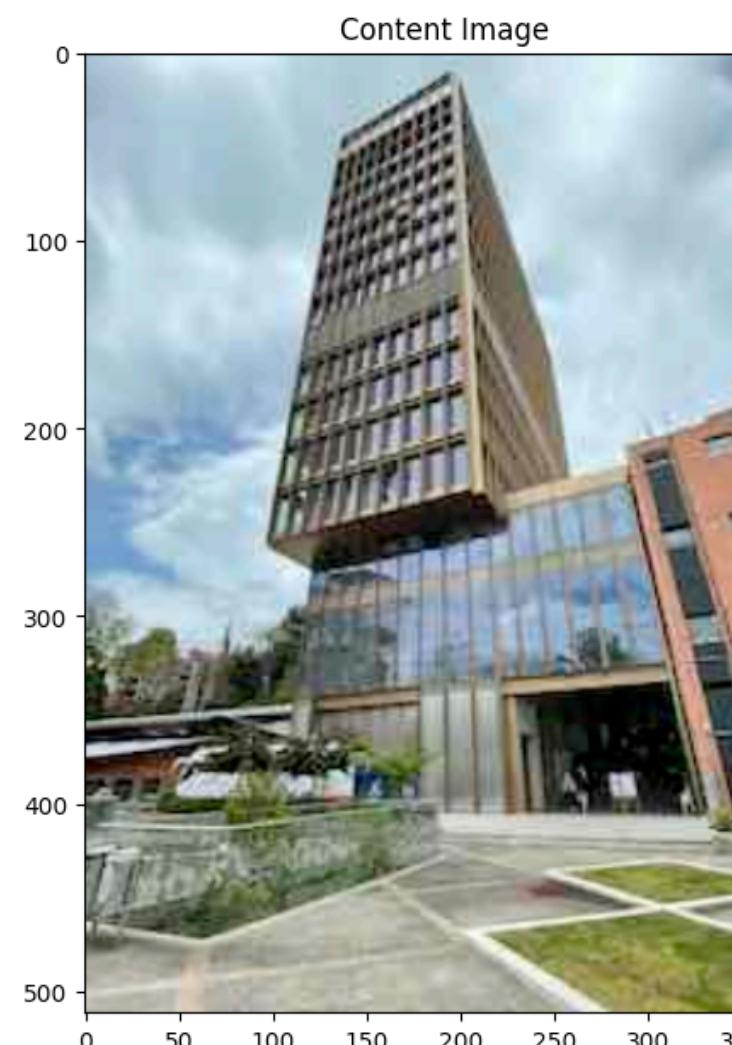
## Autoencoders, GANs, Diffusion Models

### Diffusion Models

*Modelo probabilístico  
de difusión que 'elimina'  
o 'agrega' ruido.*

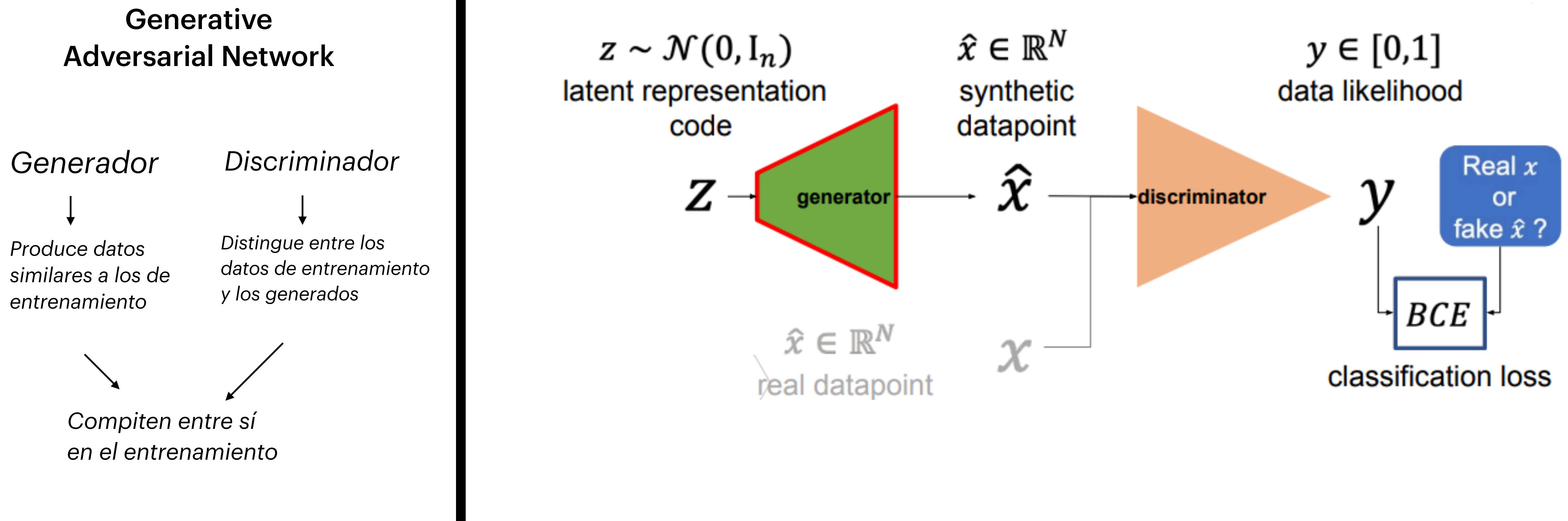


*Transferencia  
de estilo neural*



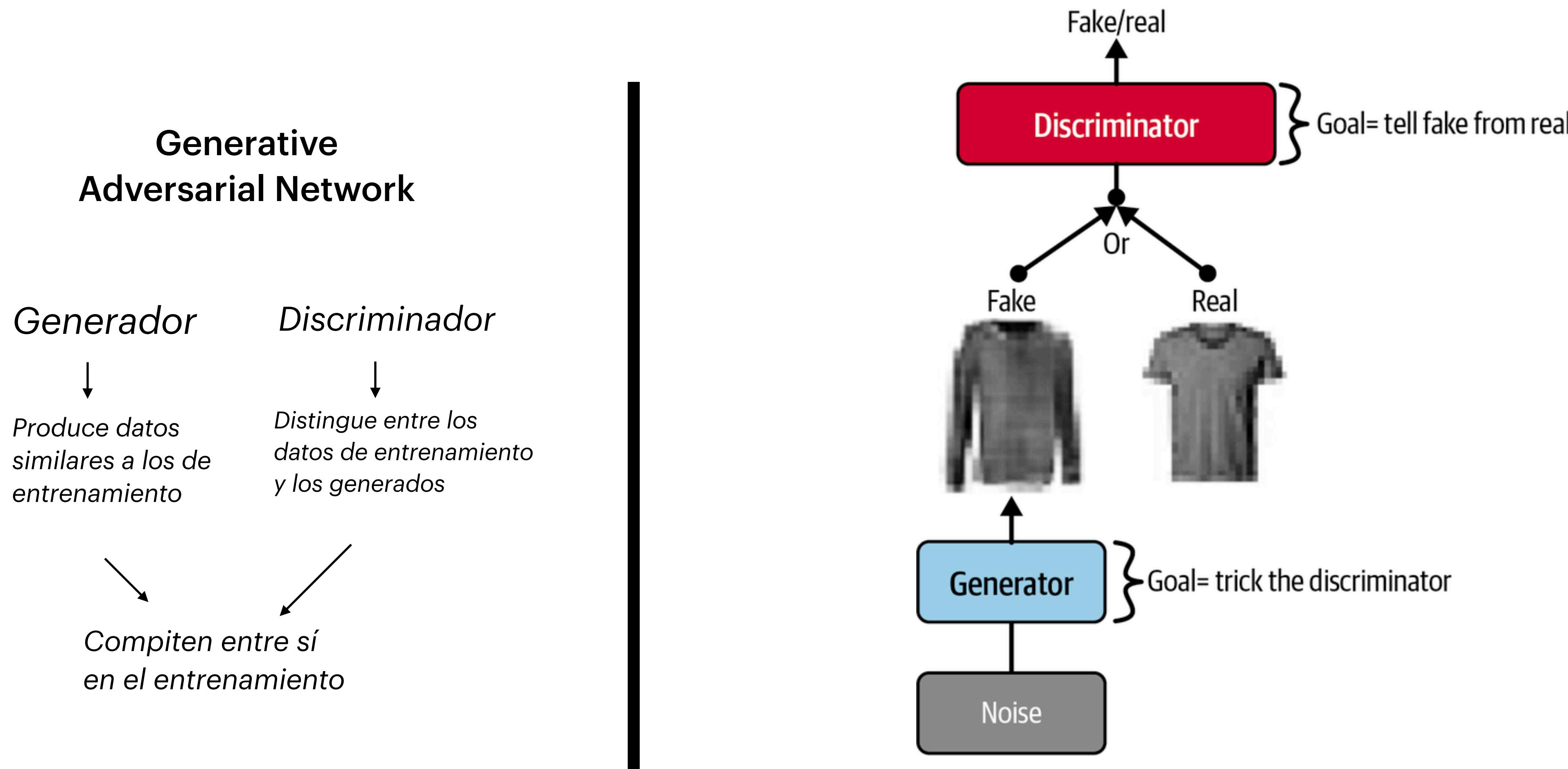
# Encoder-Decoder

## Autoencoders, GANs, Diffusion Models



# Encoder-Decoder

## Autoencoders, GANs, Diffusion Models



# Encoder-Decoder

## Autoencoders

### Autoencoder

Aprende a copiar sus entradas a sus salidas

Codificador  
(encoder)

↓  
Implica restricciones

Decoder  
(decoder)

¿Cuál de las siguientes secuencias numéricas te parece más fácil de memorizar?

40, 27, 25, 36, 81, 57, 10, 73, 19, 68

50, 48, 46, 44, 42, 40, 38, 36, 34,  
32, 30, 28, 26, 24, 22, 20, 18, 16, 14

# Encoder-Decoder

## Autoencoders

### Autoencoder

Aprende a copiar sus entradas a sus salidas

Codificador  
(encoder)

↓  
Implica restricciones

Decoder  
(decoder)

¿Cuál de las siguientes secuencias numéricas te parece más fácil de memorizar?

40, 27, 25, 36, 81, 57, 10, 73, 19, 68

50, 48, 46, 44, 42, 40, 38, 36, 34, 32, 30, 28, 26, 24, 22, 20, 18, 16, 14

Notar patrones les ayuda a almacenar información de manera eficiente.

1. Reconocer entradas
2. Construcción de un espacio latente
3. Salir con algo parecido a la entrada

# Encoder-Decoder

## Autoencoders

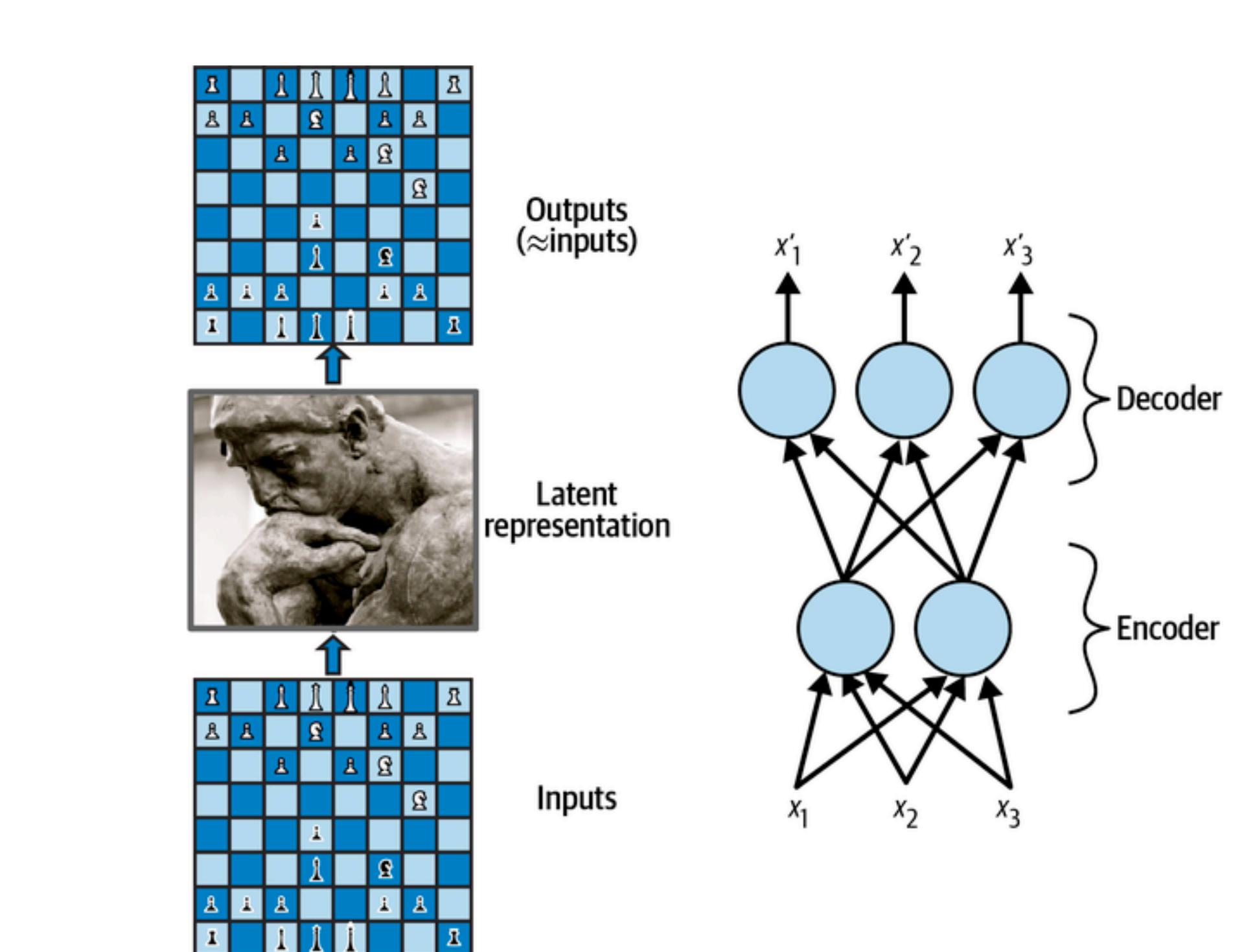
### Autoencoder

Aprende a copiar sus entradas a sus salidas

Codificador  
(encoder)

↓  
*Implica restricciones*

Decoder  
(decoder)



Los expertos en ajedrez no tienen mucha mejor memoria

simplemente ven los patrones de ajedrez más fácilmente gracias a su experiencia con el juego

# Encoder-Decoder

## Autoencoders

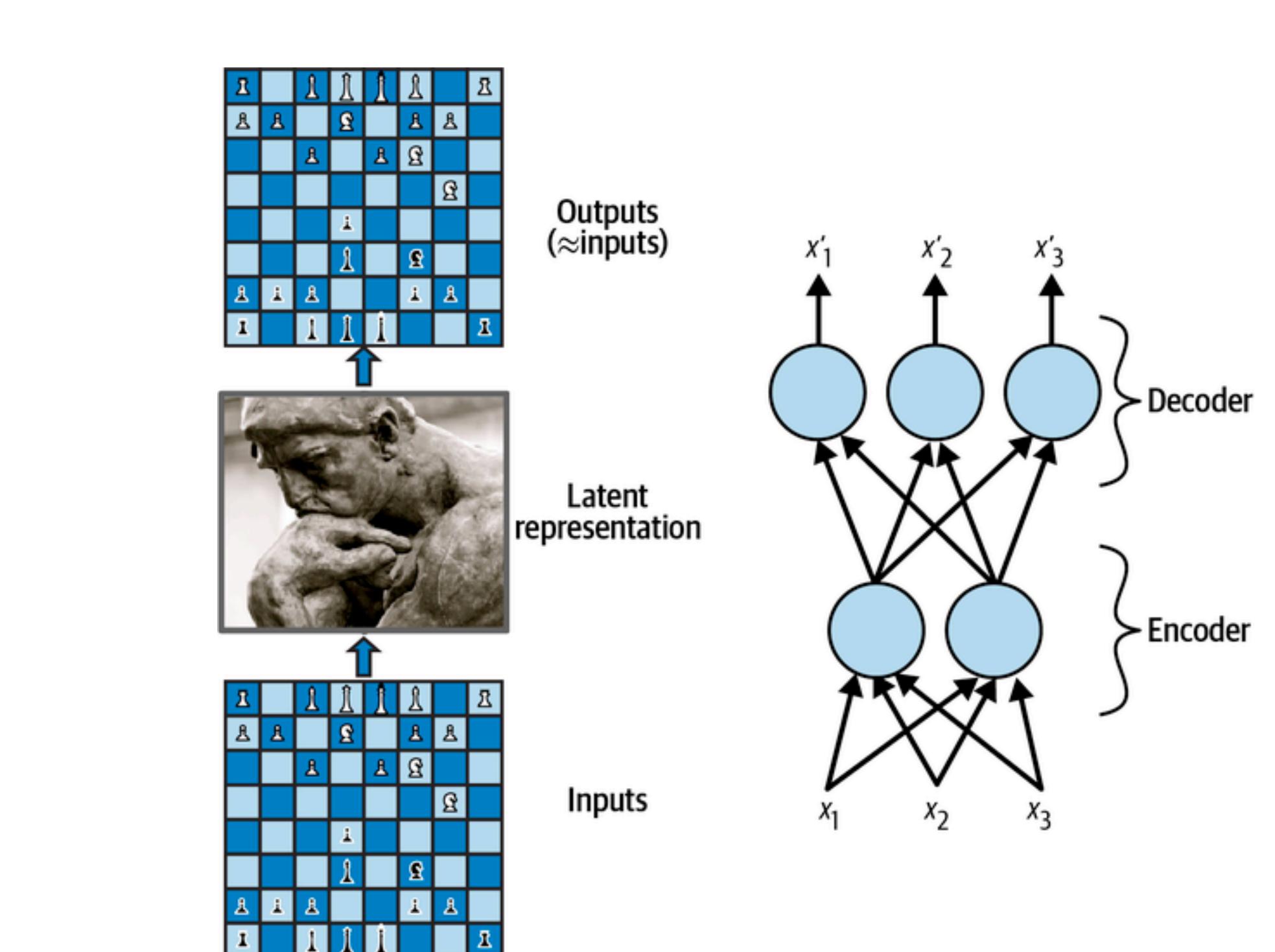
### Autoencoder

Aprende a copiar sus entradas a sus salidas

Codificador  
(encoder)

↓  
*Implica restricciones*

Decoder  
(decoder)



*"Notar patrones les ayuda a almacenar información de manera eficiente."*



*Perception in chess.  
William Chase y  
Herbert Simon*

# Encoder-Decoder

## Autoencoders

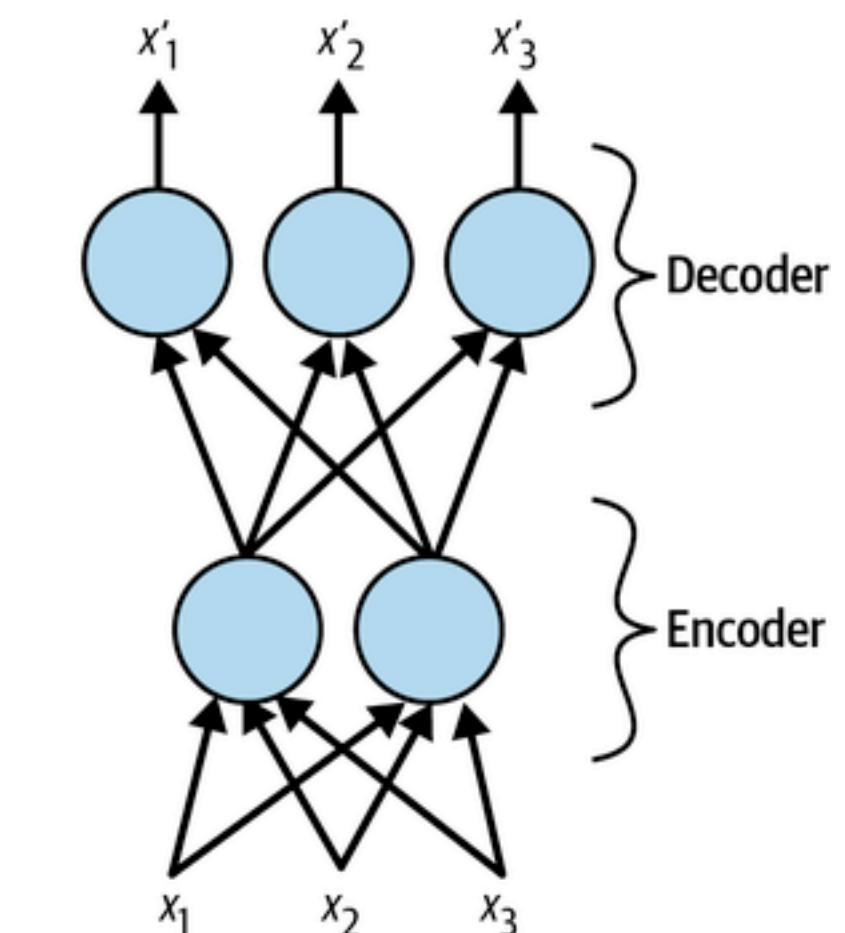
### Autoencoder

Aprende a copiar sus entradas a sus salidas

Codificador  
(encoder)

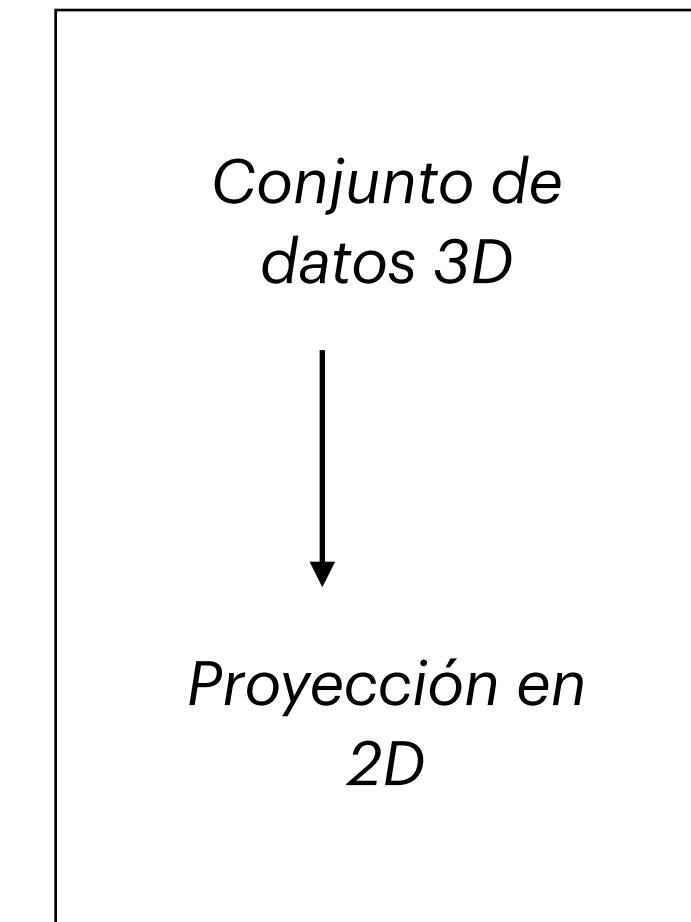
↓  
*Implica restricciones*

Decoder  
(decoder)



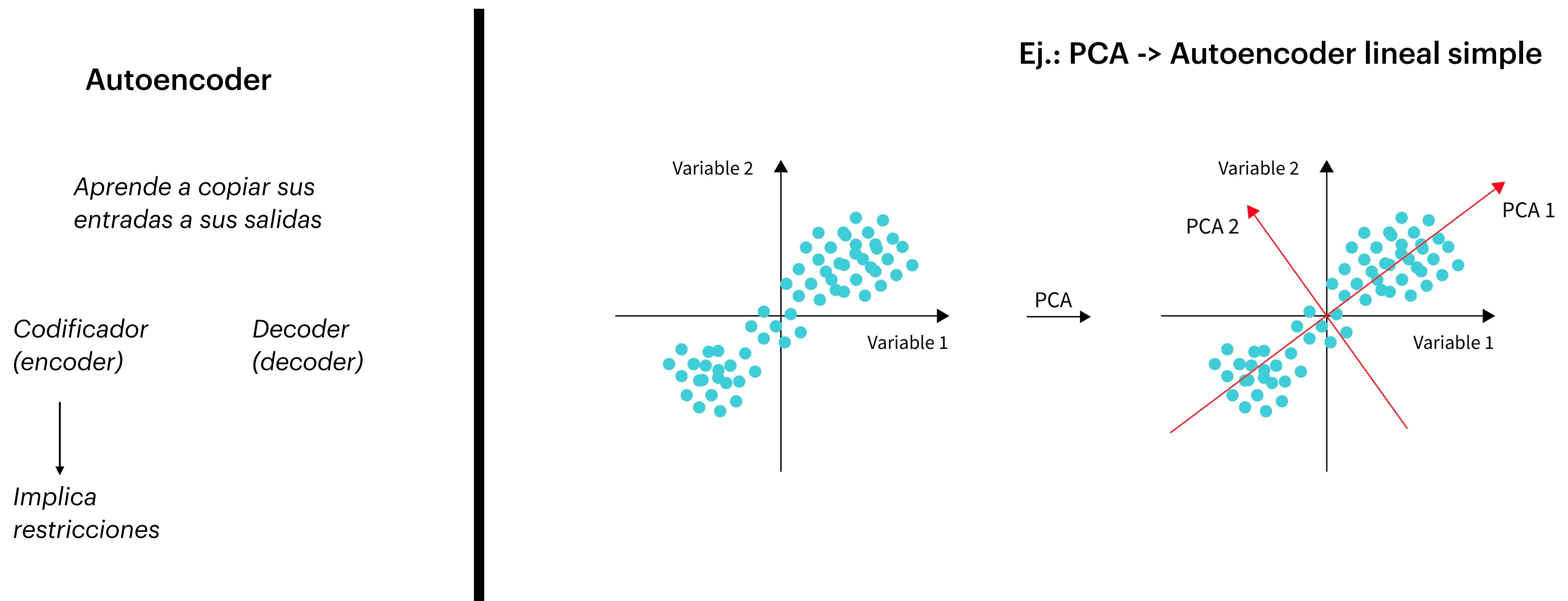
Ej.: PCA -> Autoencoder lineal simple

Realizar PCA



# Encoder-Decoder

## Autoencoders



# Encoder-Decoder

## Autoencoders

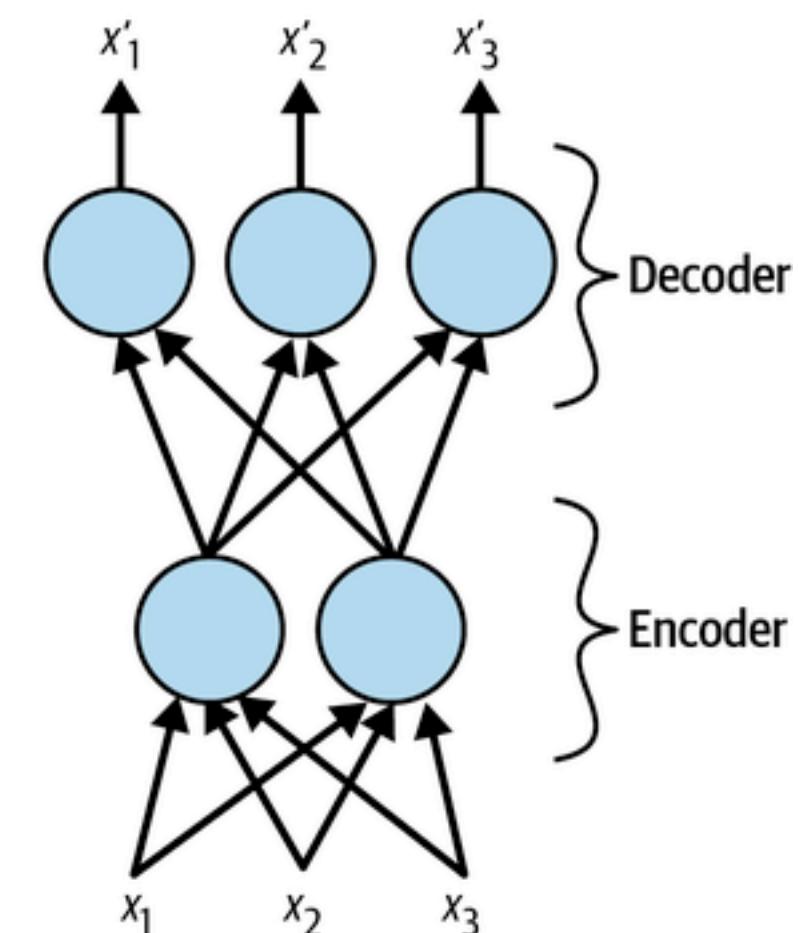
### Autoencoder

Aprende a copiar sus entradas a sus salidas

Codificador  
(encoder)

Decoder  
(decoder)

↓  
Implica restricciones



Ej.: PCA -> Autoencoder lineal simple

```
import tensorflow as tf

encoder = tf.keras.Sequential([tf.keras.layers.Dense(2)])
decoder = tf.keras.Sequential([tf.keras.layers.Dense(3)])
autoencoder = tf.keras.Sequential([encoder, decoder])

optimizer = tf.keras.optimizers.SGD(learning_rate=0.5)
autoencoder.compile(loss="mse", optimizer=optimizer)
```

```
X_train = [...] # generate a 3D dataset
history = autoencoder.fit(X_train, X_train, epochs=500, verbose=0)
codings = encoder.predict(X_train)
```

# Encoder-Decoder

## Autoencoders

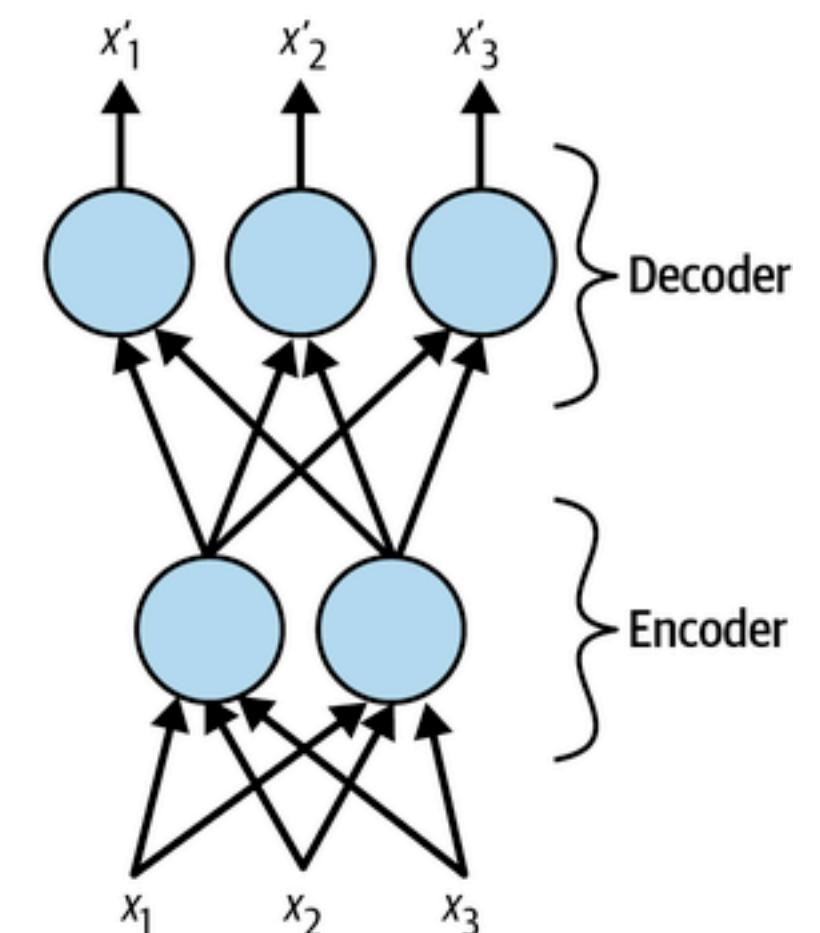
### Autoencoder

Aprende a copiar sus entradas a sus salidas

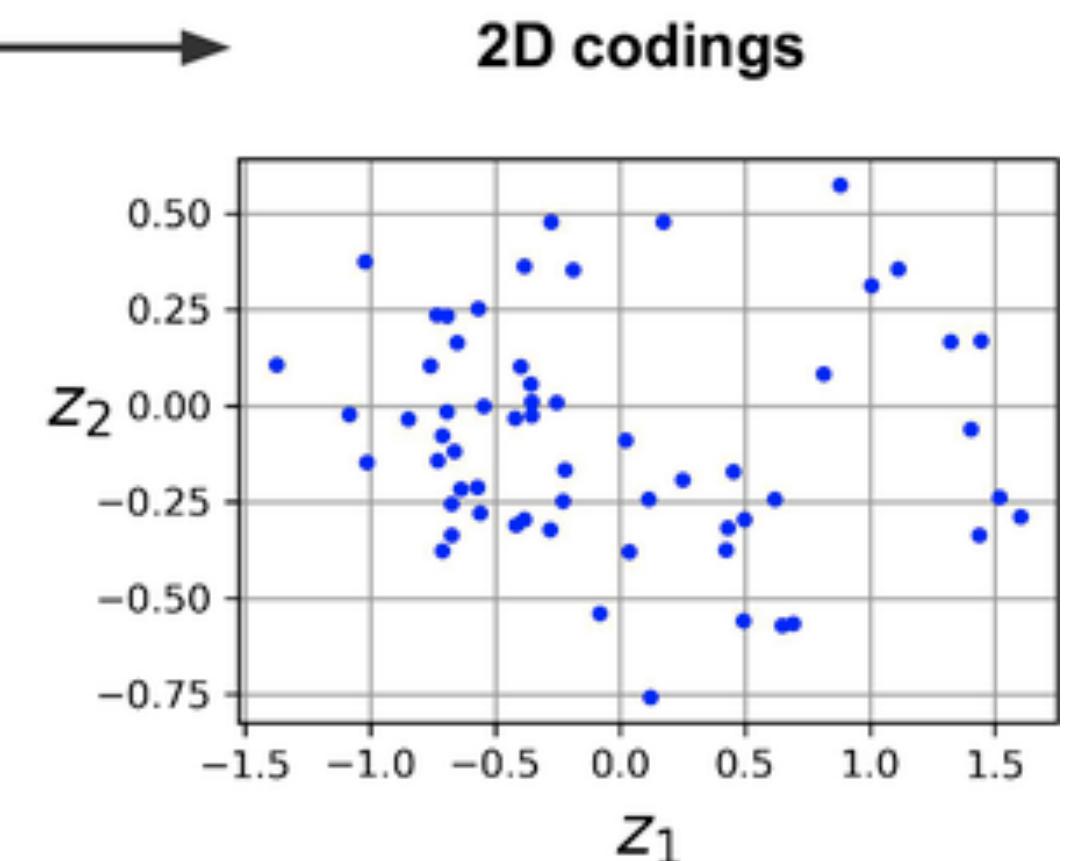
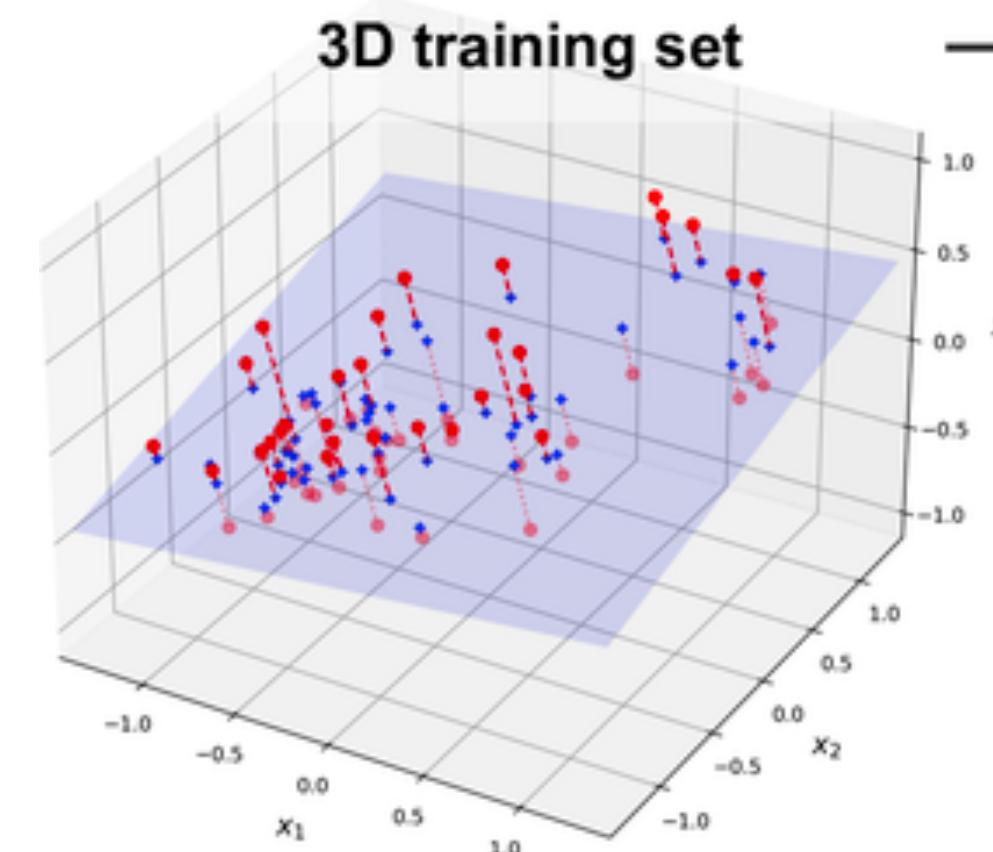
Codificador  
(encoder)

↓  
Implica restricciones

Decoder  
(decoder)



Ej.: PCA -> Autoencoder lineal simple



# Encoder-Decoder

## Autoencoders

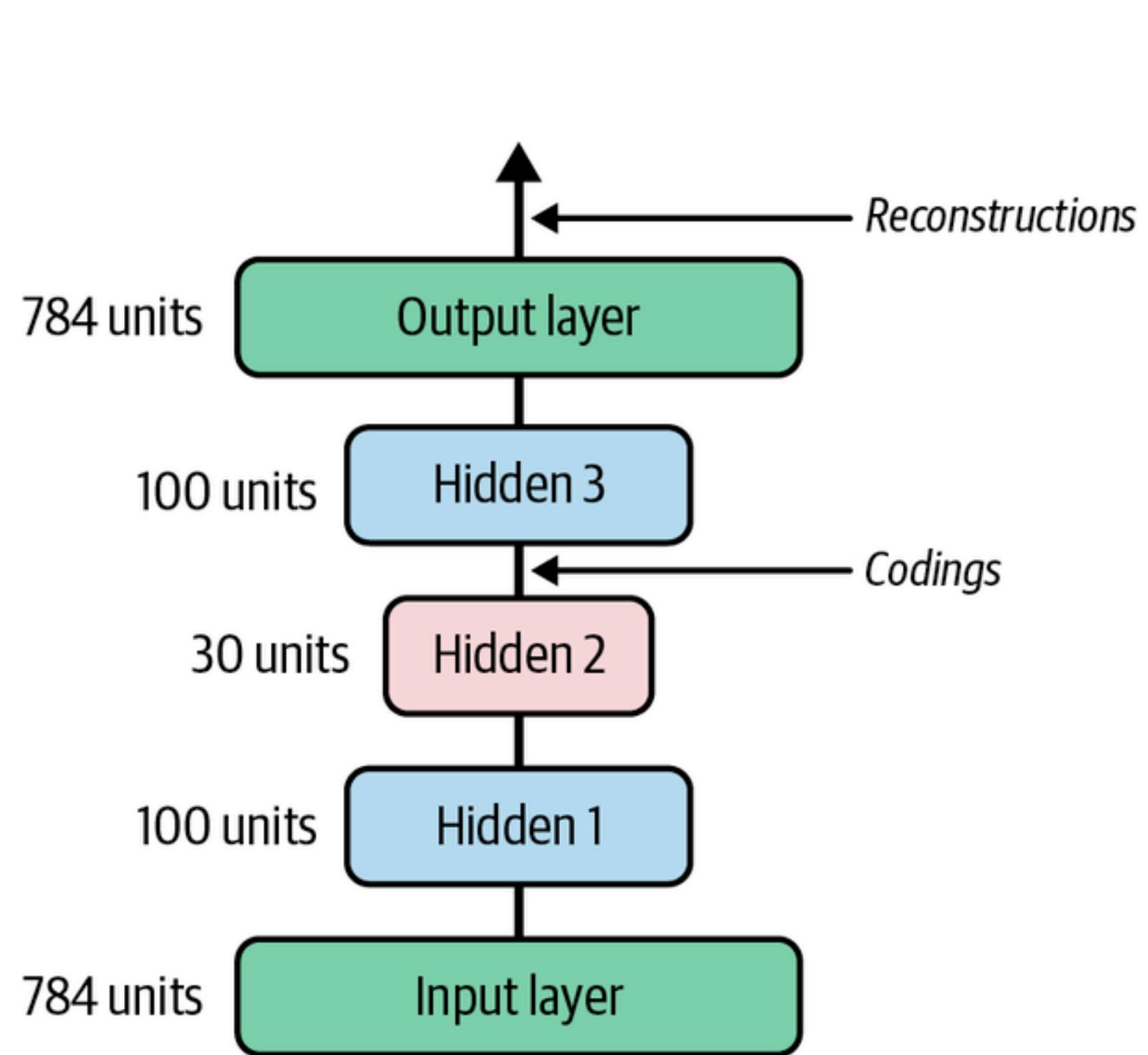
### Stacked Autoencoders

Aprende a copiar sus entradas a sus salidas

Codificador  
(encoder)

Decoder  
(decoder)

Varias capas  
ocultas



### Simetría:

1. Conservación de información.
2. Facilita el aprendizaje
3. Regularización
4. Transferencia de conocimiento

# Encoder-Decoder

## Autoencoders

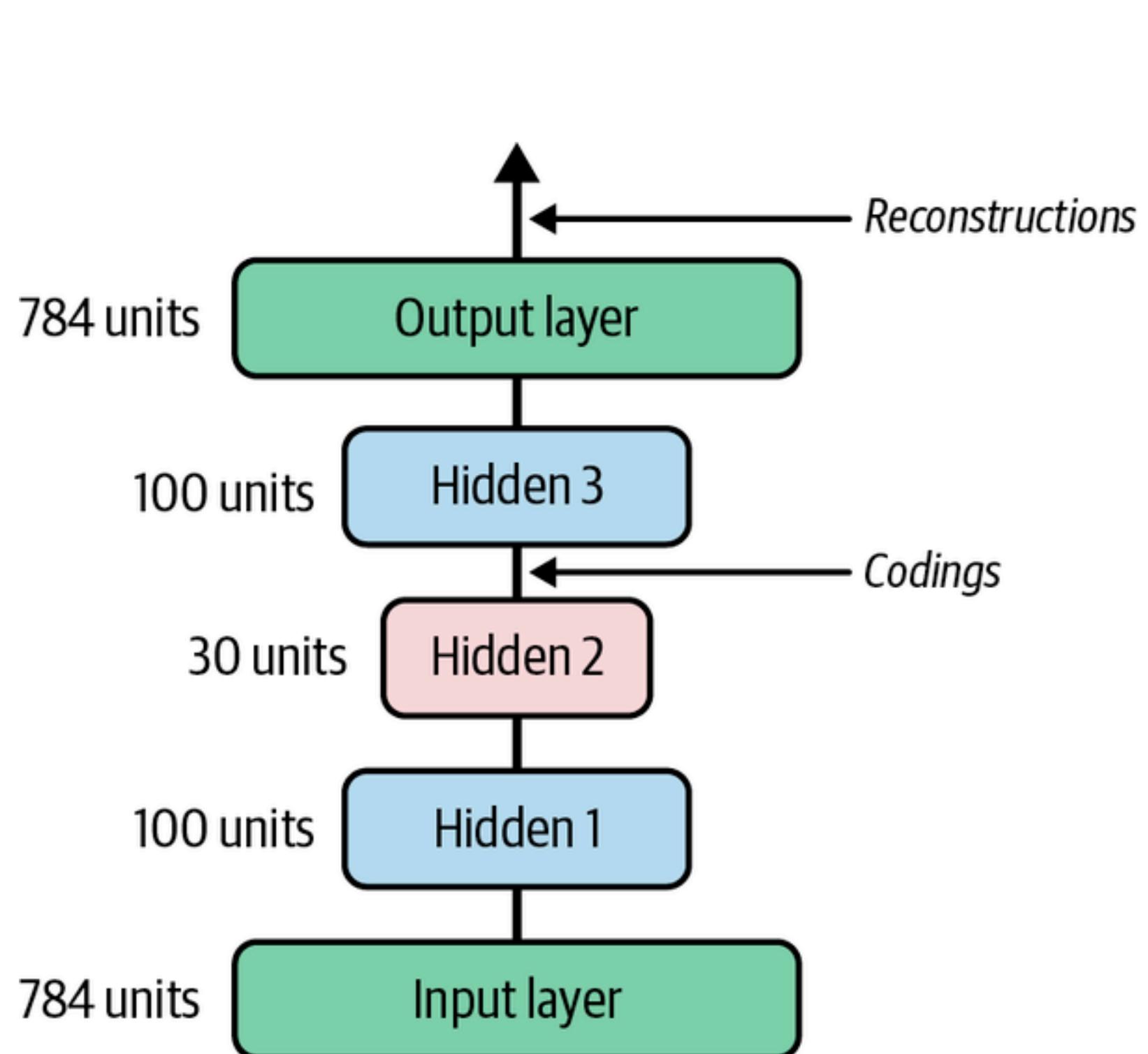
### Stacked Autoencoders

Aprende a copiar sus entradas a sus salidas

Codificador  
(encoder)

Decoder  
(decoder)

Varias capas  
ocultas



### Simetría:

1. Conservación de información.

Si no fueran simétricos habría perdida de información

Relación directa entre la representación de entrada y la representación codificada

# Encoder-Decoder

## Autoencoders

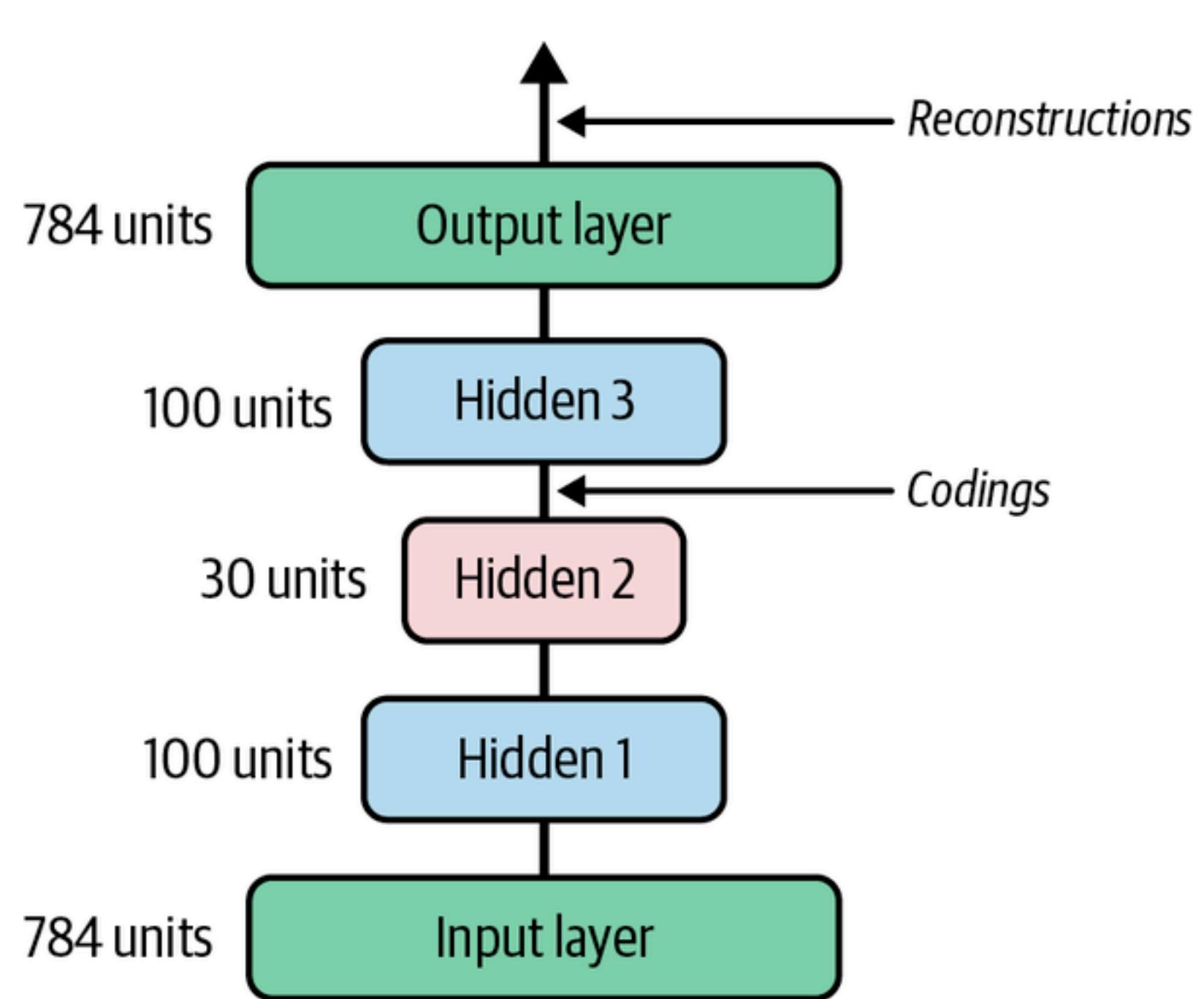
### Stacked Autoencoders

Aprende a copiar sus entradas a sus salidas

Codificador  
(encoder)

Decoder  
(decoder)

Varias capas  
ocultas



### Simetría:

2. Facilita el aprendizaje

Si no fueran simétricos sería más complejo entrar configuración adecuada de hiperparámetros

Convergencia a una solución óptima

# Encoder-Decoder

## Autoencoders

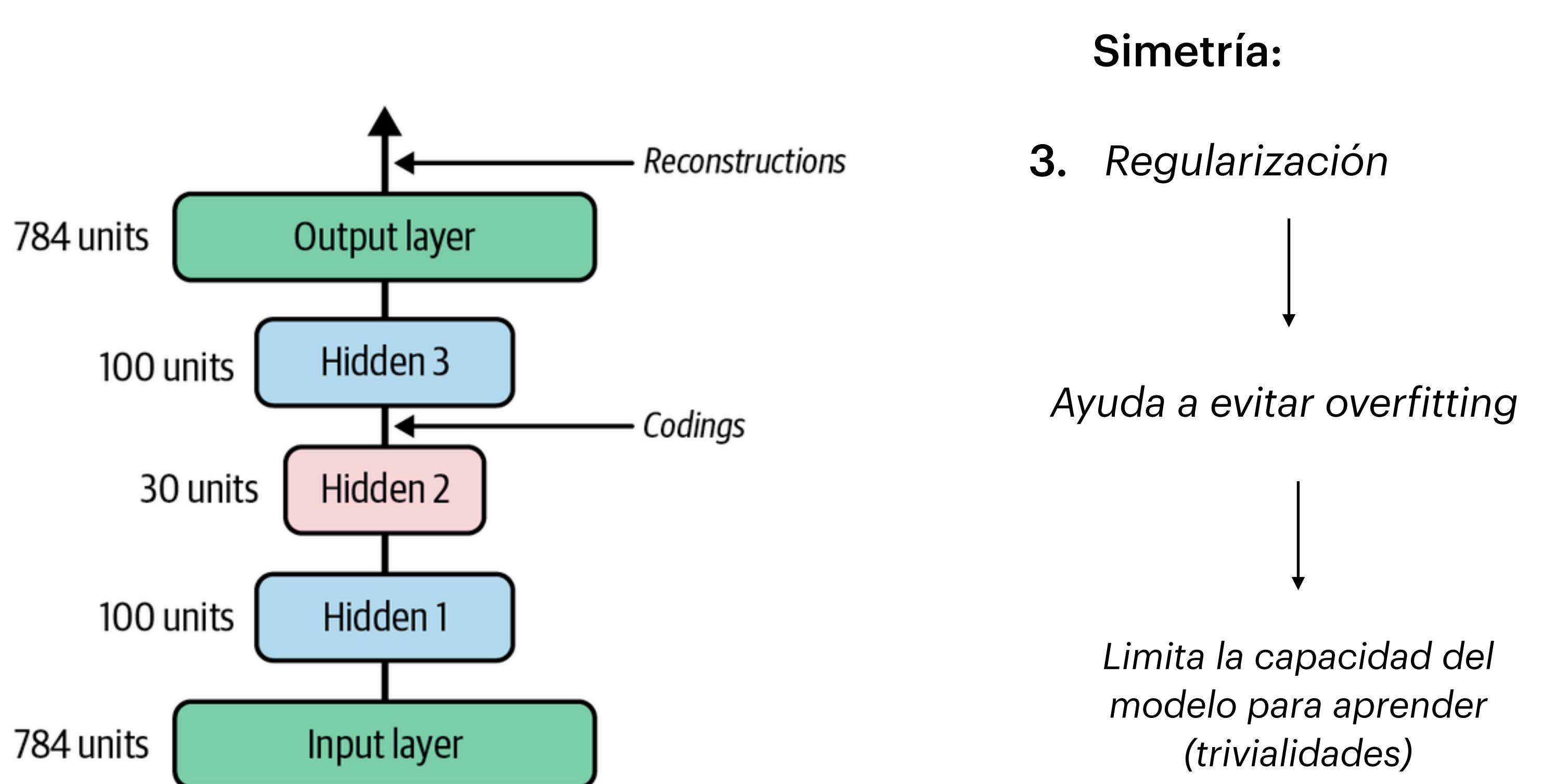
### Stacked Autoencoders

Aprende a copiar sus entradas a sus salidas

Codificador  
(encoder)

Decoder  
(decoder)

Varias capas  
ocultas



# Encoder-Decoder

## Autoencoders

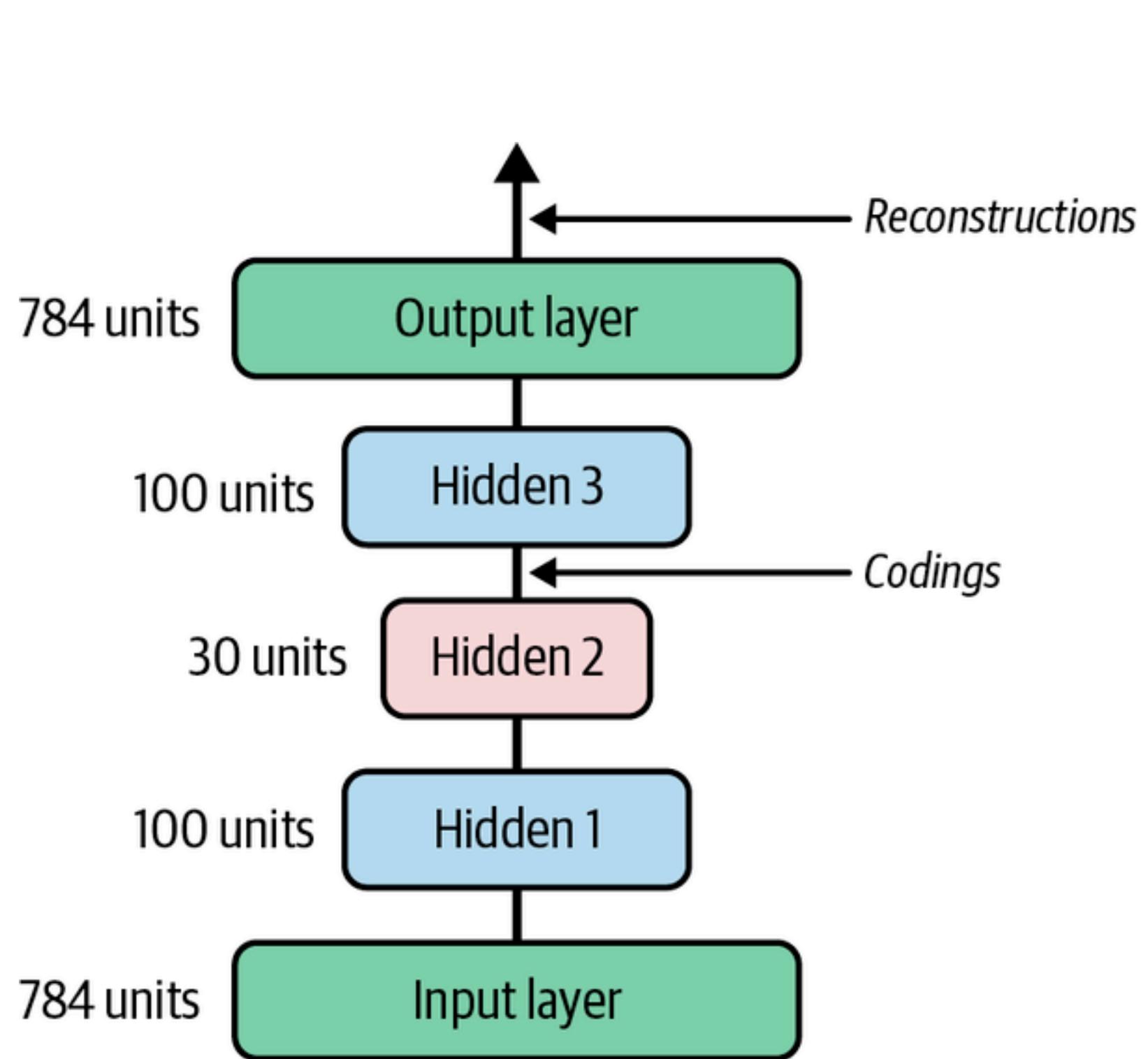
### Stacked Autoencoders

Aprende a copiar sus entradas a sus salidas

Codificador  
(encoder)

Decoder  
(decoder)

Varias capas  
ocultas



### Simetría:

4. Transferencia de conocimiento

↓  
Permite adaptarse a otras tareas de deep learning

↓  
Uso común de capas del codificador en 'transfer learning'

# Encoder-Decoder

## Autoencoders

### Stacked Autoencoders

Aprende a copiar sus  
entradas a sus salidas

Codificador  
(encoder)

Decoder  
(decoder)

Varias capas  
ocultas



Simetría:

4. Transferencia de conocimiento



Permite adaptarse a otras  
tareas de deep learning



Uso común de capas del  
codificador en 'transfer  
learning'

# Encoder-Decoder

## Autoencoders

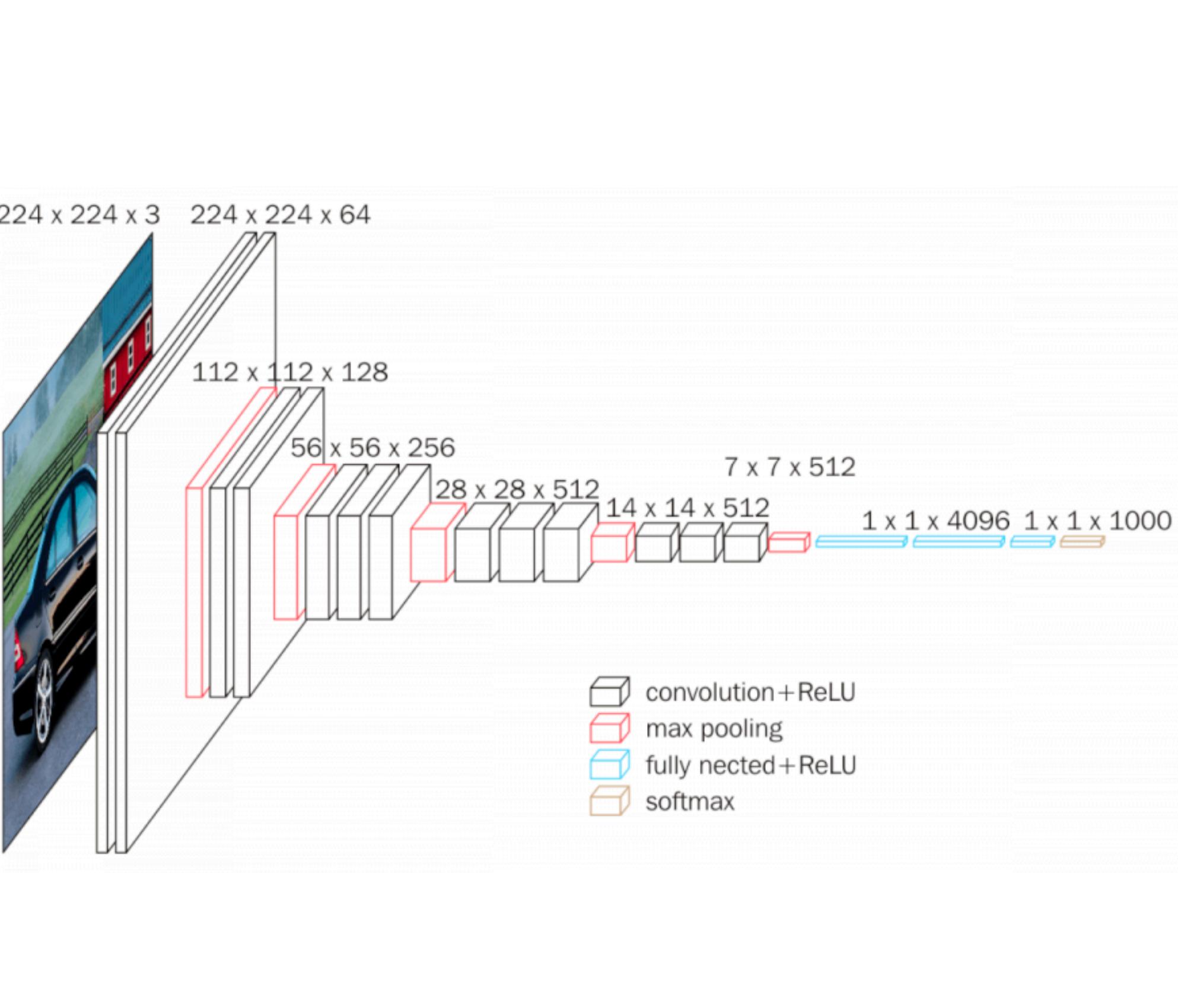
### Stacked Autoencoders

Aprende a copiar sus entradas a sus salidas

Codificador  
(encoder)

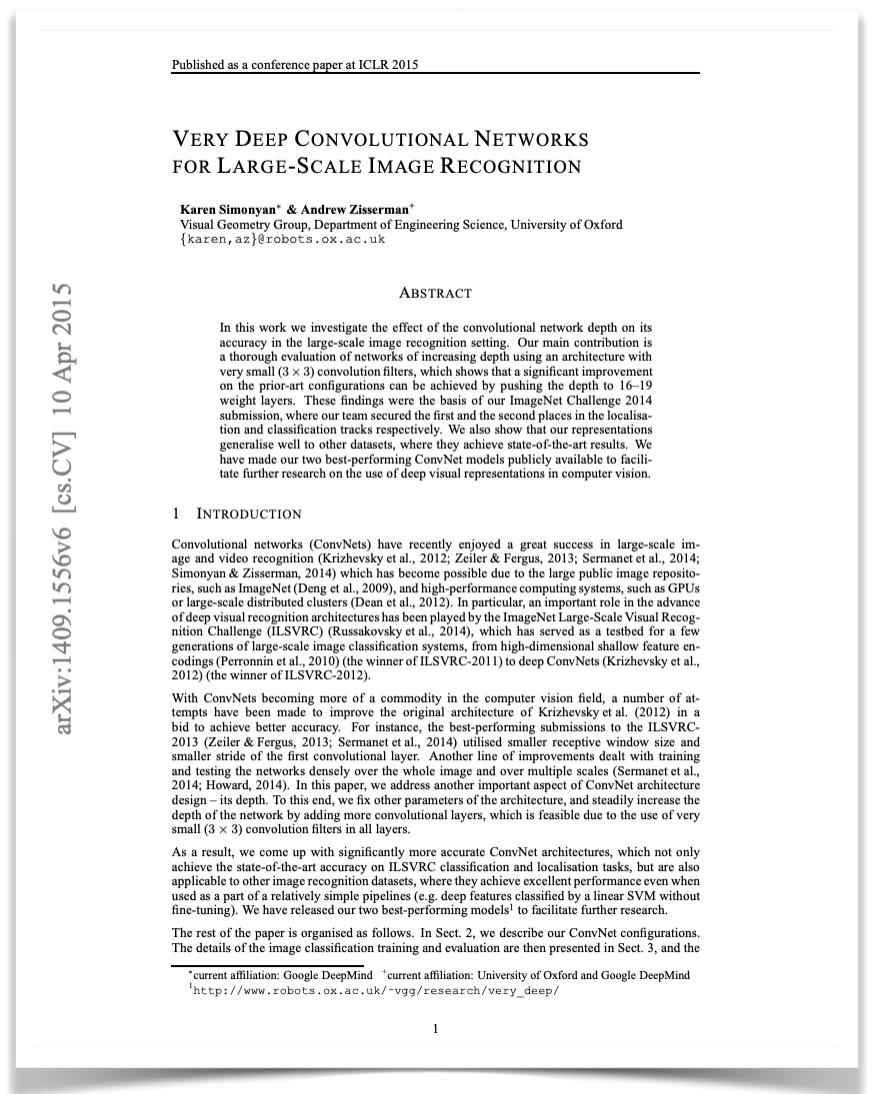
Decoder  
(decoder)

Varias capas  
ocultas



### Transfer Learning

Ej.: VGG-16 -> Uso de contorno, formas y colores.



# Encoder-Decoder

## Autoencoders

### Stacked Autoencoders

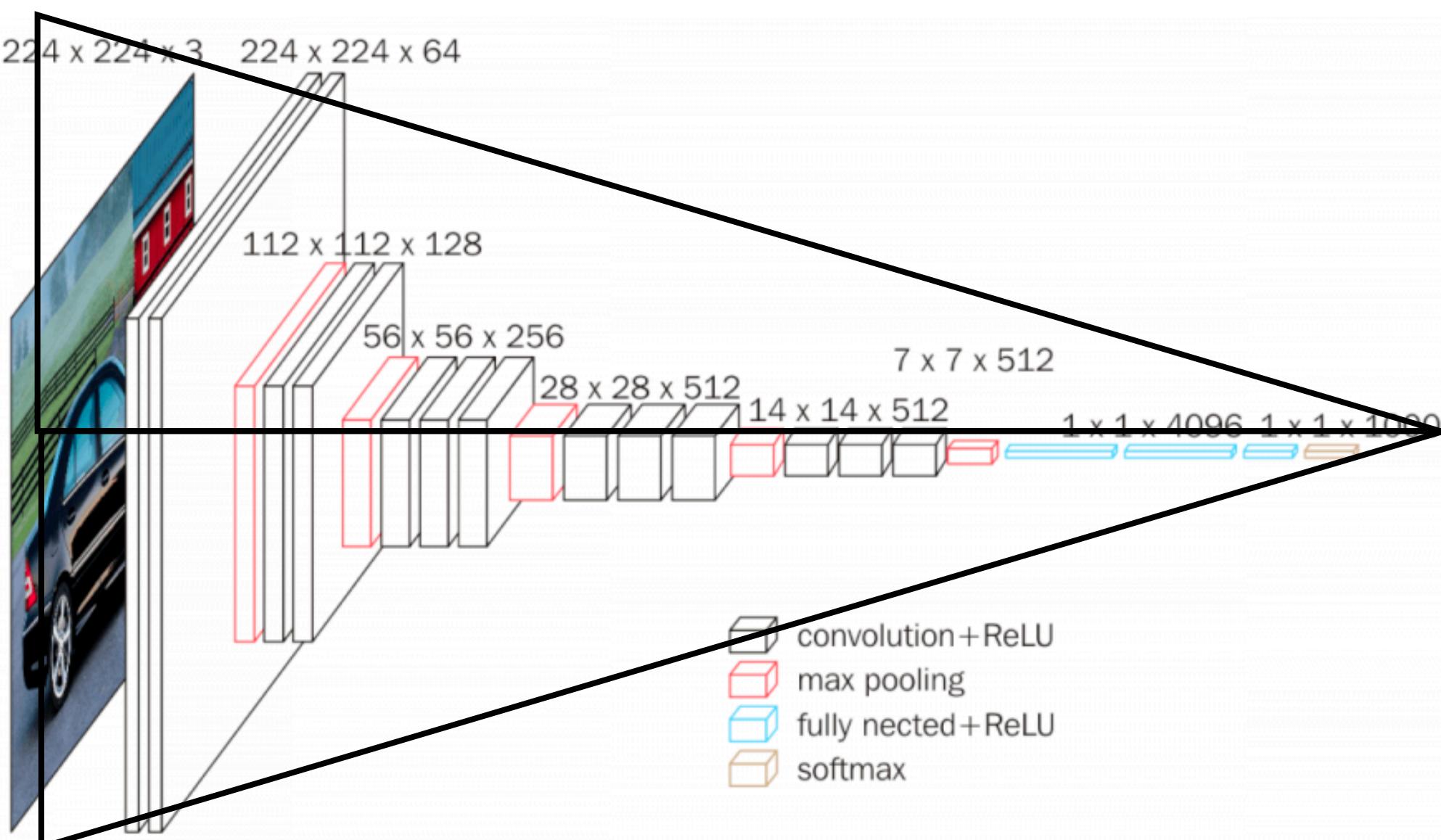
Aprende a copiar sus entradas a sus salidas

Codificador  
(encoder)

Decoder  
(decoder)

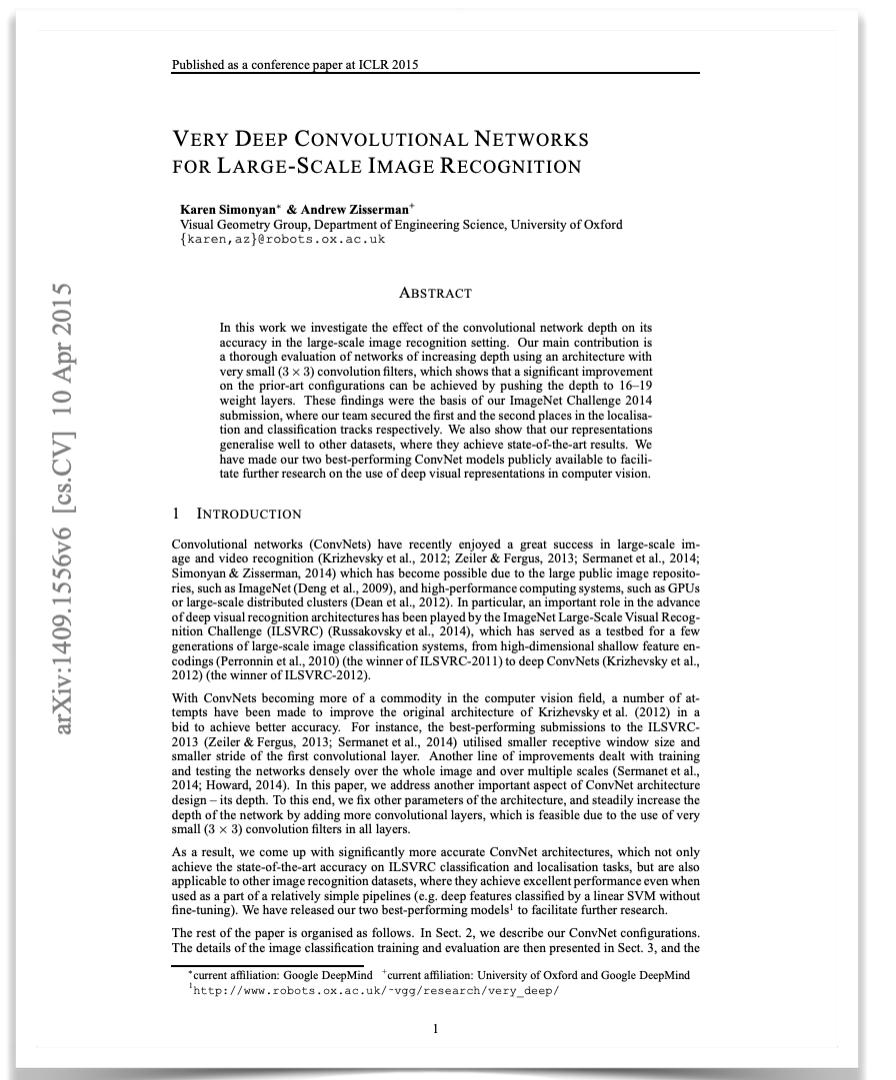
Varias capas ocultas

*La arquitectura de la VGG-16 le permitió ganar la competición ILSVRC(Imagenet) 2014*



### Transfer Learning

Ej.: VGG-16 -> Uso de contorno, formas y colores.



# Encoder-Decoder

## Autoencoders

### Stacked Autoencoders

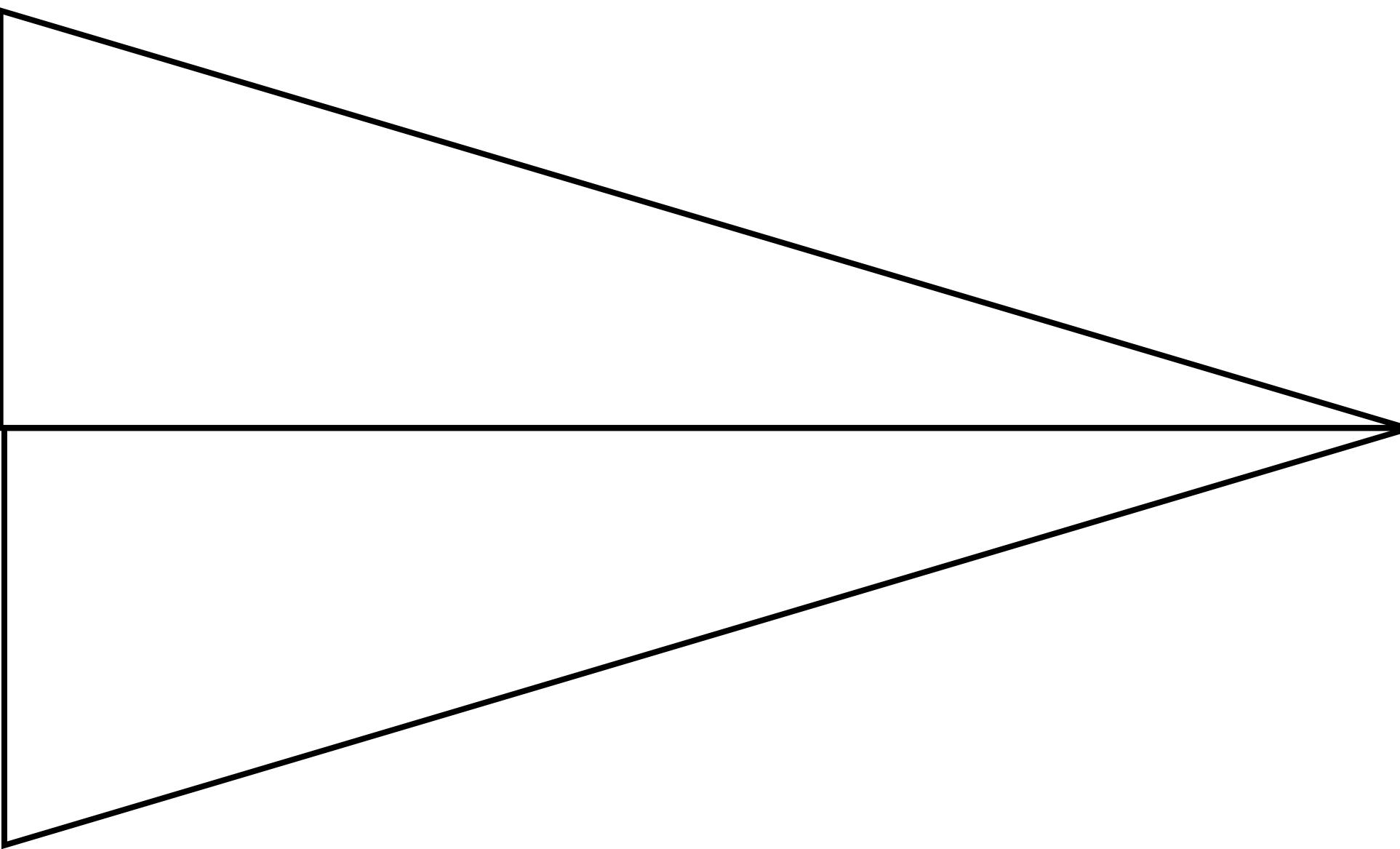
Aprende a copiar sus  
entradas a sus salidas

Codificador  
(encoder)

Decodificador  
(decoder)

Varias capas  
ocultas

### Diseño de codificador



### Destilación:

Eliminar parte del contenido  
para tener una versión más  
'concentrada'.

- Comprensión de información
- Extracción de características
- Generalización
- Capacidad de reconstrucción
- Simetría de los datos

# Encoder-Decoder

## Autoencoders

### Stacked Autoencoders

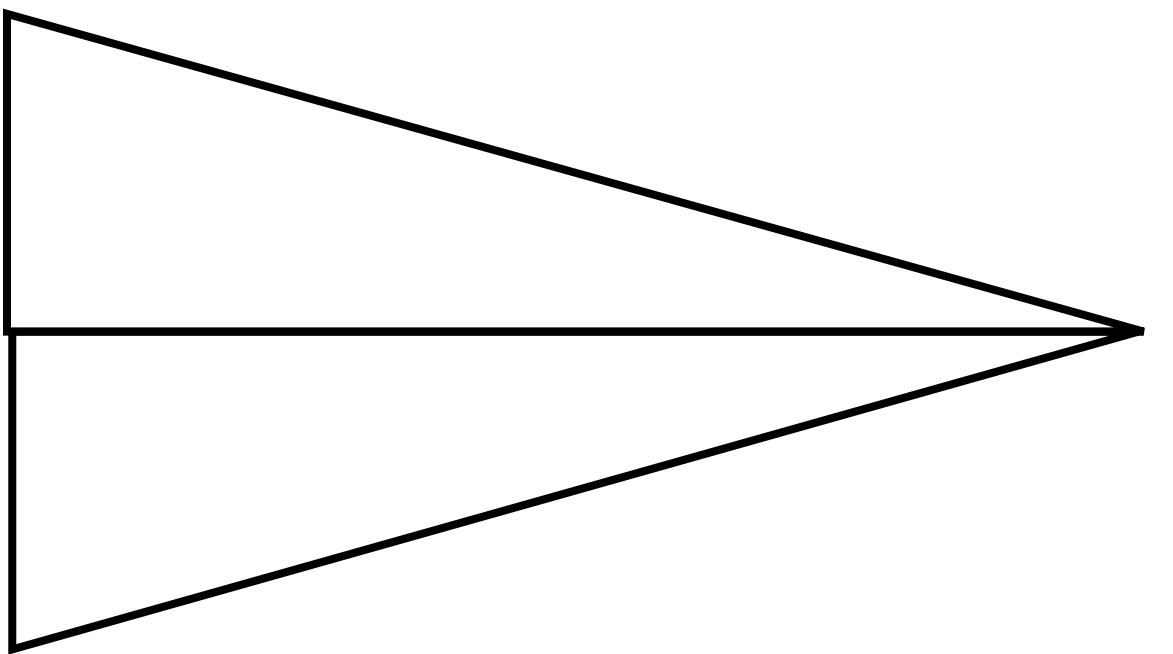
Aprende a copiar sus  
entradas a sus salidas

Codificador  
(encoder)

Decodificador  
(decoder)

Varias capas  
ocultas

### Diseño de codificador



Comprensión de  
información

1. Disminuye la cantidad  
de neuronas por capa

2. Aprender representaciones  
más compactas/esenciales

# Encoder-Decoder

## Autoencoders

### Stacked Autoencoders

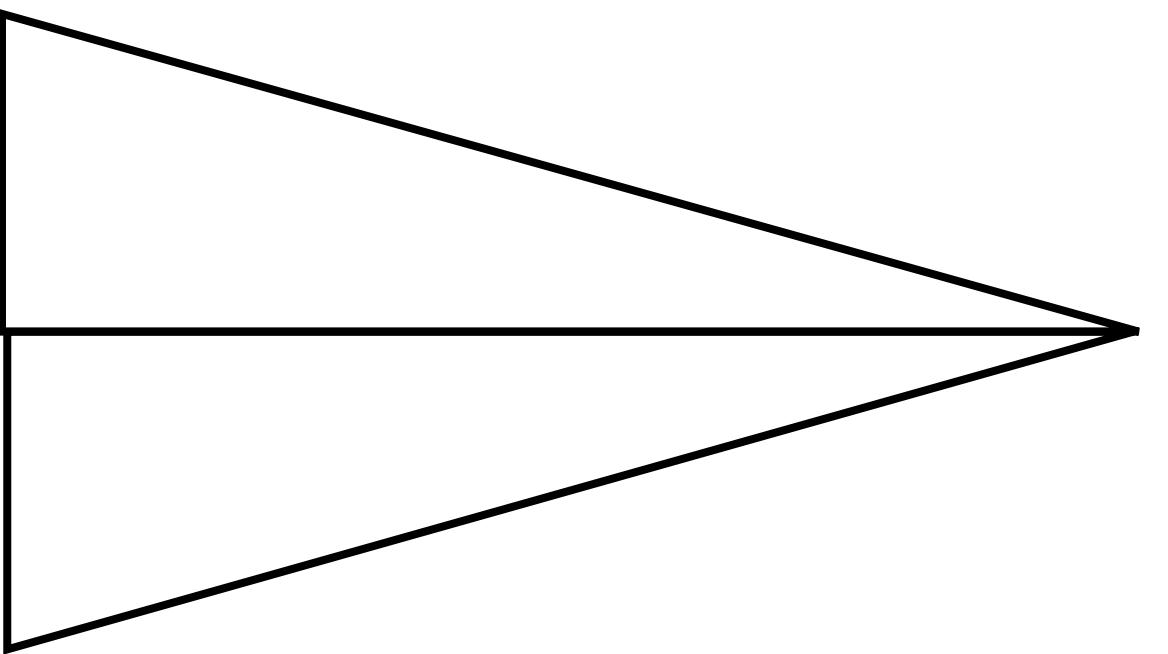
Aprende a copiar sus  
entradas a sus salidas

Codificador  
(encoder)

Decodificador  
(decoder)

Varias capas  
ocultas

### Diseño de codificador



Extractor de características jerárquico

Extracción de  
características

1. Capas iniciales aprenden características simples y locales.
2. Capas posteriores aprenden características abstractas y globales.

# Encoder-Decoder

## Autoencoders

### Stacked Autoencoders

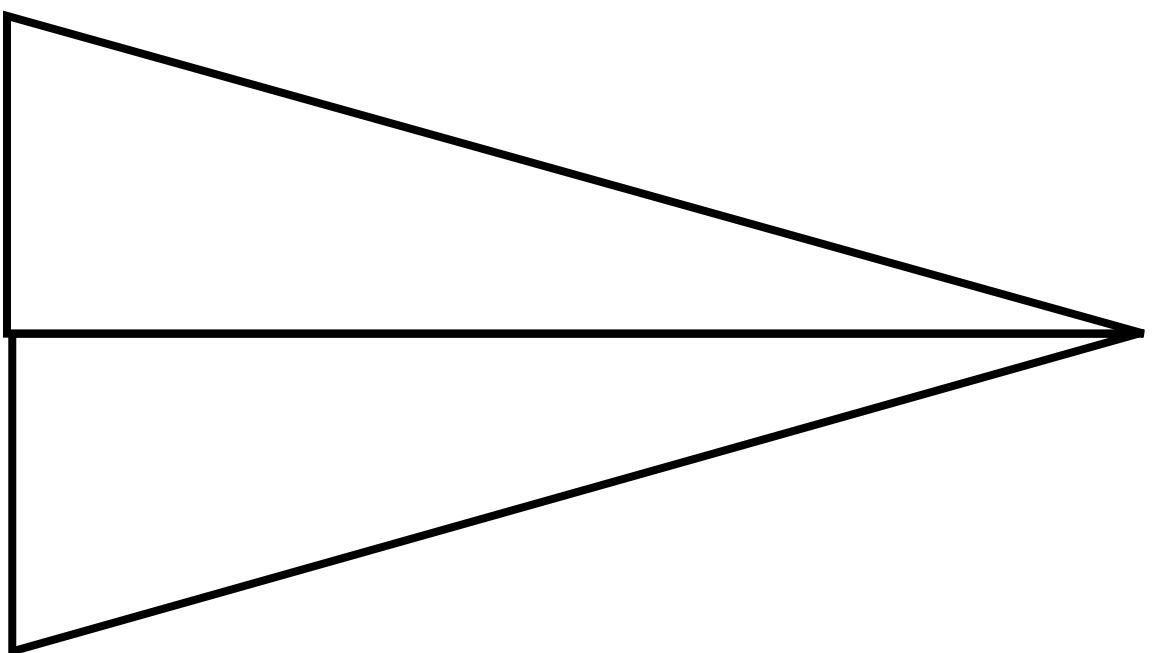
Aprende a copiar sus  
entradas a sus salidas

Codificador  
(encoder)

Decodificador  
(decoder)

Varias capas  
ocultas

### Diseño de codificador



Generalización/  
Regularización

1. Reducción progresiva  
ayuda a reducir el riesgo de  
*overfitting*
2. Evita memorización de los  
datos
3. Aprenda patrones  
generales para generalizar

# Encoder-Decoder

## Autoencoders

### Stacked Autoencoders

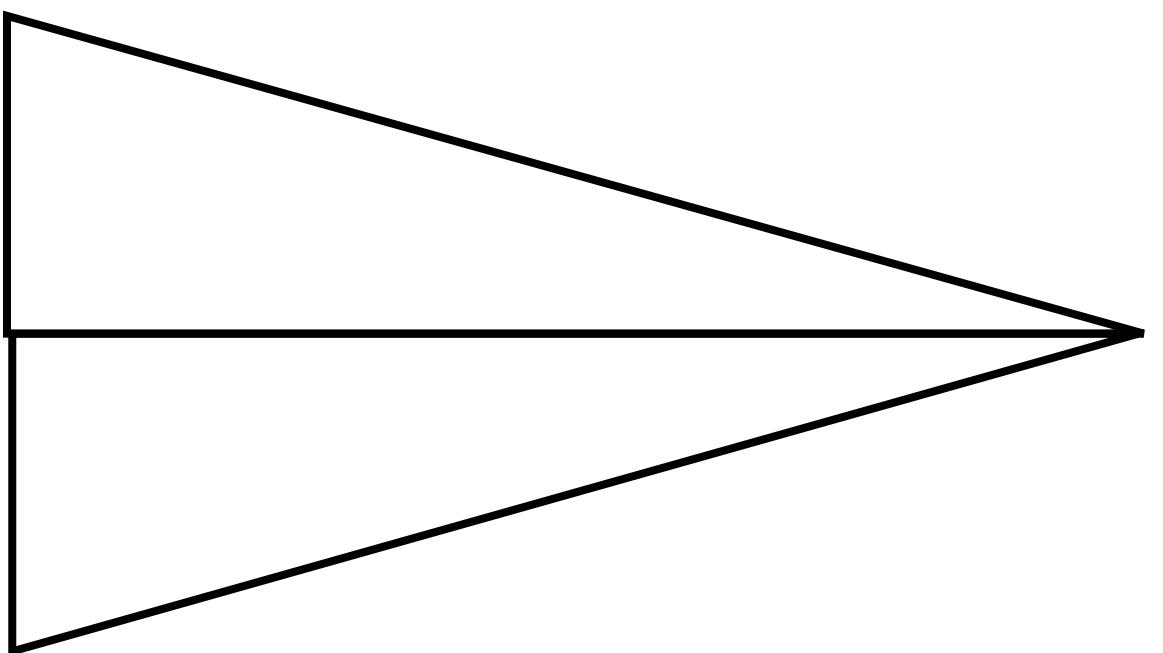
Aprende a copiar sus  
entradas a sus salidas

Codificador  
(encoder)

Decodificador  
(decoder)

Varias capas  
ocultas

### Diseño de codificador



Simetría en el  
diseño

1. Facilidad de entrenamiento
2. Eficiencia en la reconstrucción de datos
3. Reducción y aumento, forman la simetría