

Intuición sobre Autocodificadores

“AutoEncoders”

Autocodificadores

- Por la forma en que los usaremos, caen bajo la categoría de Aprendizaje no supervisado debido a que:
 - Durante el entrenamiento de ciertos modelos, a veces podremos tener datos históricos con sus correspondientes etiquetas. Sin embargo durante el uso real del modelo no será posible utilizar métricas de evaluación previa
- Usaremos Autocodificadores para:
 - reducción de dimensiones - es el uso más común, y
 - eliminación de ruido

AutoCodificadores

- El término aprendizaje no supervisado, implica que no hay etiquetas “correctas” con las que se puede comparar
- A veces, a los casos de uso que se trabajarán, se les denomina semi-supervisado porque:
 - durante el entrenamiento, pueden haber datos con etiquetas correctas
 - En el uso real del autocodificador no es posible producir métricas como exactitud o RMSE

¿Cómo funcionan los autocodificadores?

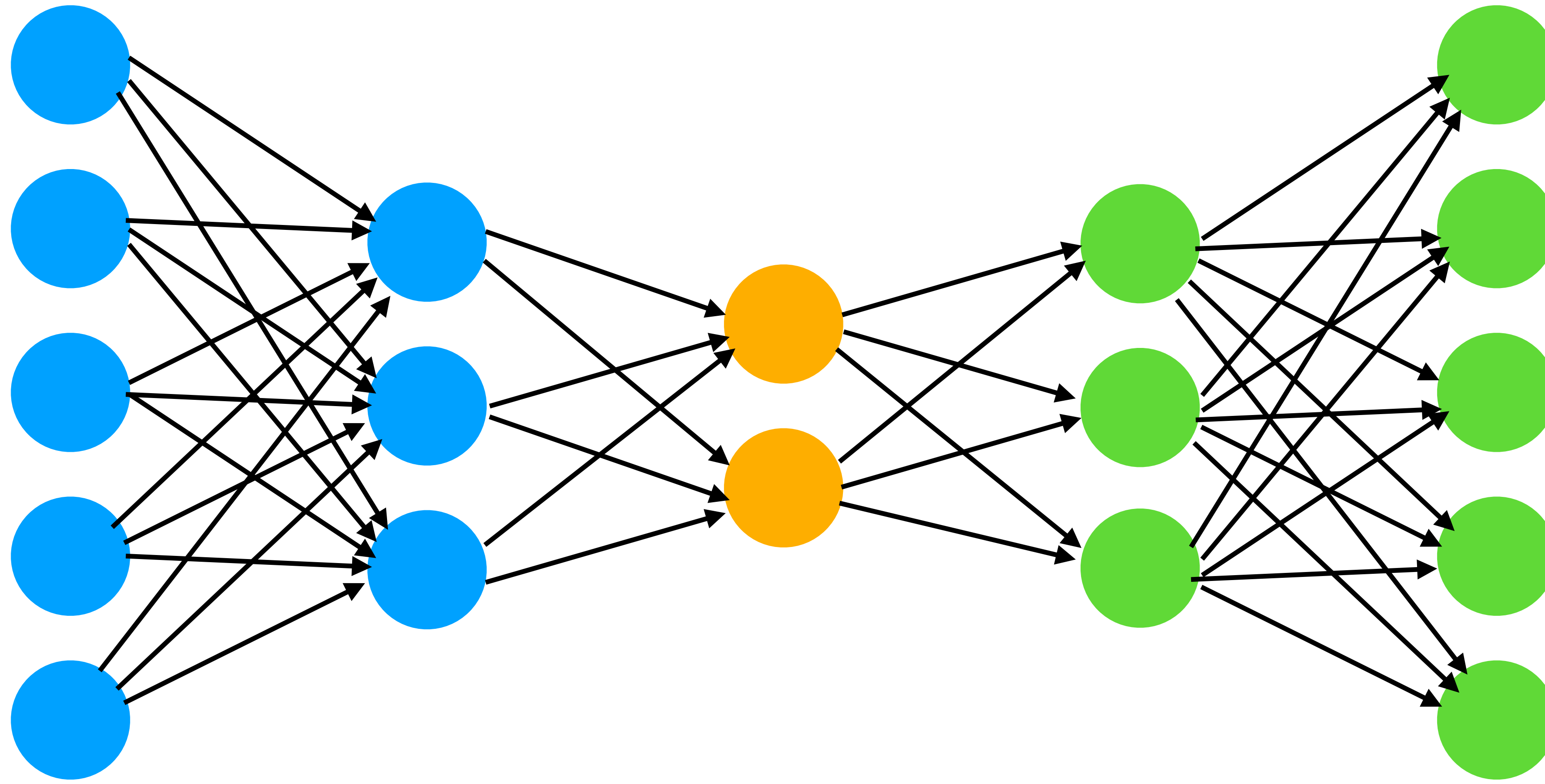
¿Cuáles son los elementos básicos?

Fundamentos de los autocodificadores

- Son redes neuronales muy simples y muy parecidas a los modelos de perceptrones multi-capas (MLPs)
- Están diseñados para reproducir los datos de entrada, en la capa de salida
- La diferencia principal entre los autocodificadores y los modelos MLPs es que el número de neuronas de entrada es igual al número de neuronas de salida

Autocodificador

Ejemplo

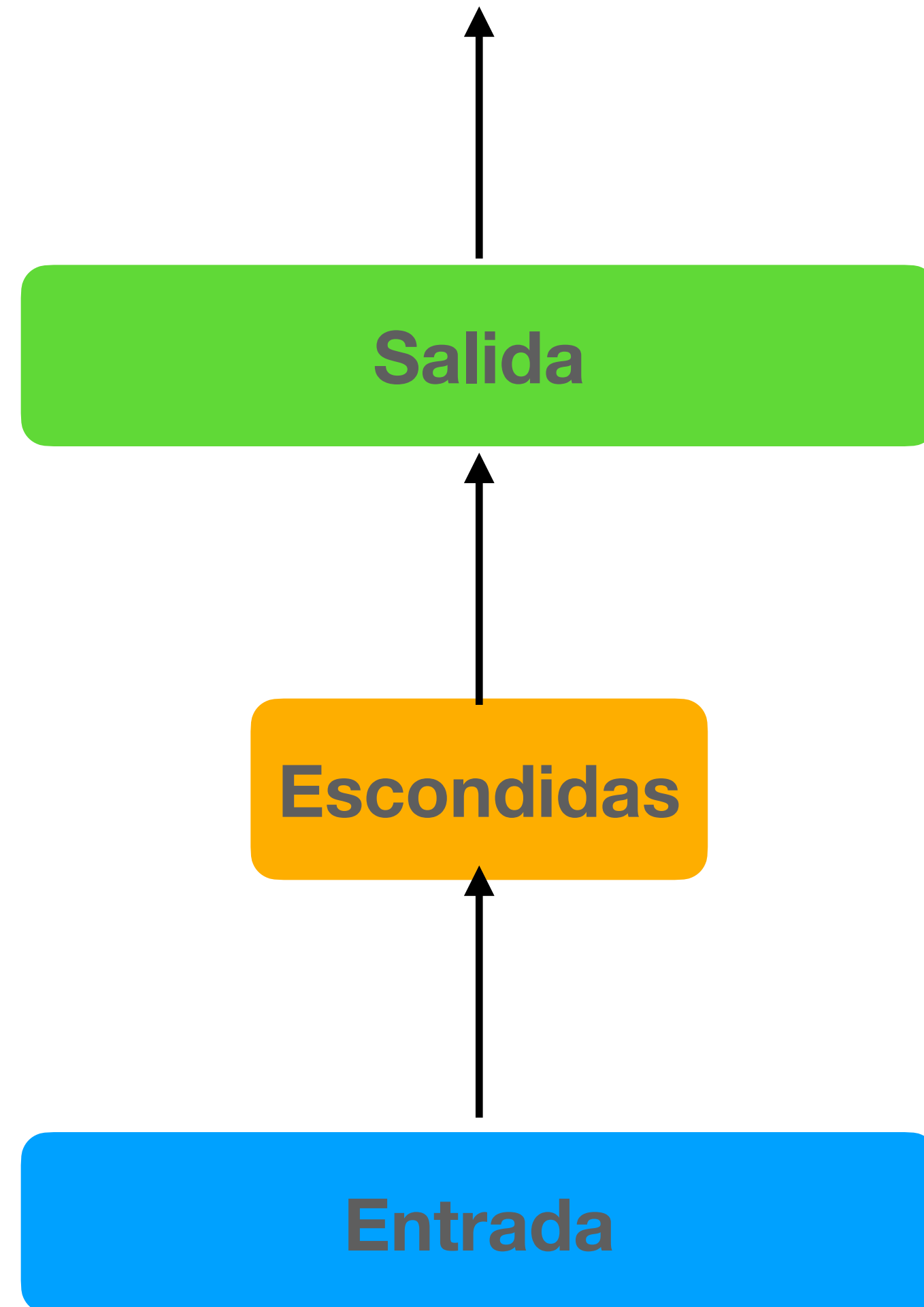


Idea básica es:

- Reducir las entradas a un número dado de neuronas (centro)
- Reproducir la salida a partir del centro
- Las capas escondidas deben aprender qué características (features) son importantes

Autocodificador

Esquemático



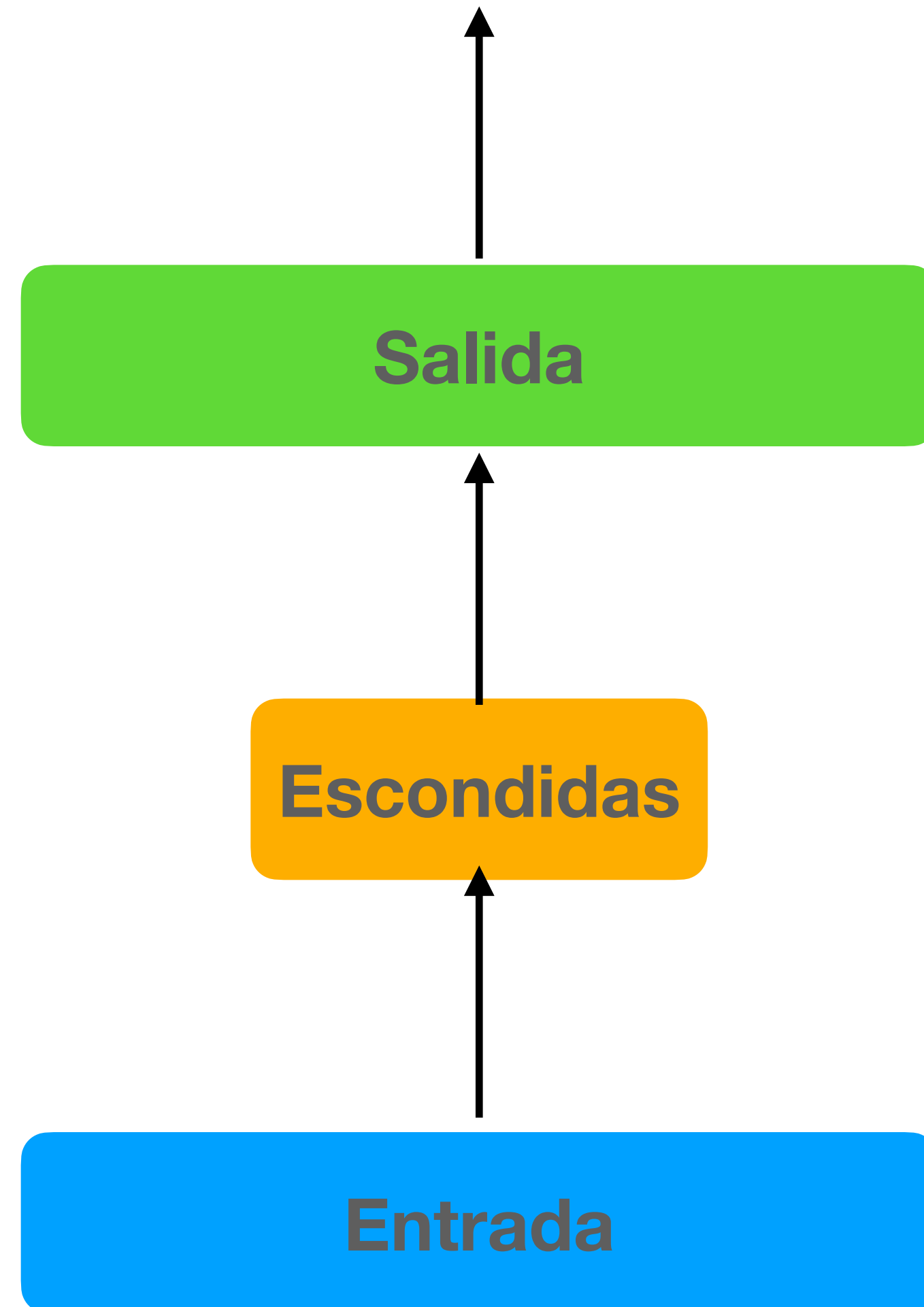
Se reducen las “dimensiones” y a partir de estas se reproduce la entrada

La representación de en medio, intenta mantener la información importante de los dato de entrada

Luego se podrá sacar provecho de la capa más pequeña para extraer ideas “insights” significativas

Autocodificador

Esquemático



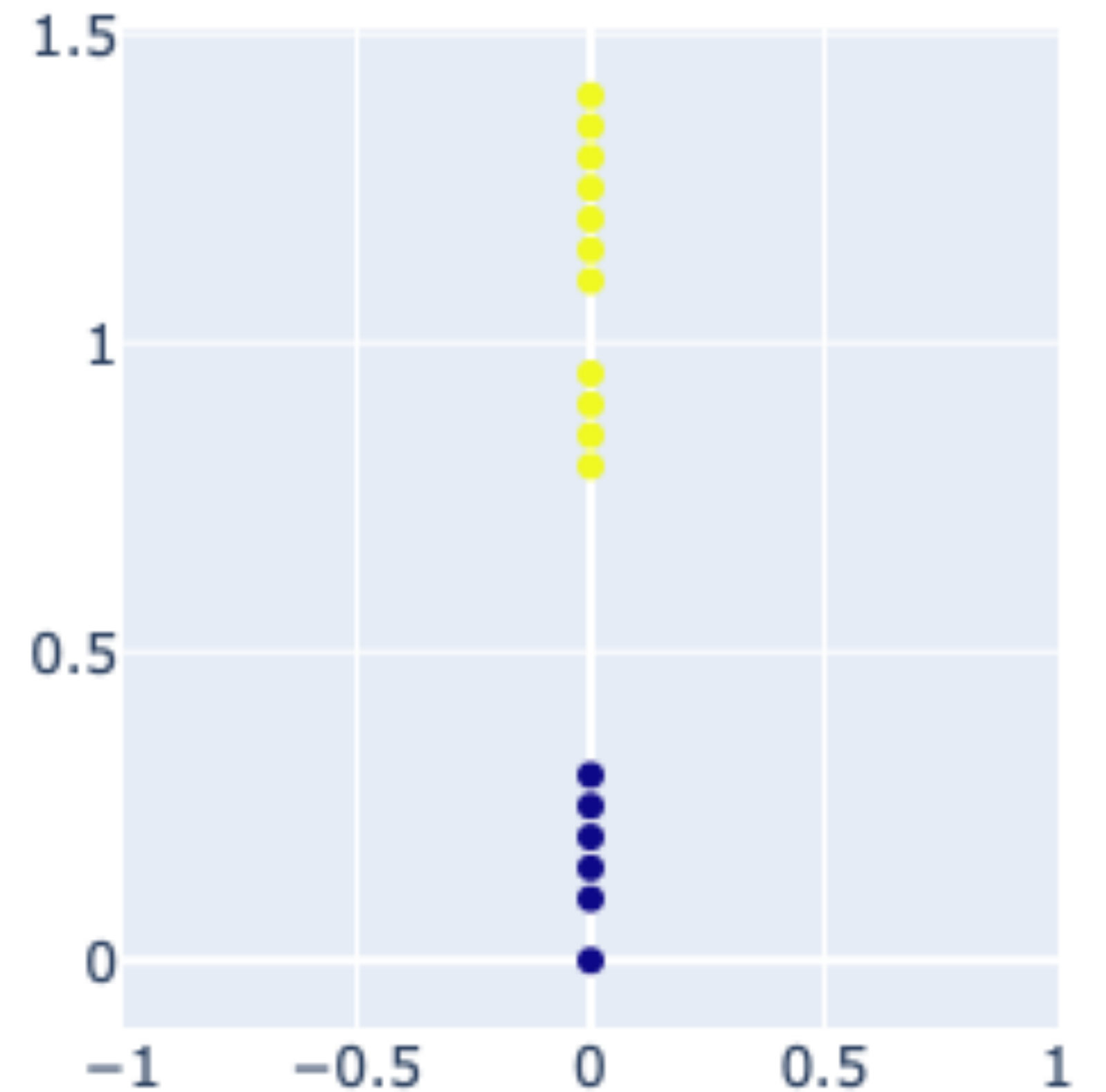
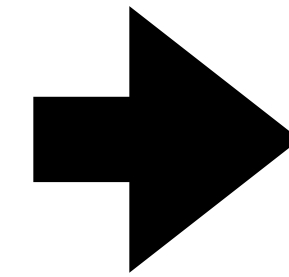
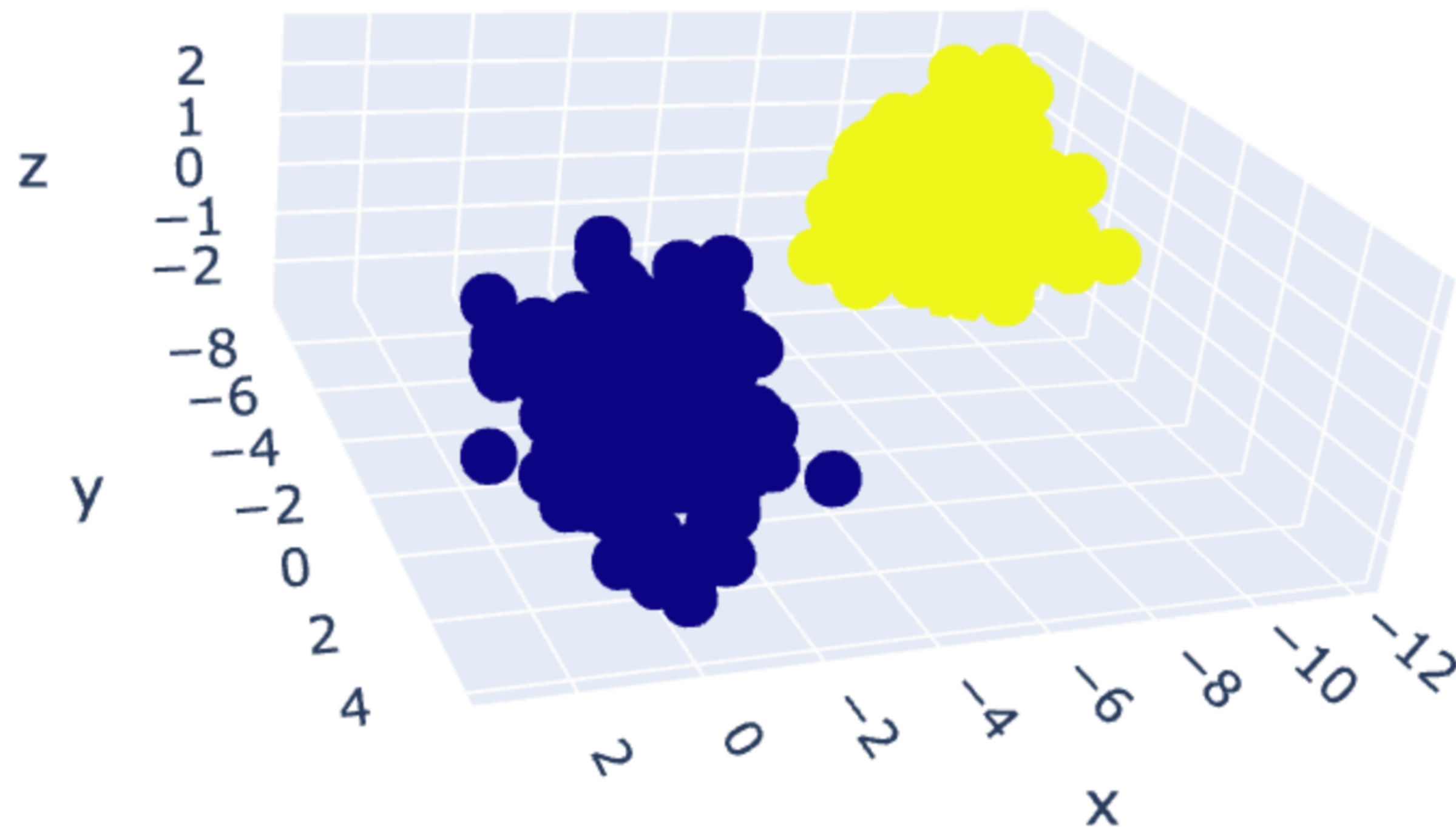
La idea es extremadamente parecida al Análisis de Componentes Principales PCA!

Es importante ver que la capa escondida más pequeña no selecciona ciertas características “features”, y deja otras fuera.

Hace cálculos combinatorios para representar los datos originales con menos “dimensiones”

Autocodificador

Reducción de dimensionalidad



**El anterior ejemplo puede parecer trivial.
Sin embargo si estamos trabajando con
20, 30 o más dimensiones, poder
bajarlo a 3, 2, o 1 dimensión ya permite
visualizarlo gráficamente**

La idea central del autocodificador es que la capa escondida central reduce la dimensionalidad aprendiendo las combinaciones más importantes de las características originales!

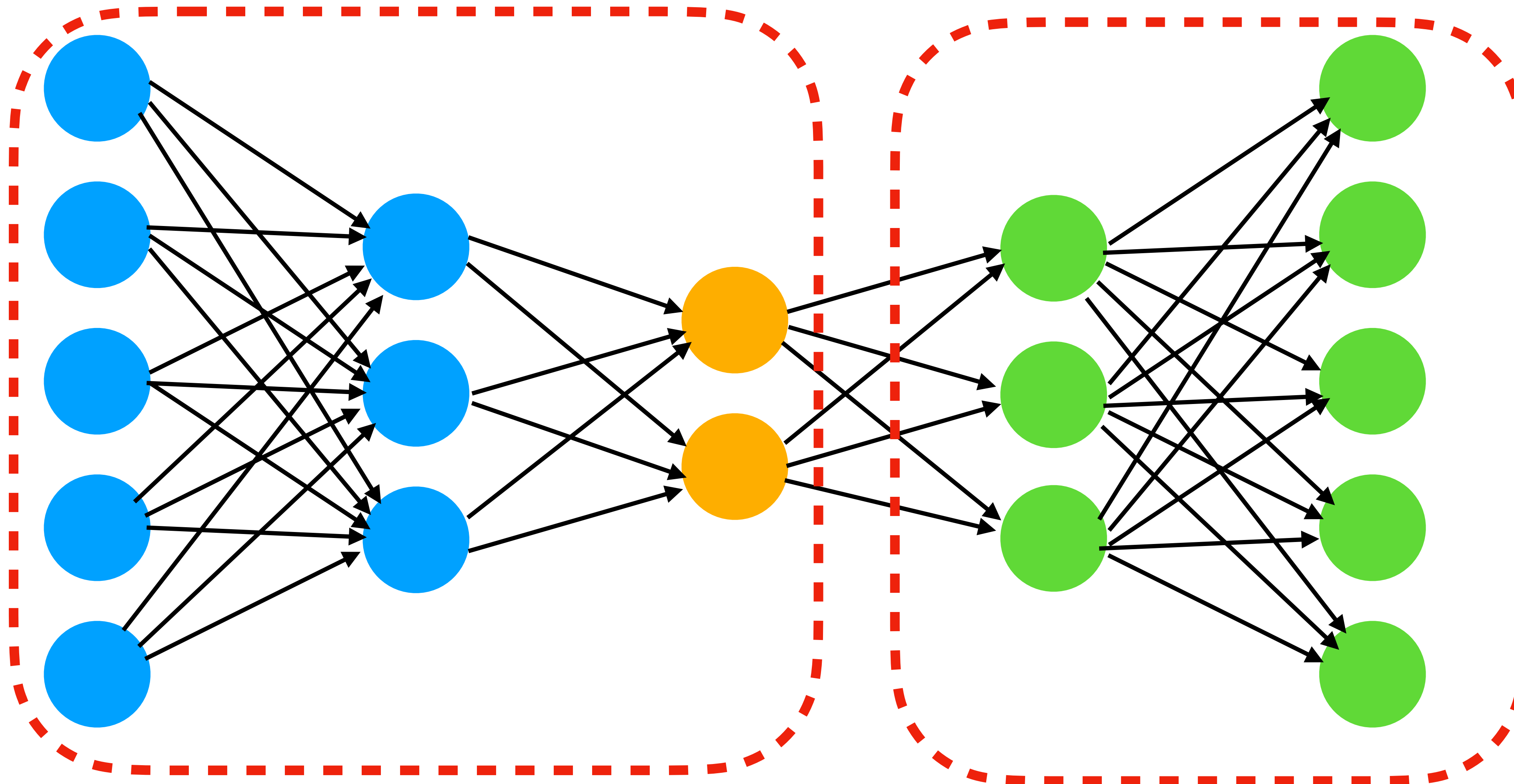
Autocodificadores

Caso 1: Reducción de dimensionalidad

- Se puede pensar en el autocodificador como constituido por dos partes
 - Codificador
 - Decodificador

Autocodificador

Ejemplo



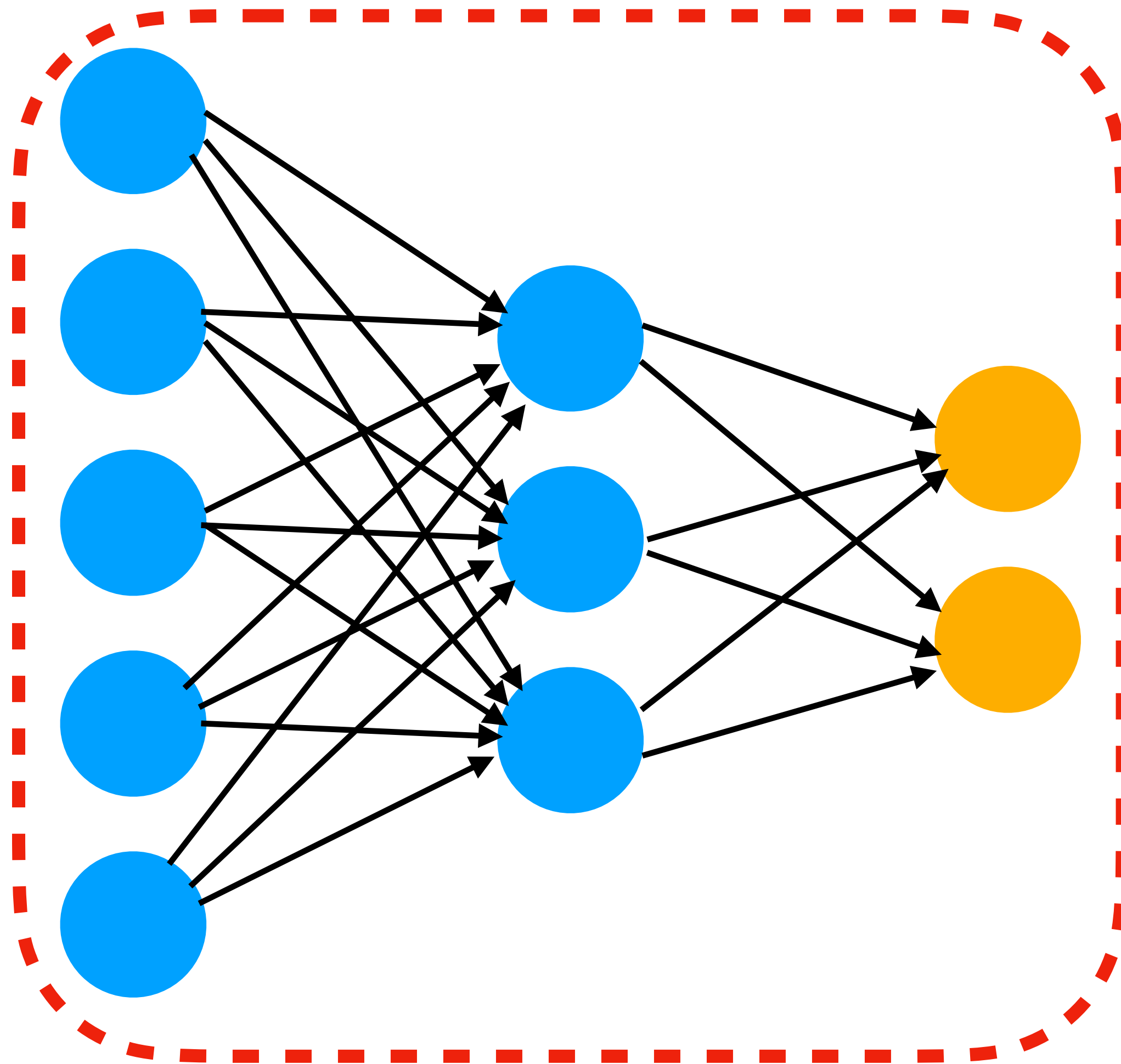
Codificador

Decodificador

Se entrena el autocodificador completo hasta que la capa escondida se vuelve muy buena en extraer la información importante y el decodificador logra producir los mismos datos que la entrada.

Autocodificador

Ejemplo



Codificador

Una vez se entrena el autocodificador, ya podemos utilizar solo la parte del codificador para la reducción de la dimensionalidad de los datos de entrada.

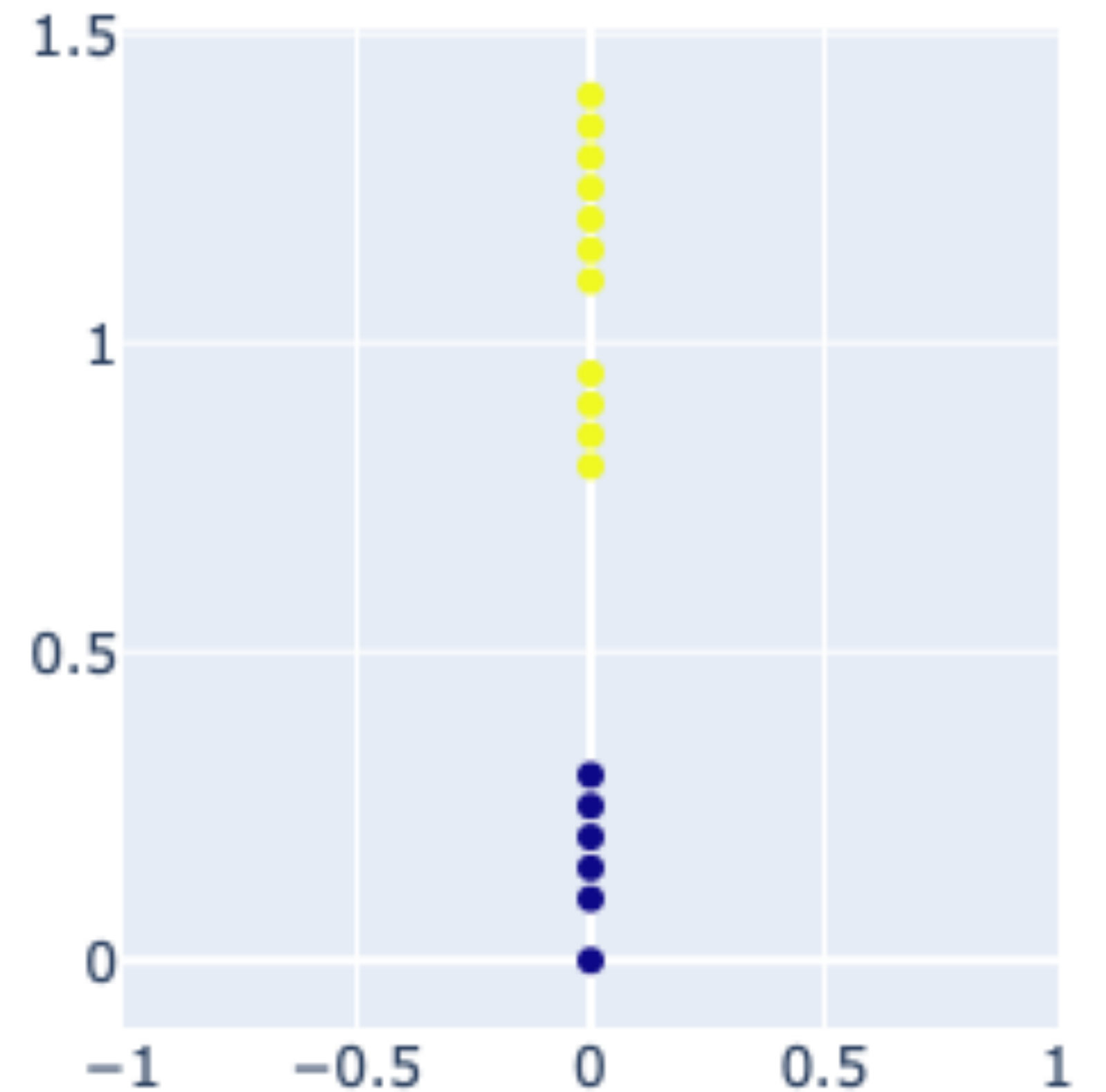
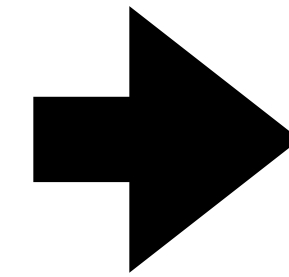
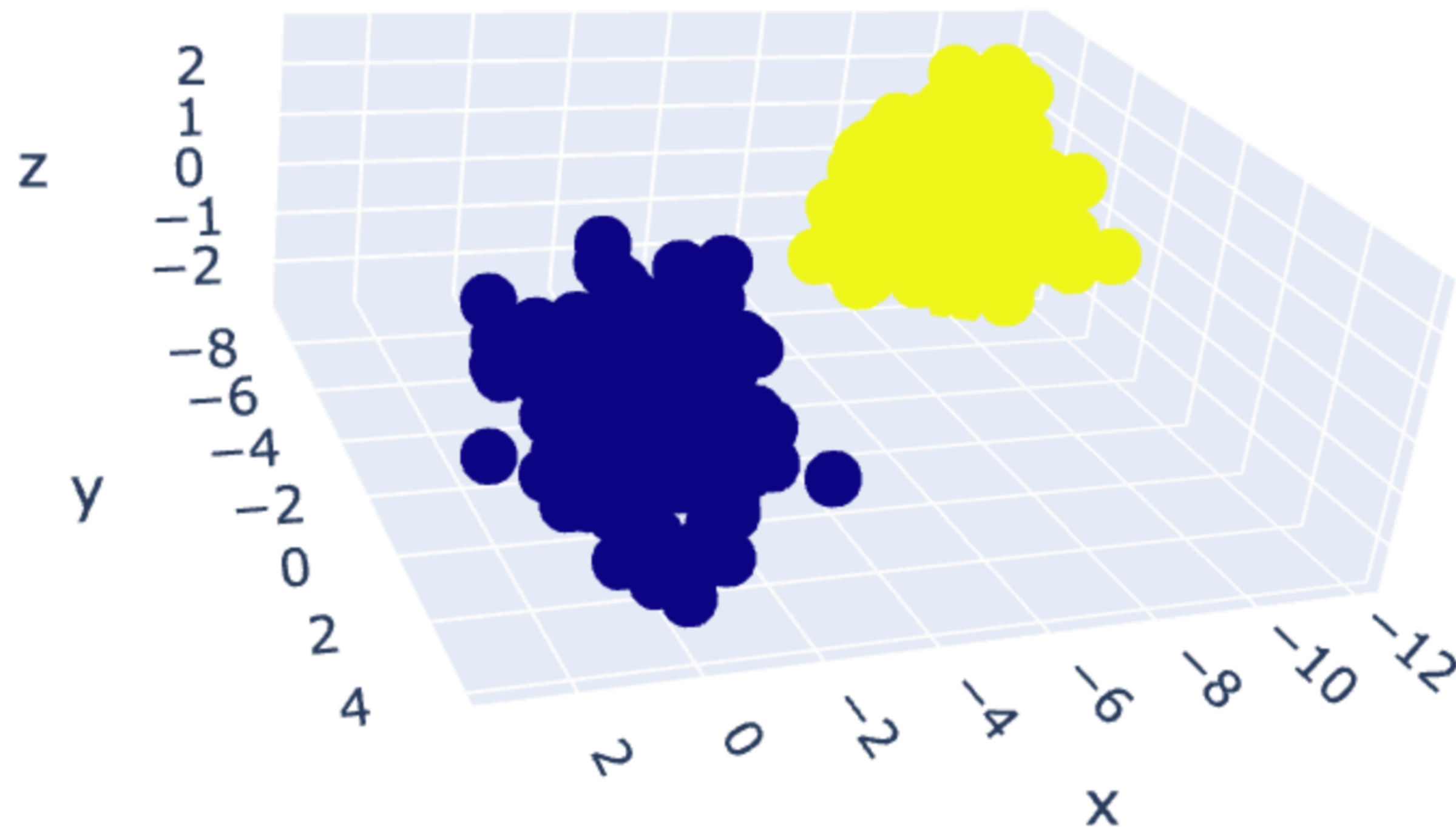
Codificador

¿Qué usos se le puede dar?

- Comprimir datos
- Visualizar datos en menos dimensiones
- Descubrir relaciones subyacentes que no pueden verse claramente con muchas dimensiones

Autocodificador

Reducción de dimensionalidad



Se trabajará este mismo ejemplo
con un Notebook

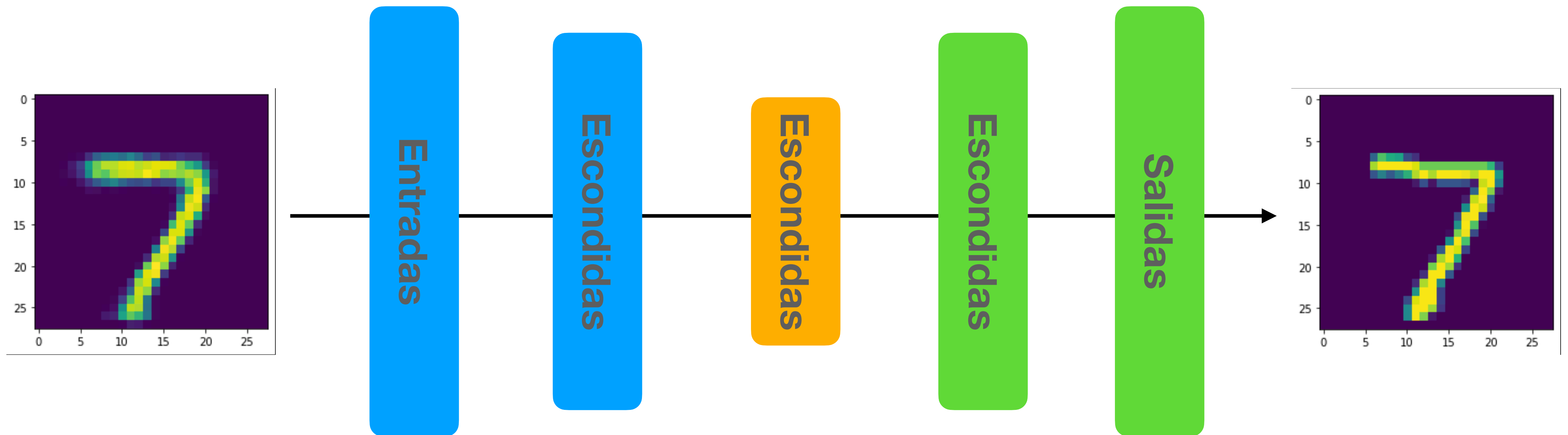
Autocodificadores

Caso 2: Imágenes

- Aún las imágenes más simples conforman conjuntos de datos de alta dimensionalidad
- Las imágenes del MNIST son de 28 X 28 pixeles que equivale a 784 features o dimensiones

Autocodificadores

Imágenes



Autocodificadores

Caso 2: Imágenes

- El autocodificador entiende qué información es importante y cuál no
- Por ejemplo, con las imágenes del MNIST lo más probable es que entienda que la información de las orillas no es importante...recordar que estas imágenes están todas centradas

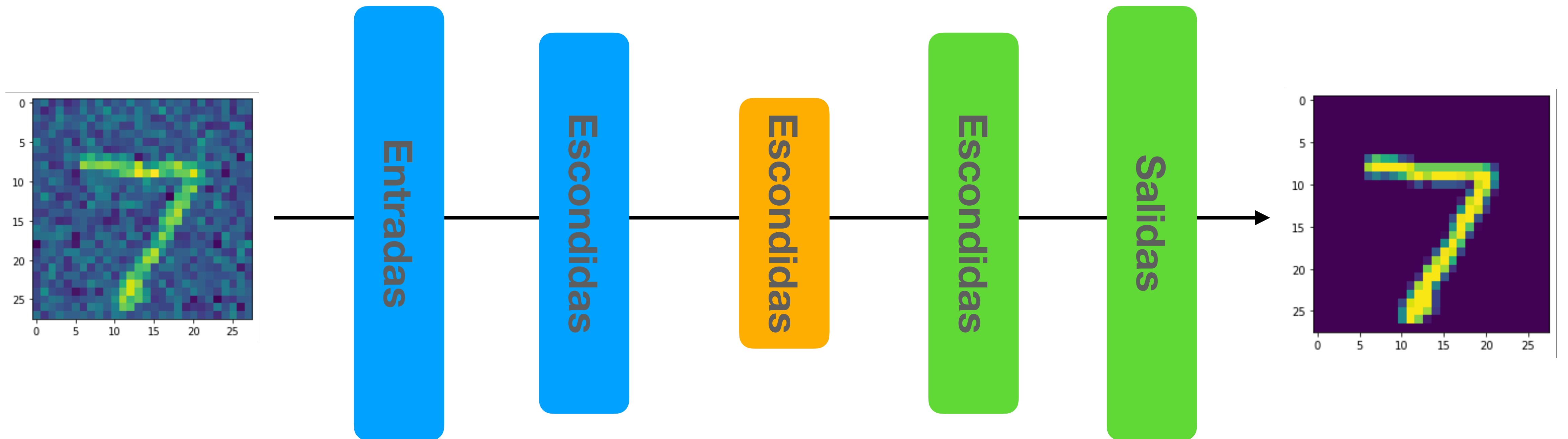
Autocodificadores

Caso 2: Imágenes

- En la vida real, se puede usar esta misma lógica para reducir el ruido ya que el autocodificador entiende qué características son importantes.

Autocodificadores

Imágenes - eliminación de ruido



Autocodificadores

Caso 2: Imágenes

- Se verán dos autocodificadores
 - El primero, para recrear la imagen original
 - El segundo, que tome una imagen con ruido para que la “limpie”

Se trabajará este mismo ejemplo
con un Notebook