

# Minería de Datos

## IIC2433

Feature Embeddings  
Vicente Domínguez

# ¿Qué veremos esta clase?

- Qué es un *Feature Embedding*
- Auto-encoders

# Features

- ¿Cómo obtenemos atributos/features de la información?
- ¿Cuál es la mejor forma de obtenerlos?

# Features

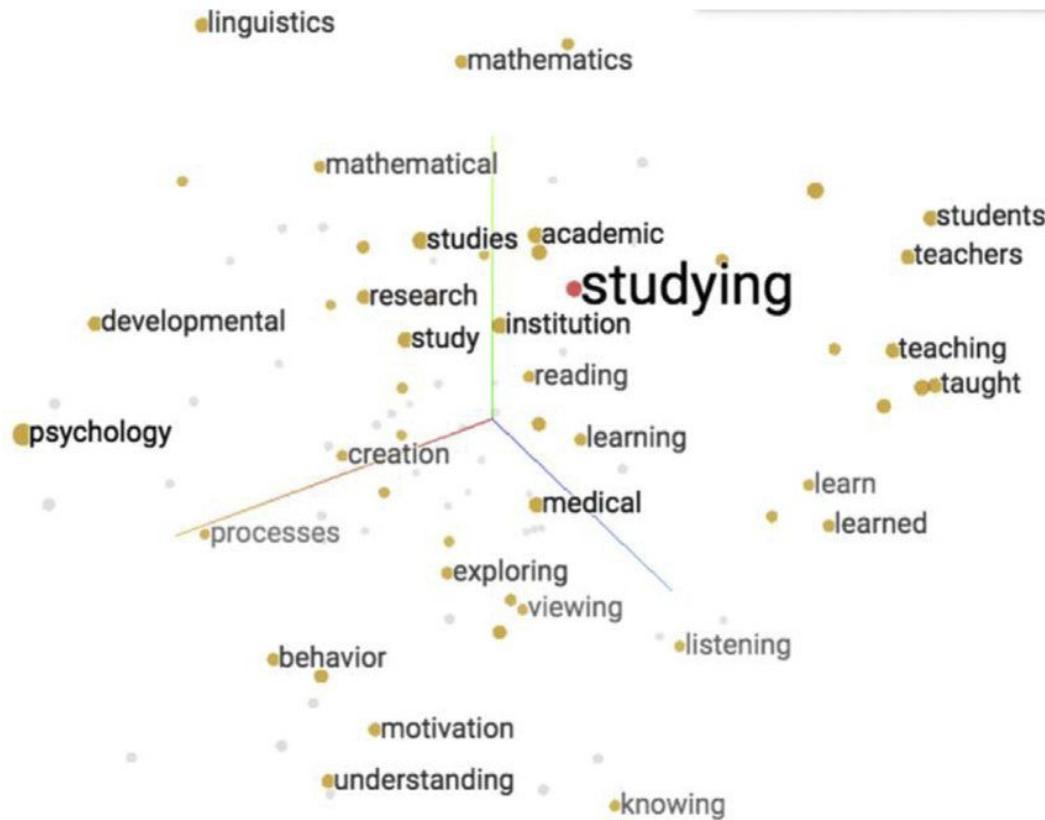
- Nos gustaría obtener atributos que sean ricos y descriptivos
- En general se puede anotar datos descriptivos de la información
- ¿Cómo sabemos que es buena la información?

# Feature engineering

- Durante años de han ido desarrollando nuevos métodos para obtener buenos features, en distintas áreas (imágenes, texto, audio, videos, transacciones)
- ¿Para qué nos gustaría tener buenos atributos?

# Feature embeddings

- Nos gustaría que los atributos, finalmente generen un **espacio** en donde los datos que comparten características similares estén cerca, y los que son diferentes también se encuentren alejados.
- Si es que tenemos features ricos en información, casi cualquier modelo de predicción debería funcionar bien.



Kastrati, Zenun & Kurti, Arianit & Imran, Ali. (2020). WET: Word Embedding-Topic distribution vectors for MOOC video lectures dataset. Data in Brief. 28. 105090. 10.1016/j.dib.2019.105090.

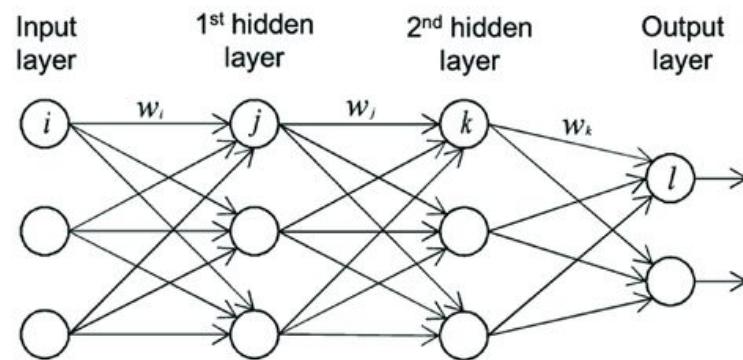
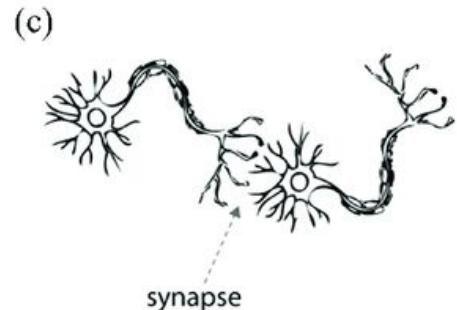
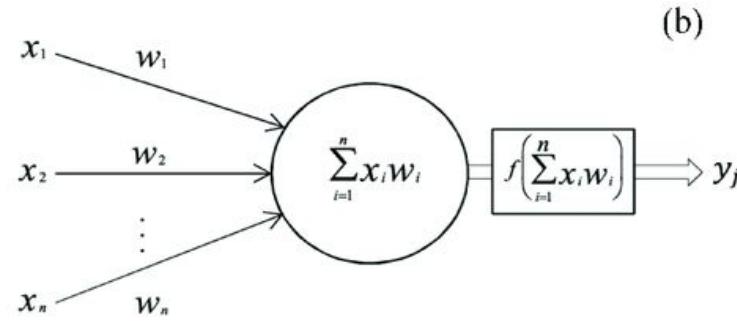
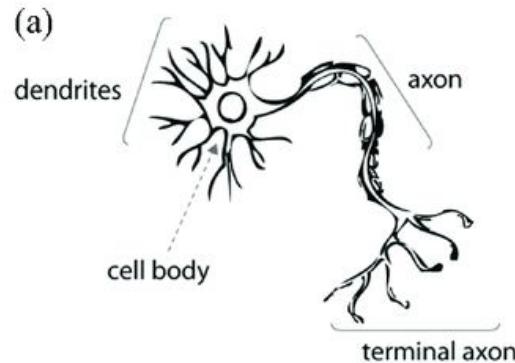
# ¿Cómo logramos esto?

- Primero, debemos llevar a un mismo espacio distintos tipos de atributos.
  - Categóricos
  - Ordinales
  - Numéricos
  - etc
- Los atributos, pueden vivir en rangos de valores completamente diferentes.
- ¿Y si no tengo atributos? Por ejemplo, tengo documentos, imágenes o sonidos ¿Qué puedo hacer?

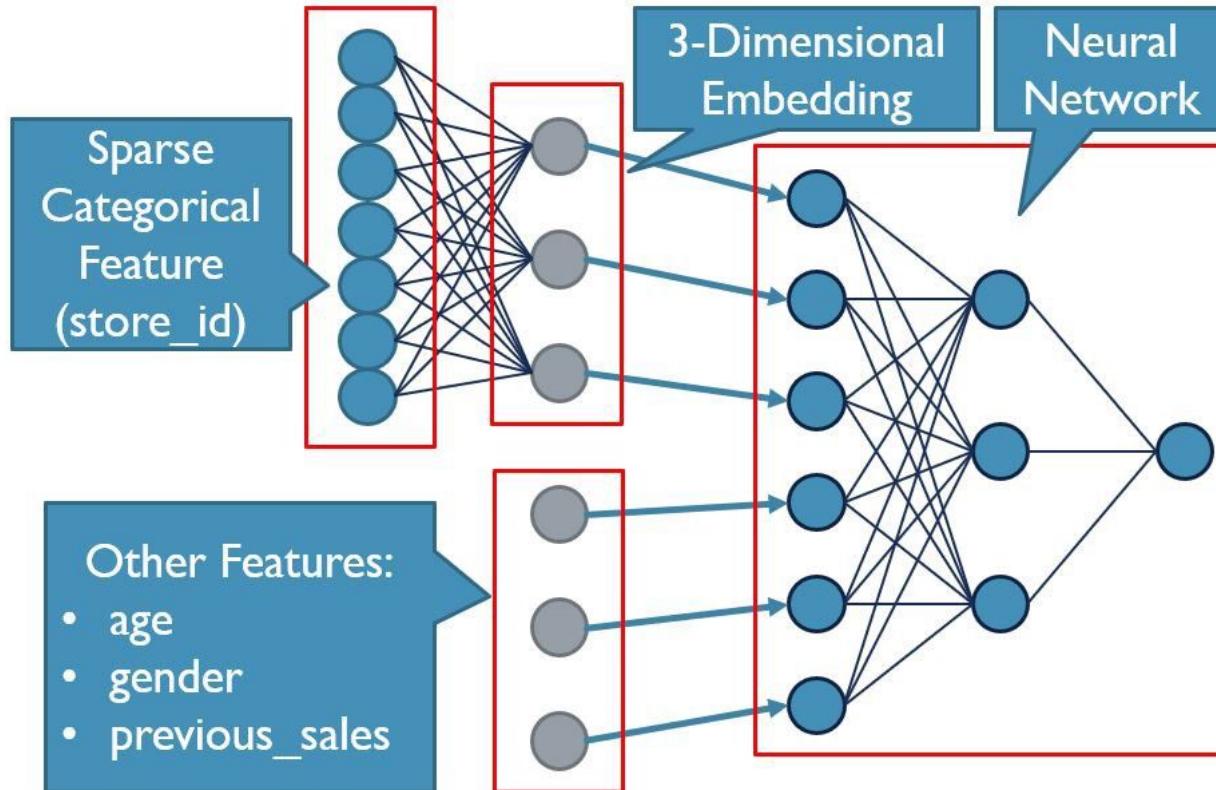
# Redes neuronales

- Modelo computacional que trata de replicar la estructura en que las neuronas cerebrales aprenden en los humanos
- Sus precursores, Yann LeCun, Geoffrey Hinton y Yoshua Bengio ganaron el Premio Turing (Premio Nobel en Computación) por sus aportes al desarrollo del modelo.

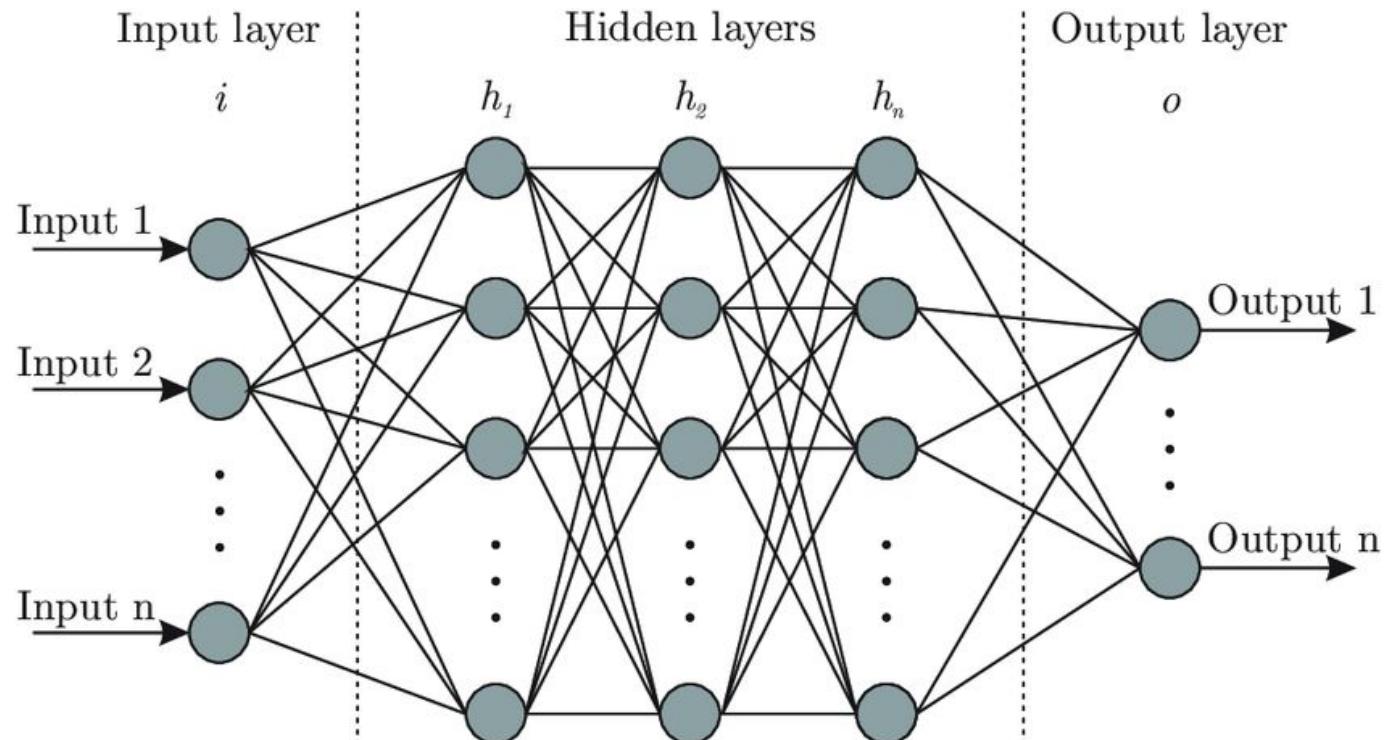
# Redes neuronales



# Redes neuronales

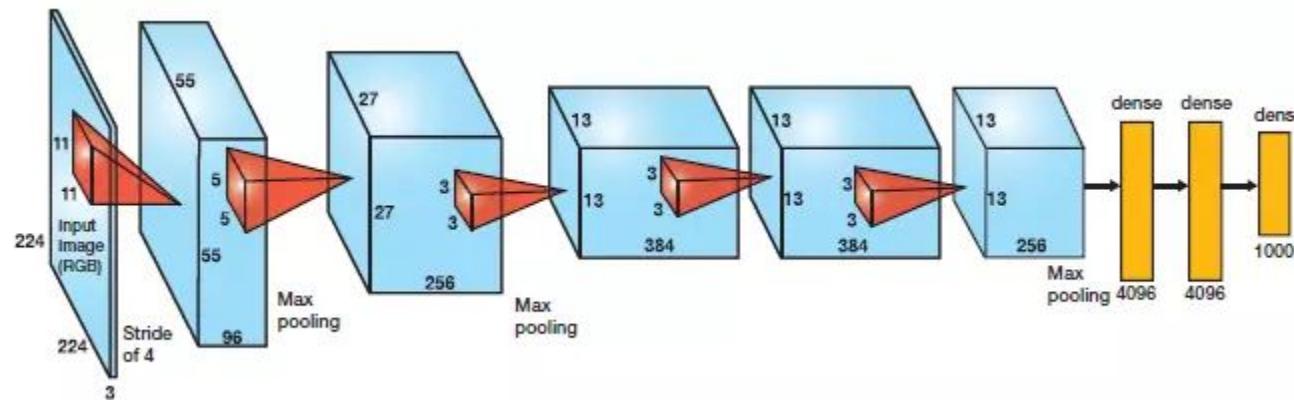


# Redes neuronales



# Redes neuronales

- Uno de sus principales logros y por los cuales empezó a ganar interés el hecho fue vencer a los modelos de clasificación de imágenes en competencias de clasificación.
- El modelo que destacó fue la Alexnet en el Imagenet Challenge 2012



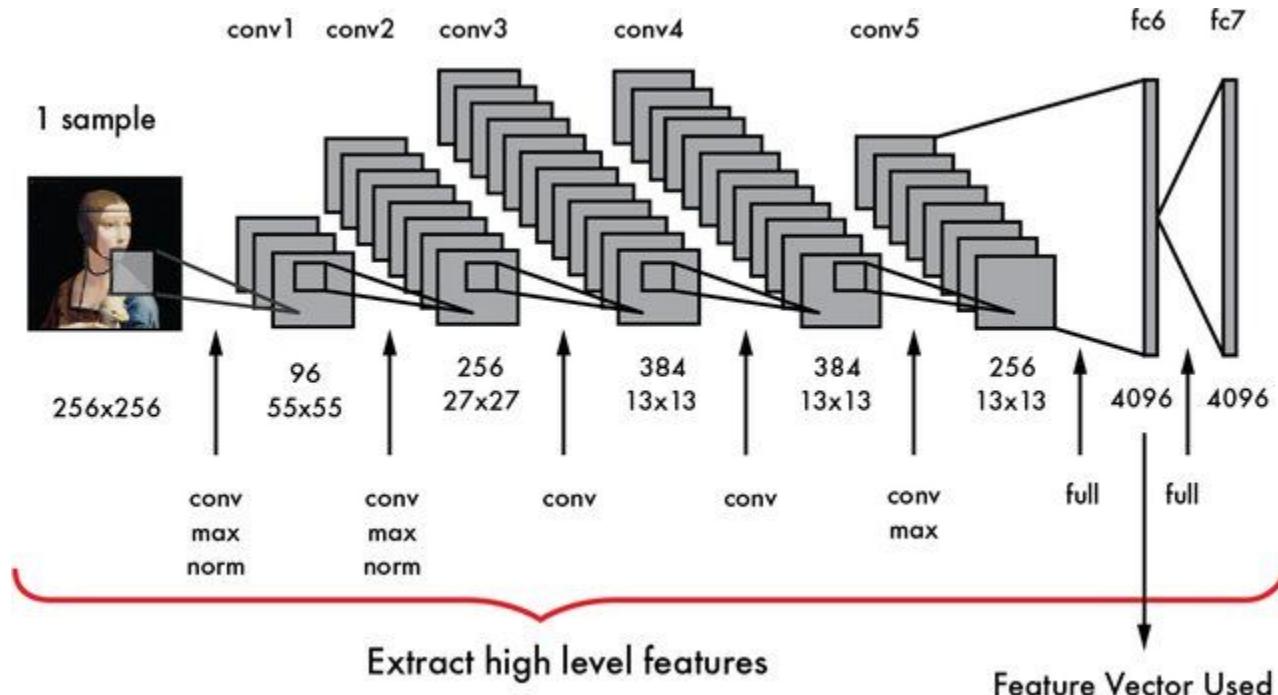
# Redes neuronales

## Transfer Learning

- Una forma muy buena de obtener un feature embedding es tomando el espacio de representación que ya creó otra red ya entrenada.
- Por ejemplo, es posible obtener los pesos que genera una capa anterior a la de clasificación y utilizarla como features

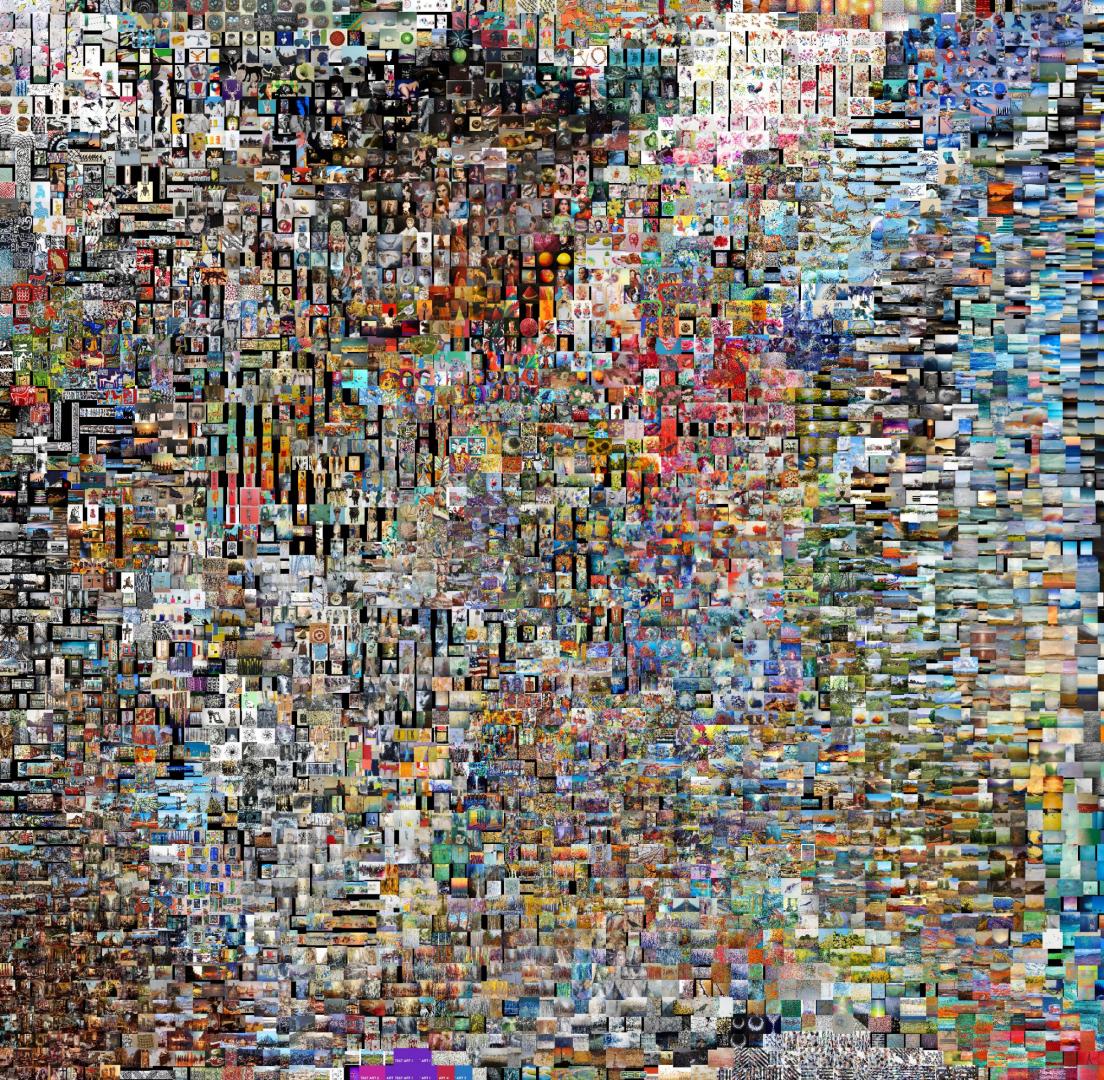
# Redes neuronales

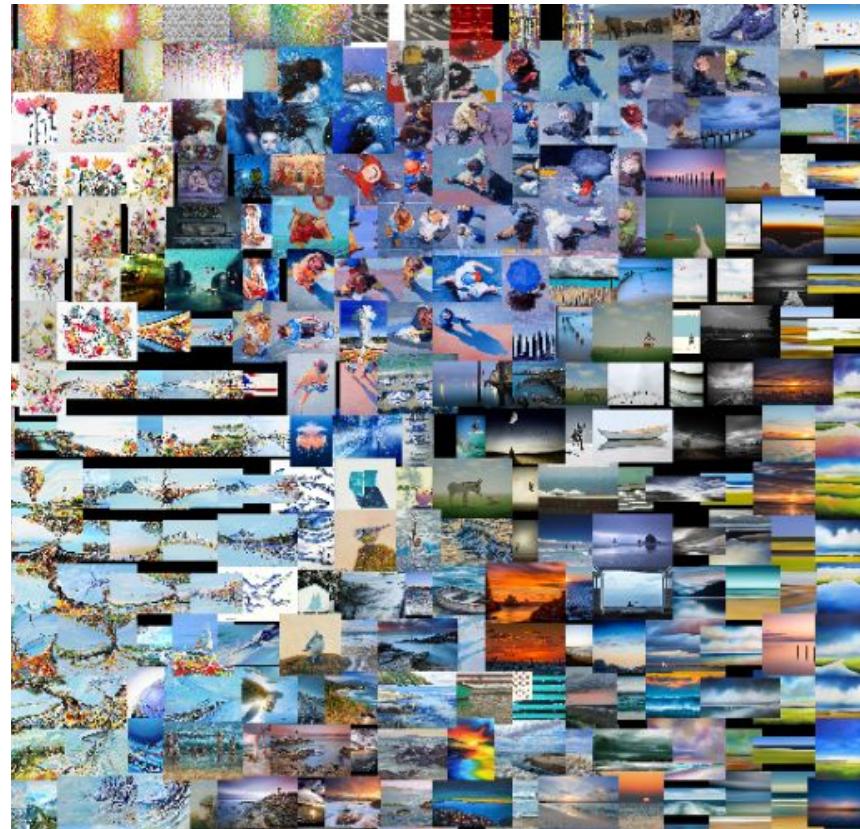
## Transfer Learning



Messina, P., Dominguez, V., Parra, D., Trattner, C., & Soto, A. (2019). Content-based artwork recommendation: integrating painting metadata with neural and manually-engineered visual features. *User Modeling and User-Adapted Interaction*, 29(2), 251-290.

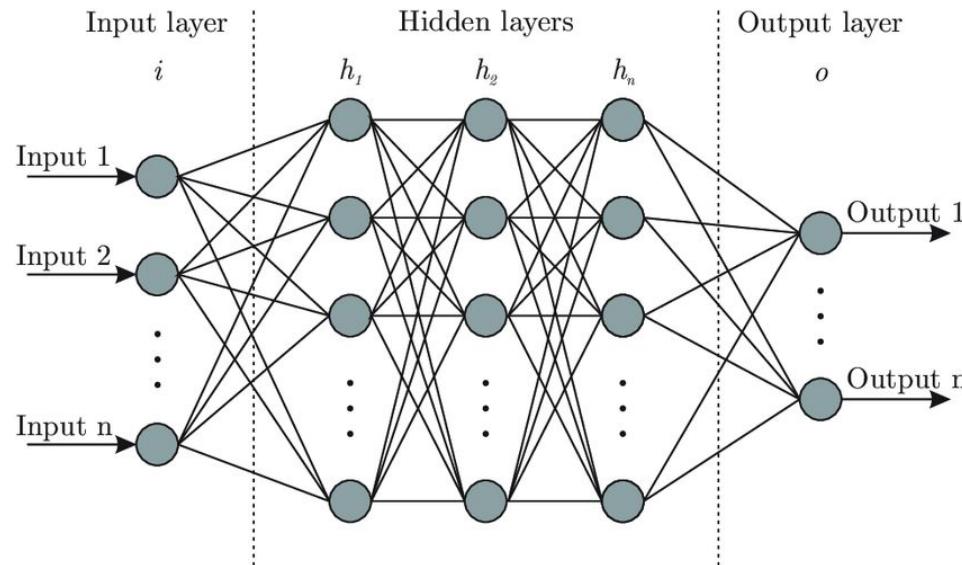






# Autoencoders

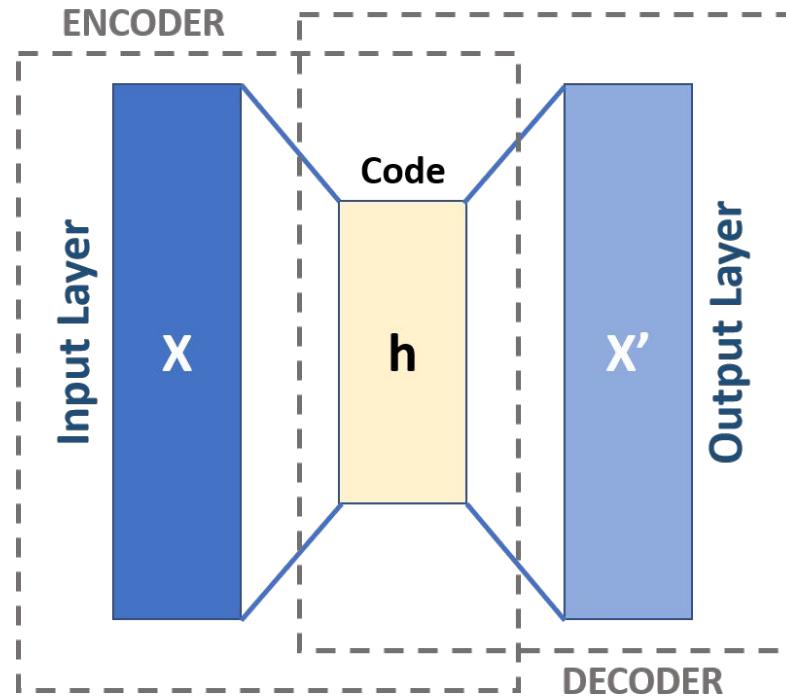
- Ahora ¿cómo podría generar mi propia representación, sin utilizar la aprendida por otro modelo?
- ¿Qué parte de la red creen que es la más importante para este objetivo?



# Autoencoders

- Modelo enfocado en obtener una codificación de los datos.
- Buscan representar un espacio donde se reduzca la dimensionalidad y se obtenga una buena representación de las características de los datos.
- Al tratar de encontrar las características que más aportan a la representación del dato, también logran reducir el ruido.

# Autoencoders



[https://en.wikipedia.org/wiki/File:Autoencoder\\_schema.png](https://en.wikipedia.org/wiki/File:Autoencoder_schema.png)

# Autoencoders

- Más detalladamente, los autoencoders se componen de dos partes: Encoder y Decoder, también conocidas como  $\phi$  y  $\psi$  (phi y psi)
- Lo que se busca es encontrar dos funciones tales que ocurra lo siguiente

$$\phi : \mathcal{X} \rightarrow \mathcal{F}$$

$$\psi : \mathcal{F} \rightarrow \mathcal{X}$$

$$\phi, \psi = \arg \min_{\phi, \psi} \|\mathcal{X} - (\psi \circ \phi)\mathcal{X}\|^2$$

# Autoencoders

- En su versión más simple, se toma un input  $\mathbf{x} \in \mathbb{R}^d = \mathcal{X}$  y lo mapea a  $\mathbf{h} \in \mathbb{R}^p = \mathcal{F}$  tal que:

$$\mathbf{h} = \sigma(\mathbf{W}\mathbf{x} + \mathbf{b})$$

- Con la imagen  $\mathbf{h}$  también conocida como la codificación, variables latentes o espacio latente.  $\sigma$  es la función de activación element-wise,  $\mathbf{W}$  la matriz de pesos y  $\mathbf{b}$  un vector de bias.
- Estos pesos son encontrados con el método backpropagation.

# Autoencoders

- En la fase de decodificación, se busca reconstruir desde la imagen  $\mathbf{h}$  a un  $\mathbf{x}'$  tal que:

$$\mathbf{x}' = \sigma'(\mathbf{W}'\mathbf{h} + \mathbf{b}')$$

- Donde  $\sigma'$ ,  $\mathbf{W}'$  y  $\mathbf{b}'$  pudiendo no estar relacionados con sus correspondientes para el codificador.

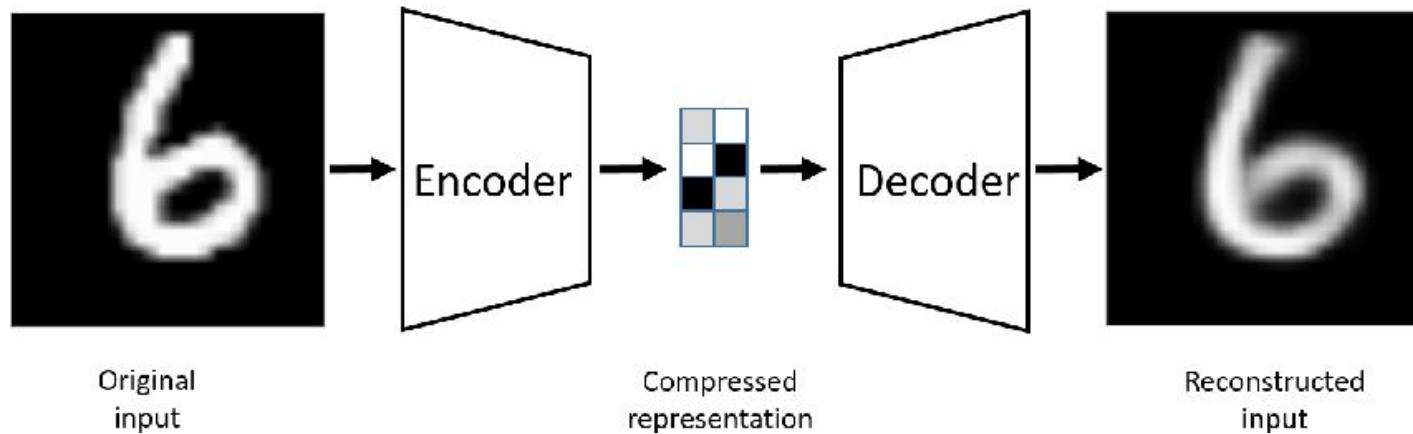
# Autoencoders

- Finalmente, el autoencoder busca minimizar el error de la reconstrucción. De esta forma ajusta los pesos para lograr un Feature Space.

$$\mathcal{L}(\mathbf{x}, \mathbf{x}') = \|\mathbf{x} - \mathbf{x}'\|^2$$

$$\mathcal{L}(\mathbf{x}, \mathbf{x}') = \|\mathbf{x} - \sigma'(\mathbf{W}'(\sigma(\mathbf{W}\mathbf{x} + \mathbf{b})) + \mathbf{b}')\|^2$$

# Autoencoders



Bank, D., Koenigstein, N., & Giryes, R. (2020). Autoencoders. *ArXiv*, *abs/2003.05991*.

# Autoencoders

- Los autoencoders no solo pueden reducir, sino tener una representación más grande.
- Han salido muchas nuevas versiones (variacionales, reductores de ruido, etc)
- Son la base de los modelos generativos.

# Autoencoders

