

# Aprendizaje profundo

## PRESENTACIÓN DEL CURSO

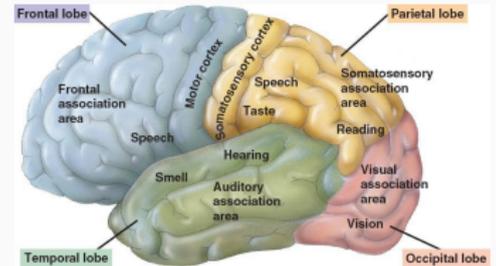
---

Gibran Fuentes-Pineda

Agosto 2019

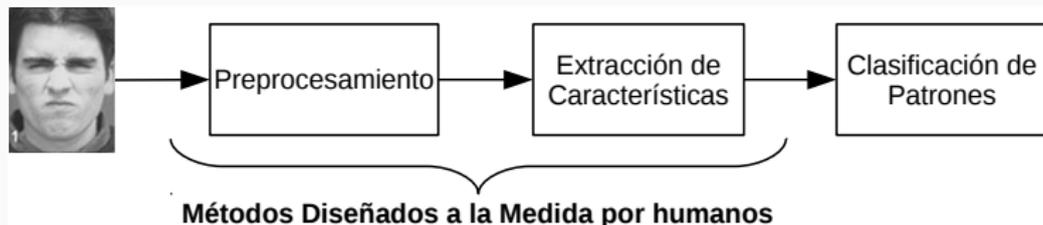
# La hipótesis del algoritmo único (Mountcastle)

- La corteza cerebral parece homogénea
  - **Mismo algoritmo** en muchas partes
- Pruebas de re-cableado neuronal
  - Corteza auditiva aprende a ver (Roe et al. 1992)
  - Corteza somatosensorial aprende (Metin & Frost 1989)

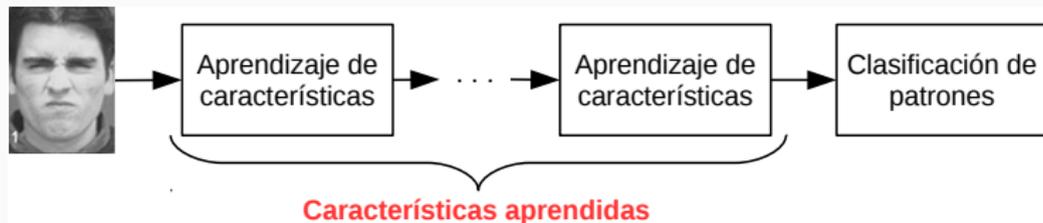


# Aprendizaje de características

## Con extracción de características

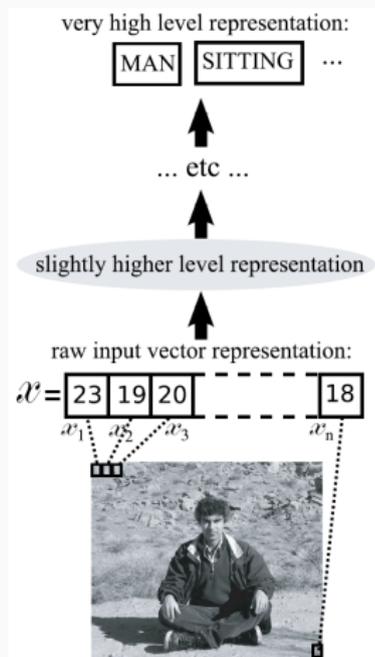


## Con aprendizaje de características

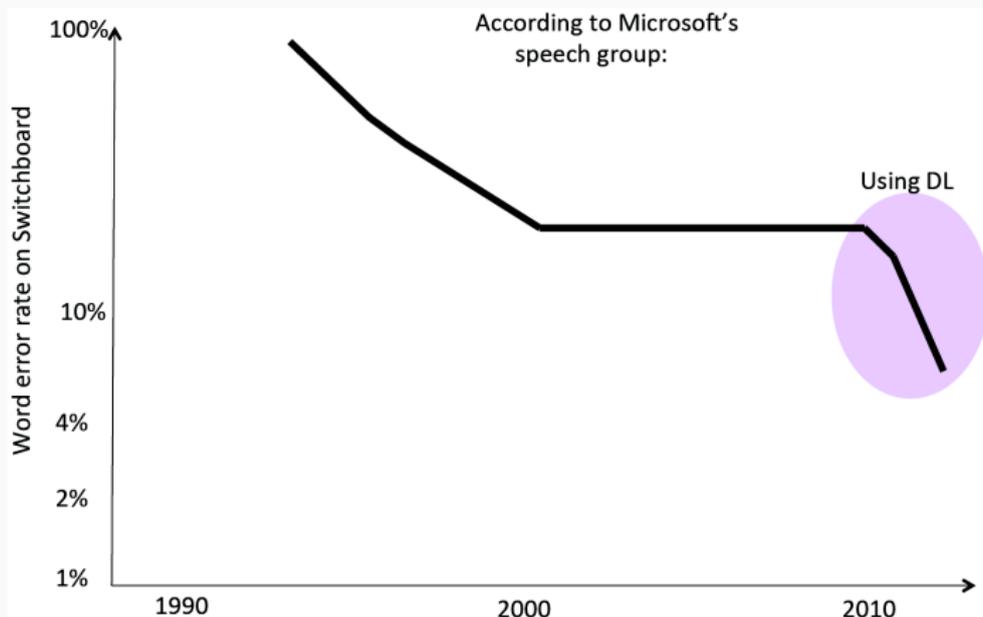


# Jerarquías de Características

- Características se **aprenden** directamente de la entrada
  - Niveles más altos se forman a partir de niveles más bajos
- Aproximación de funciones complejas a partir de composición de múltiples niveles (**arquitectura profunda**)
- Se pueden **compartir** características aprendidas en varias tareas relacionadas



# Rendimiento en reconocimiento de habla

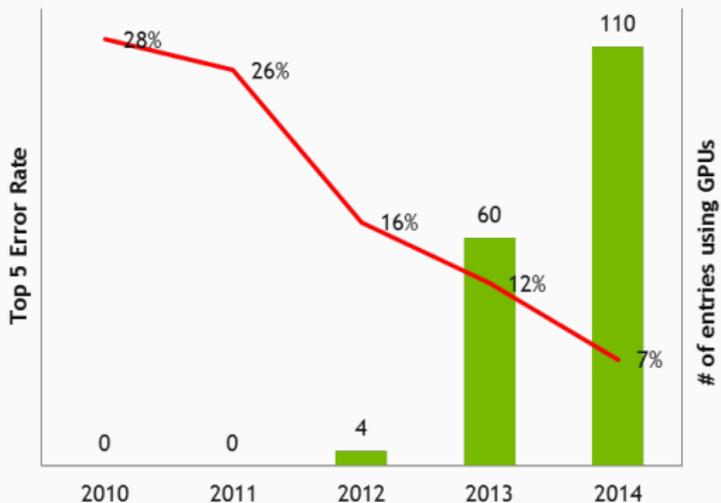


# Mejor en algunas tareas que los humanos

- **Verificación de rostros** (Sun et al. 2014):
  - Base de datos: Learning Faces in the Wild
  - Tarea: Verificar si un rostro es de cierta persona
  - CNN: 99.15 %
  - Humano: 99.20 %
- **Reconocimiento de escritura china** (Ciresan 2013):
  - Base de datos: CASIA-HWDB & CASIA-OLHWDB
  - Tarea: Clasificar caracteres chinos (3755 clases)
  - CNN: 4.21 % error
  - Humano: 3.87 % error
- **Reconocimiento de señales** (Ciresan et al. 2012)
  - Base de datos: German Traffic Sign Recognition Benchmark
  - Tarea: Clasificar 50000 señales de tráfico en 43 clases
  - CNN: 99.46 %
  - Humano : 98.84 %

# Importancia de los GPUs

IMAGENET



Tomada de <https://devblogs.nvidia.com/>

`nvidia-ibm-cloud-support-imagenet-large-scale-visual-recognition-challenge/`

1. Perceptrón multicapa
2. Redes convolucionales
3. Redes recurrentes
4. Regularización y entrenamiento de redes profundas
5. Mecanismos de atención y memoria externa
6. Auto-codificadores
7. Redes generativas

# Introducción

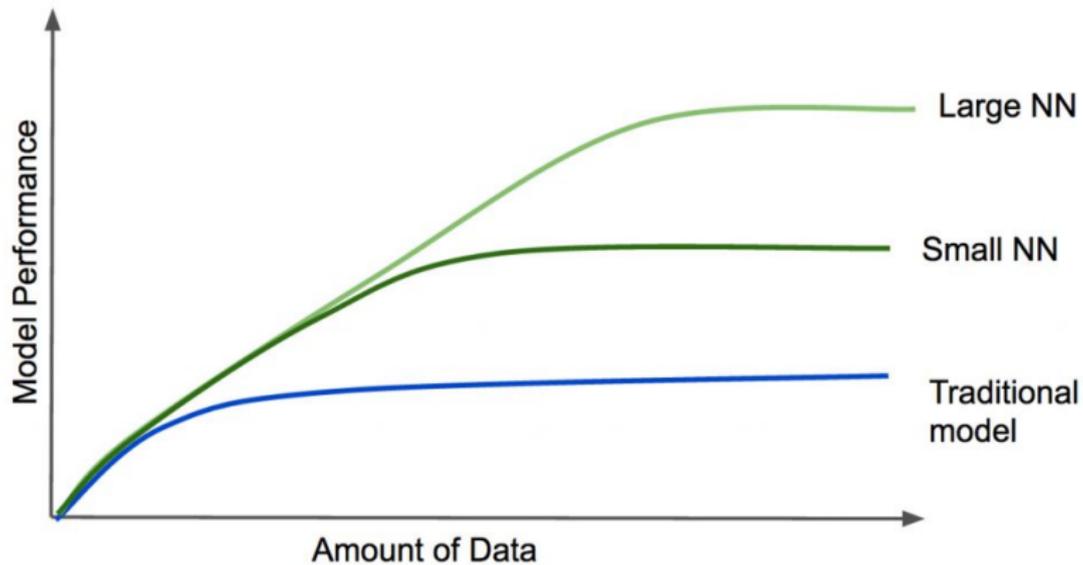
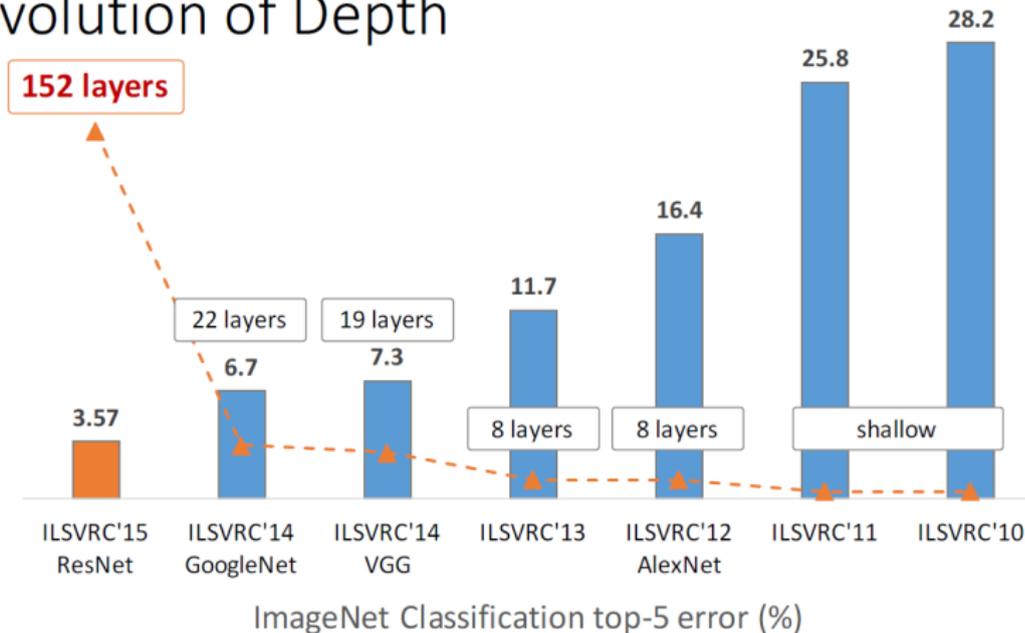


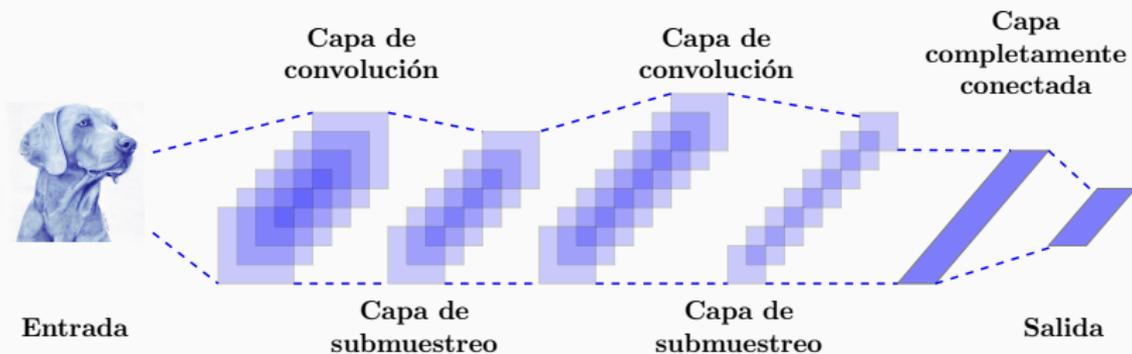
Figura de Krizhevsky et al. 2012

## Revolution of Depth



# Redes convolucionales

- Propuesta originalmente por LeCun et al. 1998 para reconocimiento de dígitos



# Redes profundas en clasificación de imágenes

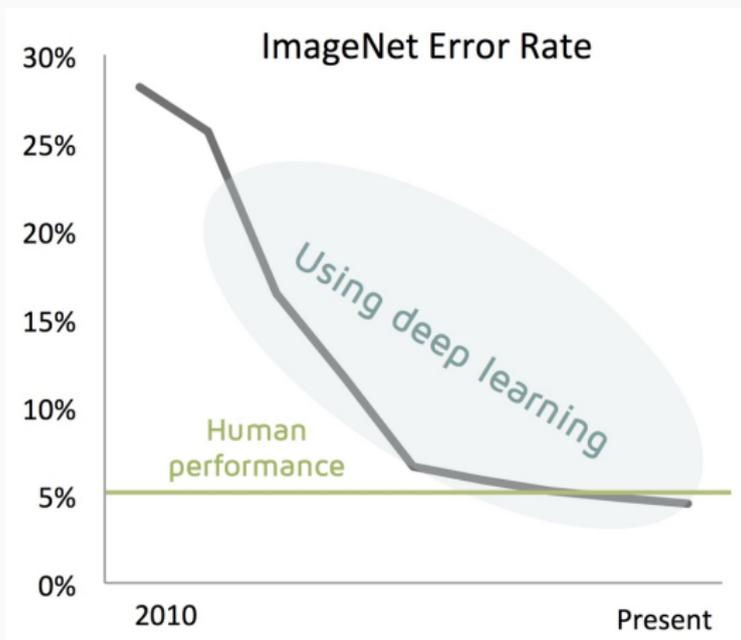
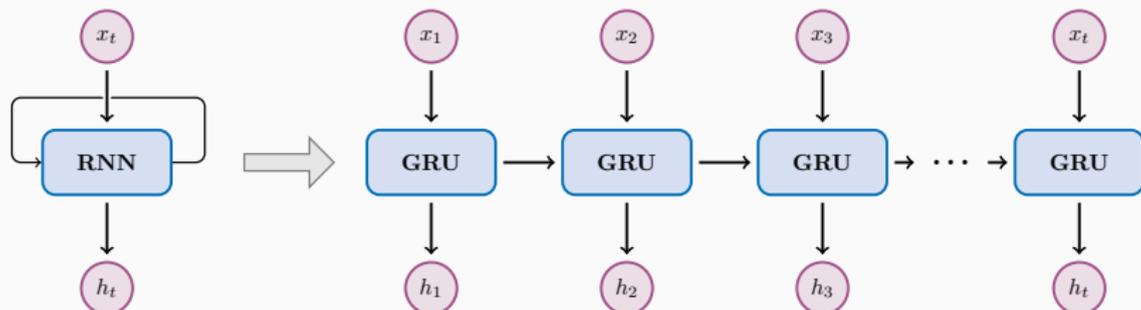


Figura de <https://www.nervanasys.com/deep-learning-and-the-need-for-unified-tools/>

- Una de las variantes más utilizadas es la *Long-Short Term Memory* propuesta por Hochreiter y Schmidhuber en 1997.



# Generación de relatos cortos a partir de imágenes



*Today was graduation day. The students were excited. My parents were so happy. He was very happy to be graduating. Everyone was so proud of him.*



*We went to a fancy restaurant. I made a lot of food for everyone. The food was delicious. The main course was presented beautifully. The dessert was even better.*

# Entrenamiento de redes profundas

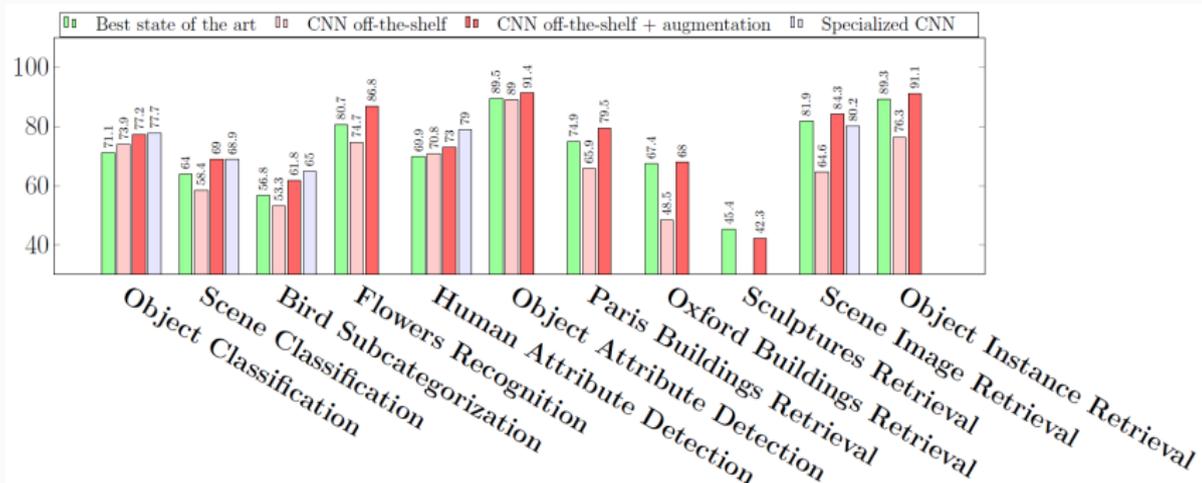
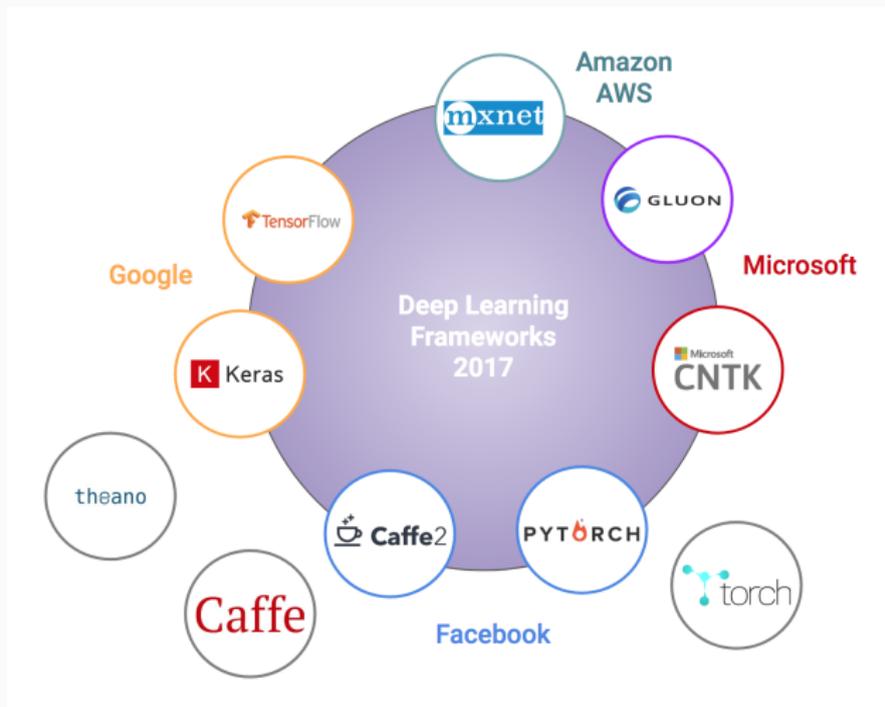


Figura tomada de Razavian et al. 2014



# Mecanismos de atención

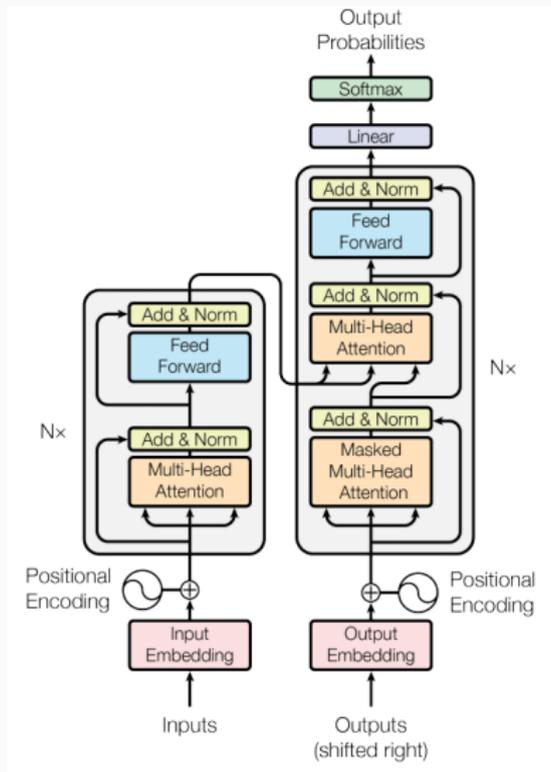


Figura tomada de Vaswani et al. *Attention is all you need*, 2017

# Traducción automática

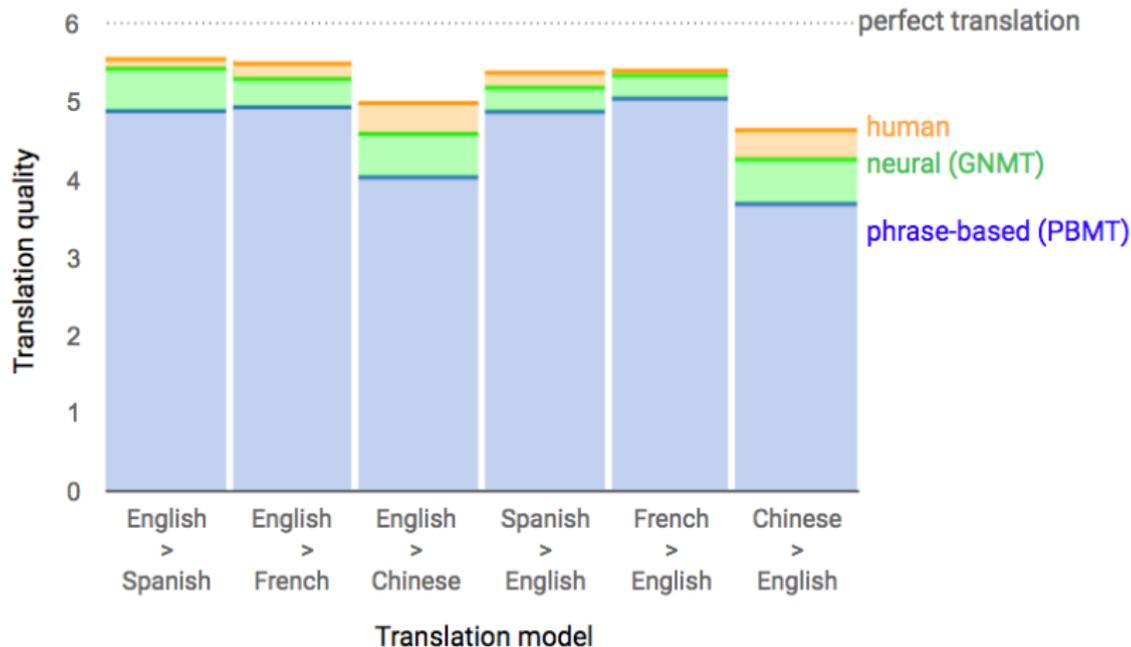
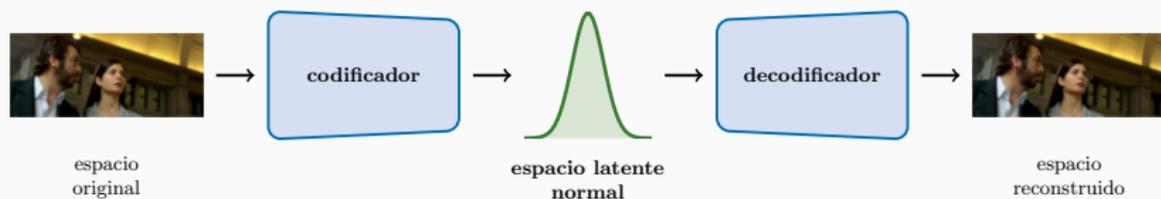


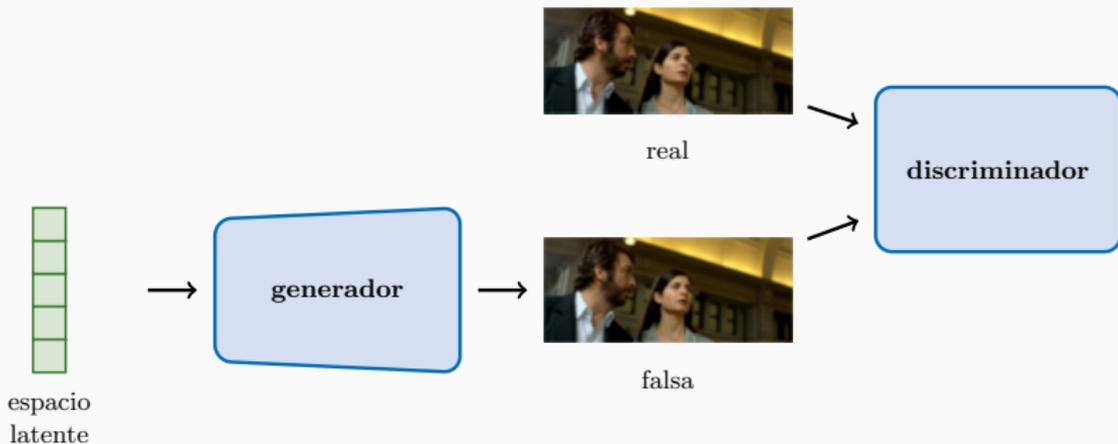
Figura tomada de <https://ai.googleblog.com/2016/09/a-neural-network-for-machine.html>

- Autocodificador generativo propuesto por Kingma y Welling en 2013



# Redes generativas antagónicas

- Modelo generativo propuesto por Goodfellow et al. en 2014



# Generación de imágenes a partir de texto

this small bird has a pink breast and crown, and black primaries and secondaries.



this magnificent fellow is almost all black with a red crest, and white cheek patch.



the flower has petals that are bright pinkish purple with white stigma



this white and yellow flower have thin white petals and a round yellow stamen



Figura tomada de Reed et al. *Generative Adversarial Text to Image Synthesis*, 2016

# Generación de rostros



Figura tomada de Ian Goodfellow (publicado por @goodfellow.ian).

# Generalización

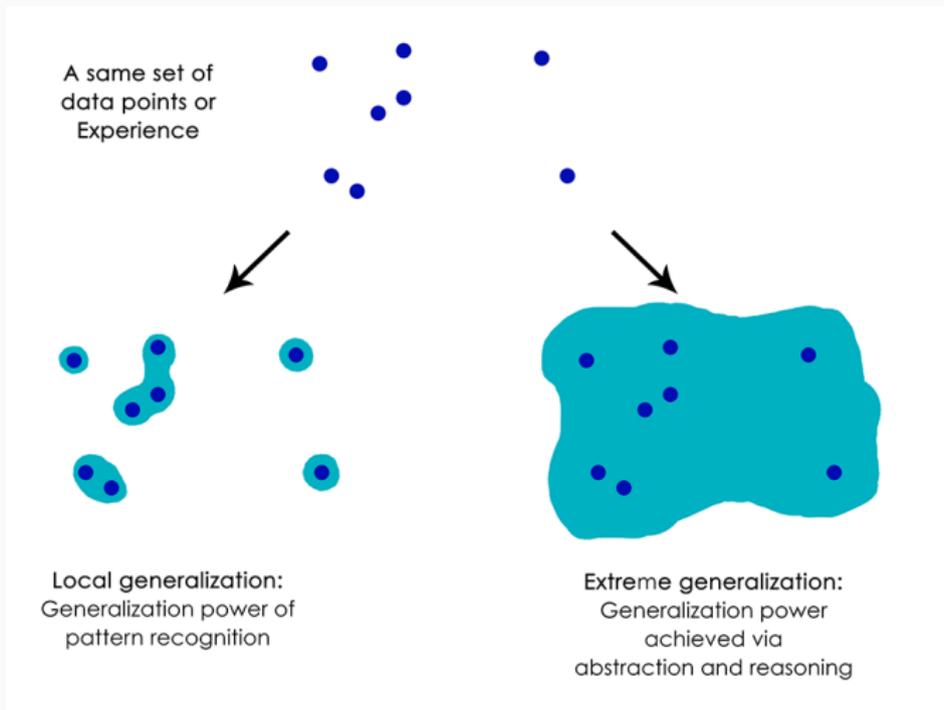
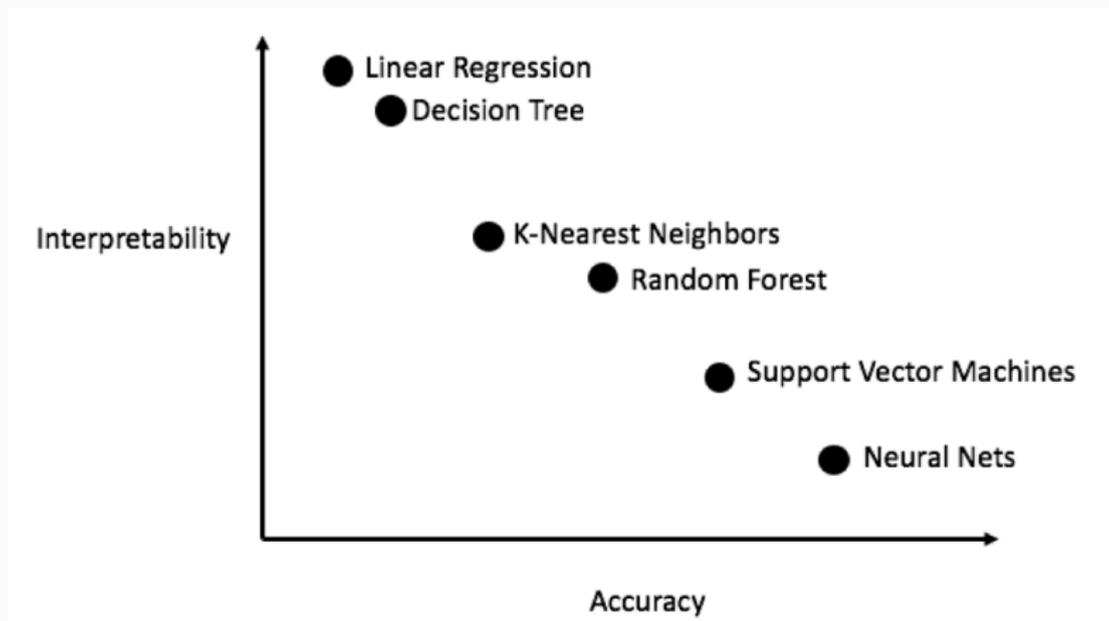


Figura por Chollet, F. 2017. Tomada de <https://blog.keras.io/the-limitations-of-deep-learning.html>

# Modelos poco interpretables

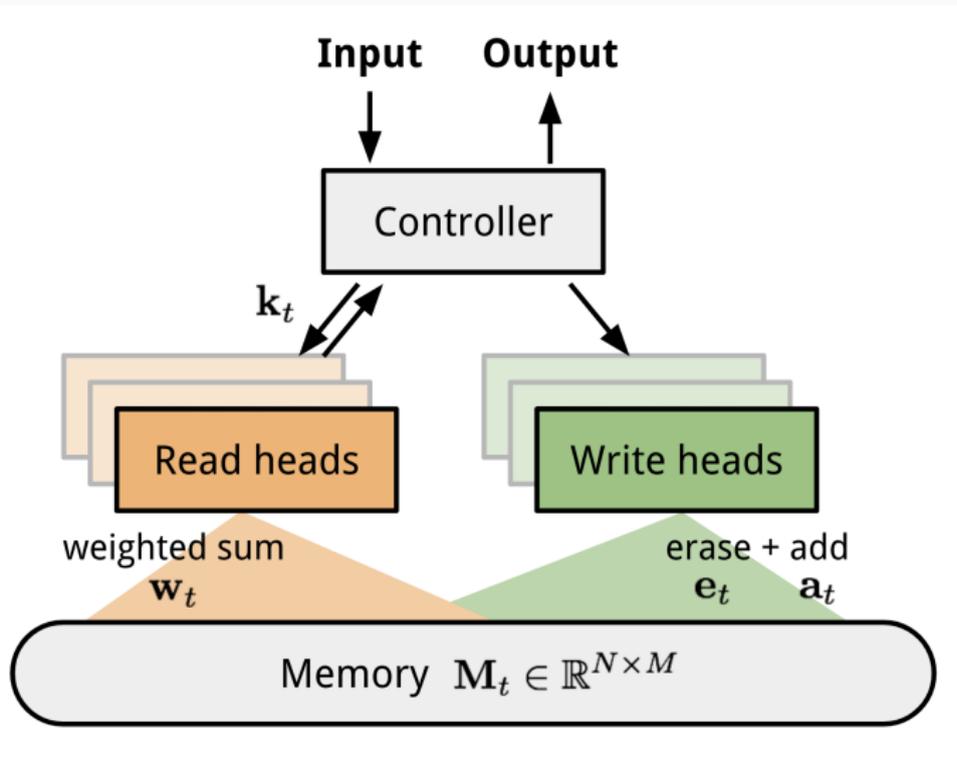


- ¿Qué complejidad necesito para realizar una tarea dada?

- ¿Qué complejidad necesito para realizar una tarea dada?
- ¿Por qué una arquitectura profunda puede ser más efectiva que una superficial?

- ¿Qué complejidad necesito para realizar una tarea dada?
- ¿Por qué una arquitectura profunda puede ser más efectiva que una superficial?
- Aprendizaje requiere minimización de función altamente no convexa, ¿por qué mínimos locales funcionan?

# Memoria externa



## Ejemplos antagónicos

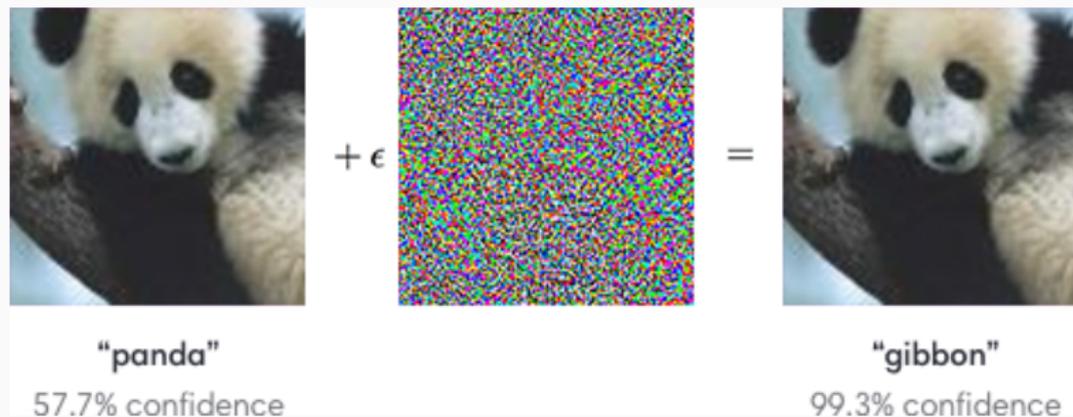


Figura de <https://blog.openai.com/adversarial-example-research/>

# Información del curso

- Horario
  - Martes y jueves de 13:00 a 14:30 hrs.  
Salón 303, anexo del IIMAS.
- Contacto
  - Profesor: Gibran Fuentes Pineda  
Oficina: IIMAS 4to piso, cubículo 417  
Correo: gibranfp@unam.mx
  - Ayudante: Blanca Vázquez Gómez
- Página del curso
  - [http://turing.iimas.unam.mx/~gibranfp/cursos/aprendizaje\\_profundo/](http://turing.iimas.unam.mx/~gibranfp/cursos/aprendizaje_profundo/)

- Exposiciones: 50 %
  - Analizar artículos importantes sobre arquitecturas y/o aplicaciones novedosas y relevantes
- Proyecto: 50 %
  - Reproducir y replicar resultados de artículos
  - Aplicar a otras tareas/condiciones, realizando un nuevo análisis o modificar la arquitectura