

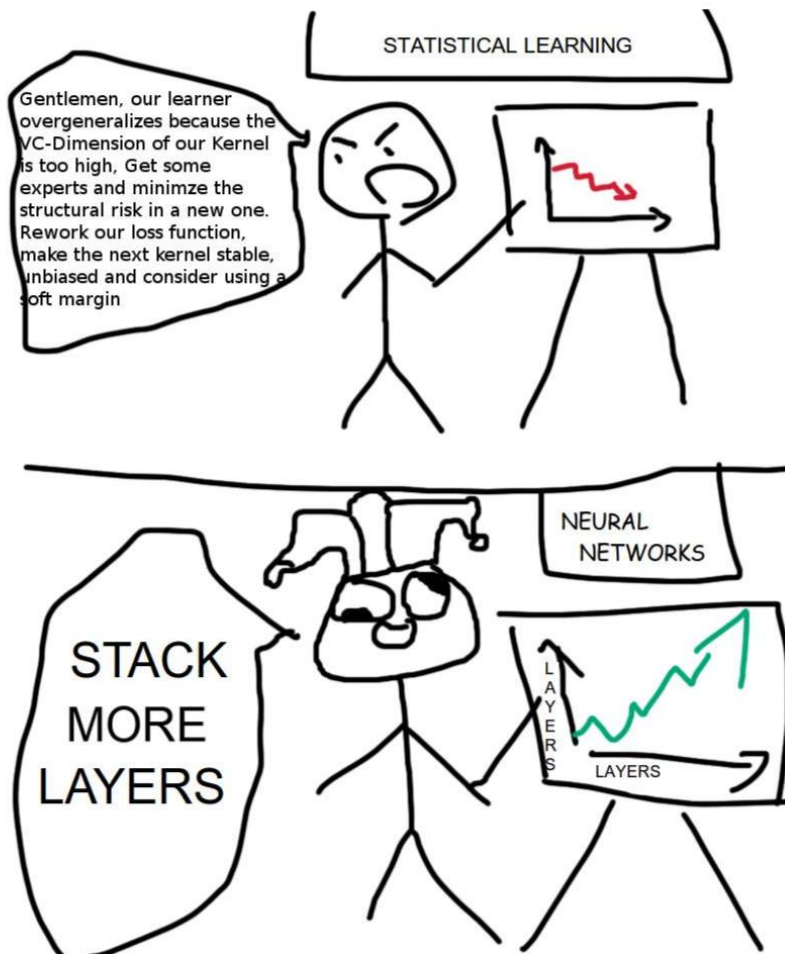


Deep Learning con Pytorch

Juan Pablo Morales
@juanpamf

¿Por qué el Deep
Learning es más
poderoso que el ML
tradicional?

Going deeper: Más capas

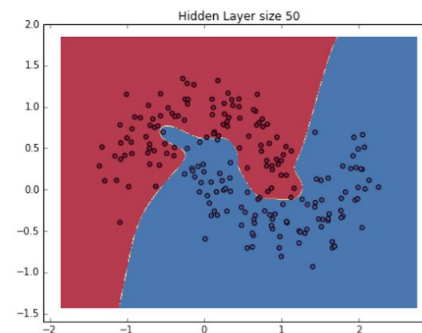
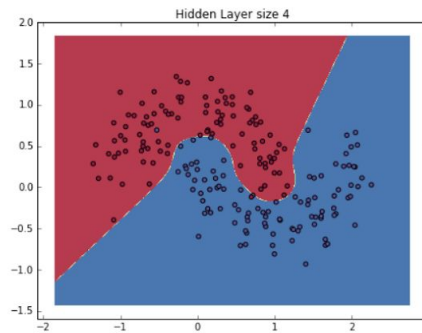
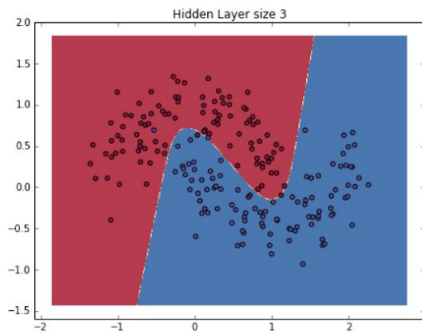
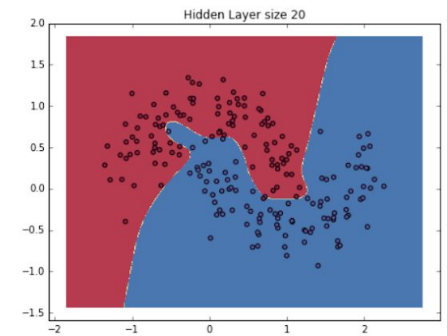
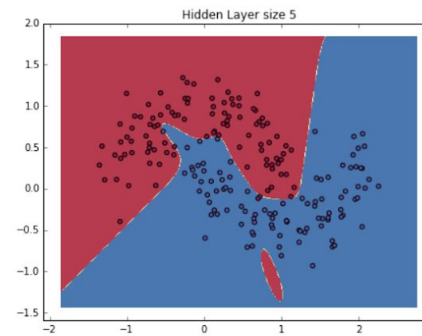
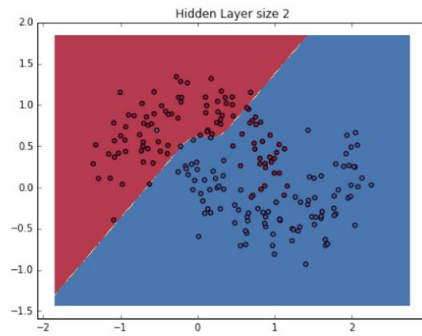
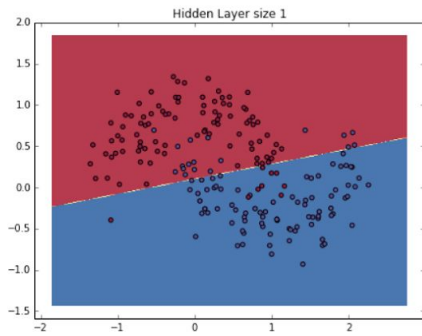


Hay razones por las cuáles con más capas obtenemos mayor performance?

Teorema de aproximación universal (Hornik et al. 1991)

- Toda función continua en $[0, 1]^d$ es aproximable por una red de una capa suficientemente grande.
- Key Takeaway: Las redes de una capa son igual de “expresivas” que las redes profundas.

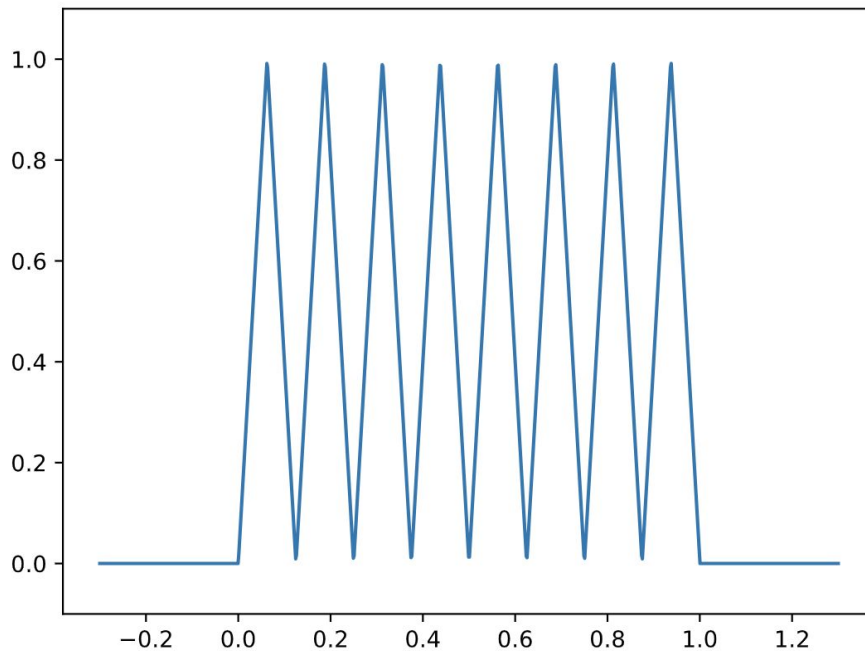
Capacidad y cantidad de neuronas



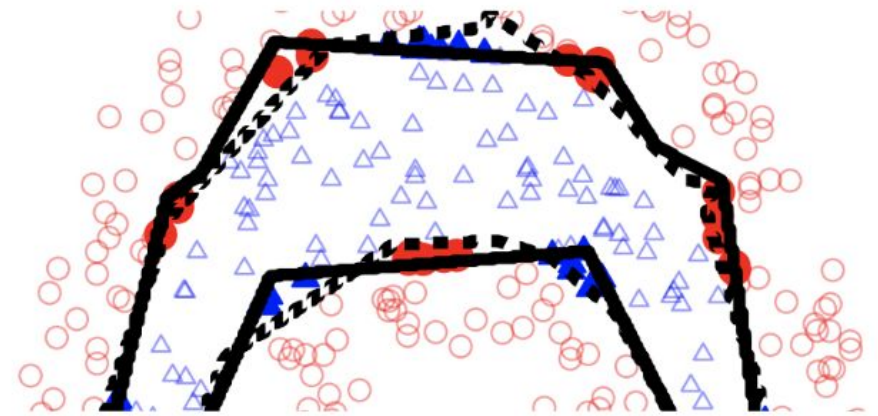
Profundidad y costo en parámetros (Telgarski et al. 2016)

- Existen funciones que pueden ser aproximadas por una red de k^3 capas que para ser aproximadas por redes de k capas necesitan al menos 2^k
- A mismo presupuesto en parámetros (optimización!) las redes profundas son altamente más expresivas.

Oscilaciones exponenciales



**1 layer adicional = 2x
oscilaciones**



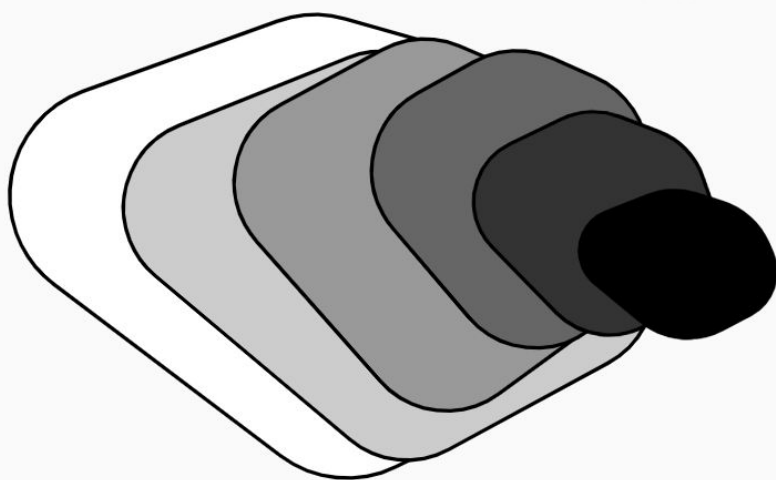
Profundidad 1 vs Profundidad 2

Los modelos profundos son más difíciles de optimizar

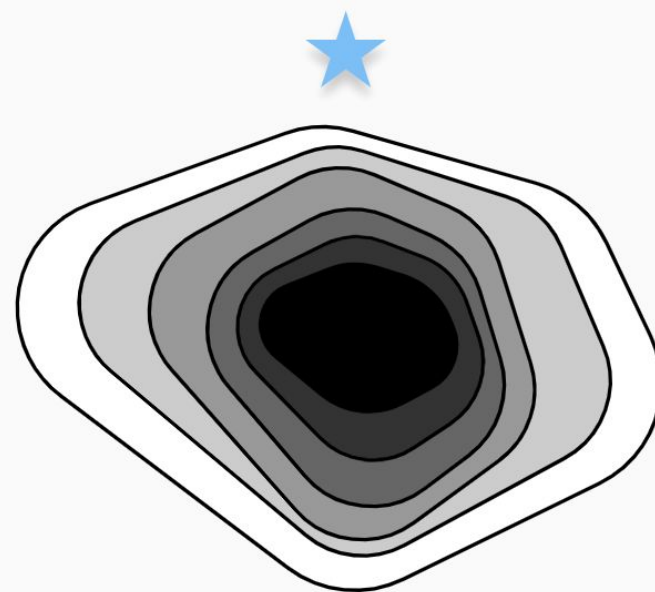


Deeper models are harder to optimize

Residual networks (He et al. 2016)



generic function classes

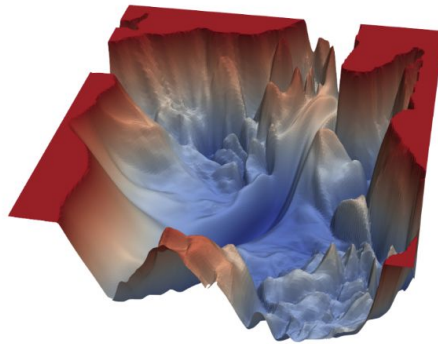


nested function classes

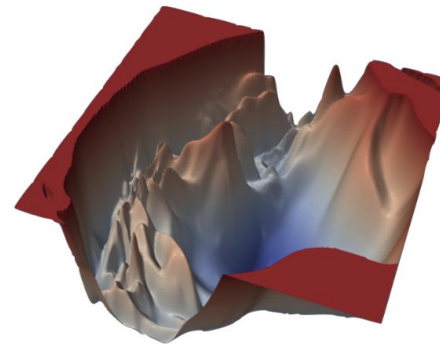
$$m(x) + x$$

Las innovaciones hacen posibles redes cada vez más profundas

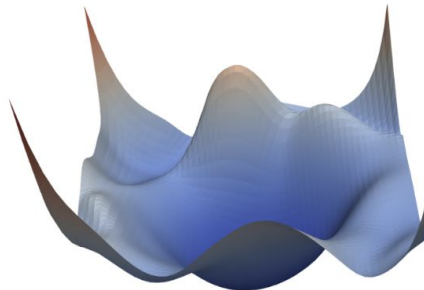
VGG-56



VGG-110



Resnet-56



Densenet-121

