



Deep Learning con Pytorch

Juan Pablo Morales
@juanpamf



Optimización de hiperparametros

Babysitting (panda)

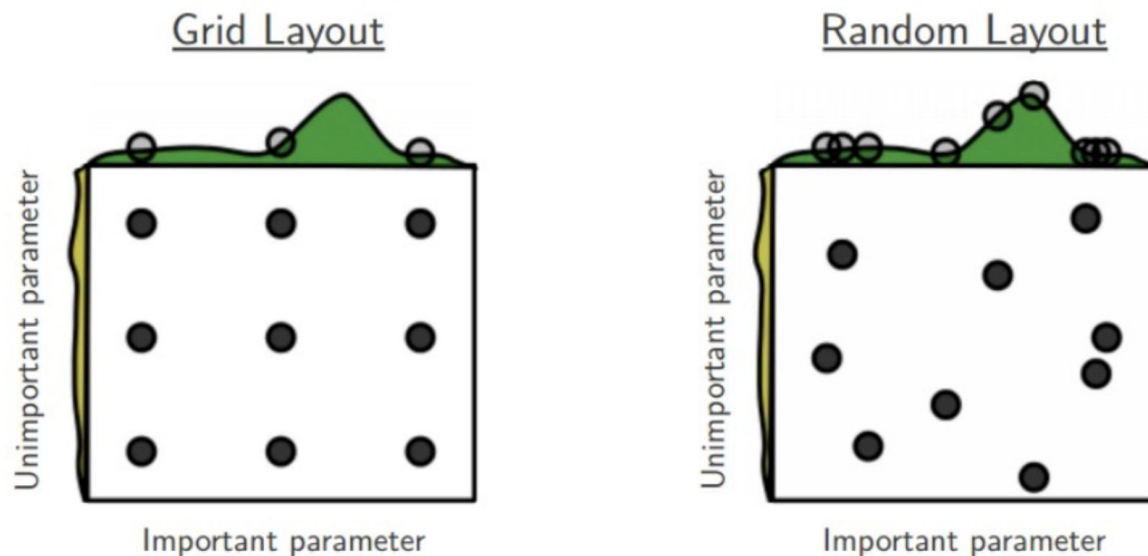
- Iteraciones sucesivas del data scientist, probando parámetros y esperando el resultado
- También conocido como Student Gradient Descent
- No es suficientemente sistemático (ej: empresas)

Grid Search (ML tradicional)

- Definimos los hiperparametros que vamos a optimizar
- Definimos el rango que pueden tomar los hiperparametros
- Definimos una cantidad de valores por probar
- Separamos el espacio de forma regular, formando un grid

Random search

- El deep learning tiene muchos más hiperparametros que el ML tradicional



Random search

- Random search es más eficiente para deep learning
- La mejor forma de utilizarlo es definiendo el rango de los parámetros con una escala correcta
 - LR: escala logaritmica
 - Batch_size: potencias de 2
- Por costo computacional podemos optar por optimizar un solo hiperparametro

Hiperparametros de mayor importancia

- Learning rate
- Momentum
- Mini-batch size
- Neuronas en cada layer
- Cantidad de layers