

Generalización y Aprendizaje de Representaciones

Qué aprende realmente una red neuronal

Cesar Garcia

2025

Introducción

- Entender qué significa **generalizar**
- Ir más allá de accuracy y loss
- Introducir el concepto de **representaciones internas**
- Preparar el terreno para autoencoders

¿Generalizar es simplemente memorizar mejor?

Más allá de los números

Dos modelos pueden tener:

- misma accuracy
- misma pérdida

Pero aprender cosas **muy distintas**.

¿Cómo podemos saber qué aprendió realmente una red?

Bias–Variance tradeoff

Intuición

- Modelos simples → alto bias
- Modelos complejos → alta varianza

Generalizar es encontrar el equilibrio.

¿Qué ocurre si aumentamos la capacidad sin control?

Profundidad y parámetros

Más capas implican:

- más parámetros
- más capacidad expresiva

Pero también:

- más riesgo de sobreajuste

¿Más grande siempre significa mejor?

Aprendizaje progresivo

Cada capa:

- transforma el espacio
- extrae características más abstractas

¿Qué tipo de información vive en una capa intermedia?

Aprendizaje de representaciones

Idea central

La red aprende una nueva geometría del espacio de entrada.

Puntos similares → más cercanos

Puntos distintos → más separados

¿Por qué esta reorganización ayuda a generalizar?

Visualizar representaciones

De datos a activaciones

En el notebook:

- compararemos modelos shallow vs deep
- visualizaremos activaciones internas

¿Qué esperarías ver en una buena representación?

Mirando hacia adelante

¿Y si entrenamos una red **solo** para aprender representaciones?

Eso motiva:

- autoencoders
- espacios latentes

¿Qué pasaría si quitamos completamente las etiquetas?

Idea clave de la sesión

Generalizar es estructurar

Una red que generaliza bien:

- no memoriza
- organiza el espacio

Aprender es reconfigurar la geometría.

¿Qué mirarías ahora además de la pérdida?