13/03/2024

Aumento de datos

Conjunto de tecnicas para evitar el sobreajuste del modelo a traves de la modificacion ligera de las imagines. Tambien se utiliza para balancear la base de datos.

¿Por que es necesario aumentar los datos?

Es importante que el modelo no aprenda relaciones obvias y realmente se fije en lo que realmente importa

Lo mas habitual es aplicar rotaciones geométricas simples (rotaciones, translaciones, zoom, etc..).

Métodos de aumentos de datos

Aumento de datos offline

* Se incrementa la base de datos directamente

Aumento de datos online

* El aumento de datos se realiza al vuelo
* Se realizan transformaciones independientes en cada batch
* Se aplican transformaciones aleatorias a paca imagen
* En cada época la misma imagen recibe imágenes aleatorias
* Es importante fijar una semilla para la reproducibilidad

Redes neuronales adversarias (GANs)

No se hace una transformación de la imagen, sino que se crea una imagen completamente nueva.

* Generador: produce imágenes a través de un ruido
* Discriminador: determina si una imagen es real o no

IMPORTANTE. En el dataset de validacion no hay aumento de datos. No se pueden validador modelos con datos que realmente no son de prueba

Early stopping

Tecnica para que el modelo deje de entranar cuando este deje de mejorar