

ASPECTOS GENERALES IA EN IMAGEN MÉDICA:

¿QUÉ ASPECTOS TRABAJAR?

Radiómica: campo emergente dentro de la radiología que implica la extracción de grandes cantidades de características cuantitativas de imágenes médicas, las cuales se utilizan para crear modelos predictivos y descriptivos. Las características radiómicas se pueden dividir en cuatro grupos principales: características basadas en el tamaño y la forma, descriptores del histograma de intensidad de la imagen, descriptores de las relaciones entre píxeles de la imagen y características fractales.

Las definiciones matemáticas de estas características son independientes de la modalidad de imagen y se pueden encontrar en la literatura → Creación de modelos que pueden predecir la respuesta al tratamiento, la progresión de la enfermedad, y otros resultados clínicos basados en las características radiómicas.

Segmentación: Por ejemplo, en neuroradiología, es común estar interesados en medir el volumen de materia gris en distintas regiones anatómicas, ya que puede ser indicador de atrofia cerebral. Herramientas como SPM (<https://www.fil.ion.ucl.ac.uk/spm>) o Freesurfer (surfer.nmr.mgh.harvard.edu) nos permiten segmentar las resonancias en T1 en diferentes tejidos (materia gris, materia blanca, líquido cefalorraquídeo, etc...). Posteriormente, podemos aplicar un atlas con regiones de interés para una patología determinada para extraer el volumen de materia gris de cada una de las regiones. Estos atlas suelen contener entre 50 o 100 regiones, por lo que el paso de píxeles a volúmenes supone una reducción de 10-100 millones a 50-100 características, un número adecuado para entrenar la mayoría de algoritmos de machine learning.

Diagnóstico: Algoritmo capaz de detectar patología, cuantificarla y categorizarla

REDUCIR TAMAÑO DE LA IMAGEN:

La imagen es una matriz que contiene los valores de la magnitud medida en cada píxel, y por tanto, se trata de información desestructurada. Existen distintas estrategias para convertir estas matrices en información estructurada. La extracción de características es una parte del proceso de reducción de dimensionalidad, en el cual un conjunto inicial de los datos sin procesar se divide y se reduce a grupos más manejables. Las imágenes médicas típicas tienen entre 10 y 100 millones de píxeles, por lo que es necesaria una importante reducción.

Aumentar el tamaño de píxel reduce la calidad de la imagen a costa de reducir de forma muy significativa el número de características.

EXTRACCIÓN DE CARACTERÍSTICAS:

El término extracción de características (feature extraction) hace referencia al conjunto de procedimientos destinados a reducir el número de recursos necesarios para describir correctamente un conjunto de datos.

La extracción de características es un término general bajo el que se agrupan diversos métodos que nos permiten sortear estos problemas reduciendo el número de combinaciones de variables explicativas en los datos sin que por ello pierdan la capacidad de describir el problema con suficiente precisión.

El objetivo general es seleccionar el subconjunto de características mínimo que exprese la cantidad máxima de varianza o, dicho de otro modo, encontrar el conjunto más compacto e informativo de características.

Llevar a cabo un proceso de extracción de características riguroso y optimizado es clave para la construcción de un modelo eficaz obteniendo mayor precisión por eliminación de características innecesarias y mayor rendimiento.

En el ámbito del aprendizaje automático, el número de atributos, características o variables de entrada de un conjunto de datos se denomina dimensionalidad. Por ejemplo, un conjunto de datos que contenga como únicos atributos la altura y peso de una serie de sujetos se puede describir como un conjunto de datos bidimensional; cualquier observación de dicho conjunto podrá representarse en un gráfico de dos dimensiones. En este sentido, integraremos los métodos de reducción de la dimensionalidad en el paraguas de la extracción de características.

Métodos de reducción de características:

ANÁLISIS DE COMPONENTES INDEPENDIENTES (ICA)

El análisis ICA (Independent Component Analysis) es también un método de reducción de la dimensionalidad lineal. Tomando como datos de entrada una mezcla de componentes (características) independientes, este método tiene como objetivo identificar correctamente cada uno de ellos reduciendo, así, la cantidad de ruido presente en los datos.

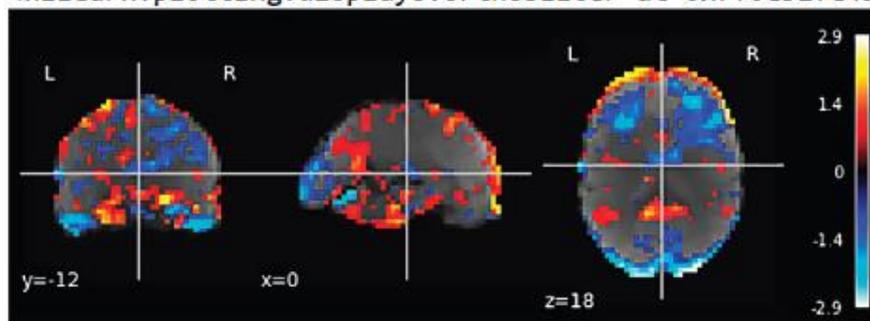
El Análisis de Componentes Independientes se utiliza habitualmente en aplicaciones médicas como el análisis de datos de electroencefalograma (EEG) y de Resonancia Magnética Funcional (fMRI), puesto que nos permite identificar las señales útiles para

el objetivo del análisis.

```
from sklearn.decomposition import FastICA
import numpy as np
# Seleccionamos 10 componentes
ica = FastICA(n_components=10, random_state=42)
components_masked = ica.fit_transform(data_masked.T).T
# Aplicamos un corte (80% señal) en los datos después de normalizar según
# la media y desviación estándar de los datos
components_masked -= components_masked.mean(axis=0)
components_masked /= components_masked.std(axis=0)
components_masked[np.abs(components_masked) < .8] = 0
# Invertimos la transformación para recuperar la estructura 3D
component_img = masker.inverse_transform(components_masked)
```

```
[70] # Finalmete, visualizamos el resultado de las operaciones de reducción
from nilearn import image
from nilearn.plotting import plot_stat_map, show
mean_img = image.mean_img(func_filename)
plot_stat_map(image.index_img(component_img, 0), mean_img)
plot_stat_map(image.index_img(component_img, 1), mean_img)
```

<nilearn.plotting.displays.OrthoSlicer at 0x7f0e32fe4950>



ANÁLISIS DE LA COMPONENTE PRINCIPAL (PCA)

El análisis de la componente principal o PCA (Principal Component Analysis) por sus siglas en inglés es una técnica de reducción dimensional lineal. Una operación PCA consiste en transformar un conjunto de variables correlacionadas (p) en un subconjunto de dimensión k ($k < p$) de variables no correlacionadas, denominadas componentes principales, conservando la mayor cantidad de varianza del conjunto de datos original.

Un análisis PCA se basa en el estudio de la correlación entre características; si la correlación es alta entre un subconjunto dado, el algoritmo PCA intentará combinar las características con una alta correlación y representar los datos asociados usando un número menor de características linealmente no correlacionadas.

MONITORIZACIÓN DE NUESTRO ALGORITMO

Neptune:

Neptune es un almacén de metadatos derivados de un proyecto de desarrollo de IA. Esta herramienta nos permite supervisar, visualizar y comparar miles de modelos de aprendizaje automático en un solo lugar. Neptune está diseñado para facilitar la colaboración; los usuarios pueden crear proyectos dentro de la aplicación, trabajar en

ellos de forma conjunta y compartir los enlaces de la interfaz de usuario entre ellos o incluso externamente. Neptune está disponible en versión cloud.

TensorBoard es un kit de herramientas de visualización para TensorFlow que permite analizar los resultados de la ejecución del proceso entrenamiento de nuestros modelos. Es una herramienta de código abierto que ofrece funcionalidades básicas que abarcan todo el flujo de trabajo de un proyecto de aprendizaje automático. Aunque cuenta con una amplia comunidad de usuarios que puede facilitar la resolución de problemas, ofrece una experiencia de trabajo más adecuada para el usuario individual.