

Clasificación de la evolución de individuos consumidores de cannabis o sanos mediante *Machine Learning*



Miriam Mota Foix
mmota.foix@gmail.com

Dirigido por: Alex Sánchez Pla
Universitat Oberta de Catalunya (UOC)

4 de enero de 2018

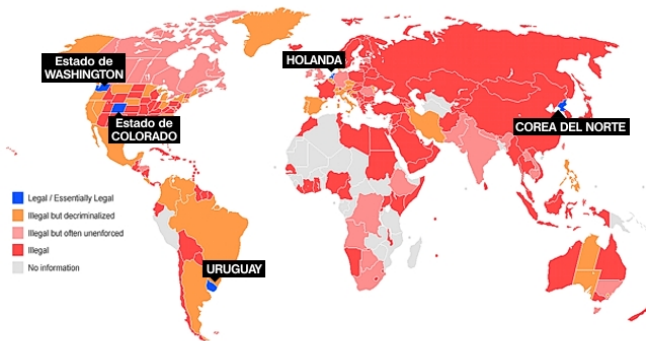
Resumen

- 1 **Introducción**
 - Cannabis
 - Objetivos
 - Datos
- 2 **Análisis**
 - Visualización de los datos
 - Pre-procesado
 - Clasificación de individuos
- 3 **Conclusiones**
 - Conclusiones
 - Limitaciones
- 4 **Bibliografía**

Cannabis

Droga blanda más usada a nivel mundial.

- Compuesto químico psicoactivo: tetrahidrocannabinol (THC)
- En España, uso restringido.



Objetivos

Generales

- Revisar las técnicas de lectura y procesado de imagen y clasificación automática.
- Clasificar los pacientes según grupo (cannabis / control).

Datos

- Datos obtenidos del estudio *Grey Matter Changes Associated with Heavy Cannabis Use: A Longitudinal sMRI Study*, Laura Koenders, 2016.
- 42 individuos → 42 sMRI
 - 20 Alto consumidores de cannabis
 - 22 Controles sanos

Tecnología: MRI

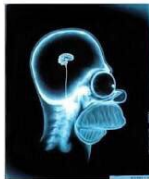
Resonancia magnética

Que es? Es una técnica de imagen que mediante imanes y ondas de radio es capaz de crear imágenes del cuerpo.

Formato. Sin formato → DICOM → Nifti

Structural MRI vs. Functional MRI

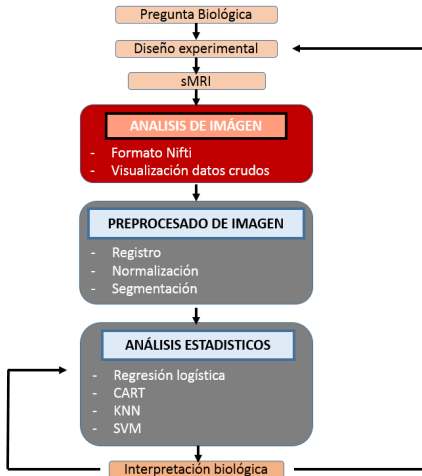
Structural MRI reveals brain anatomy.



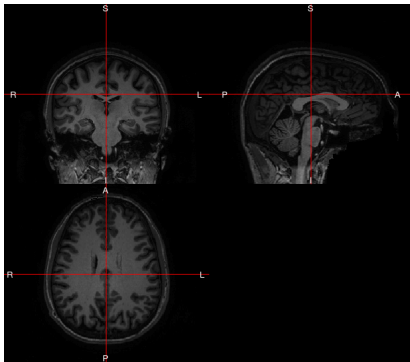
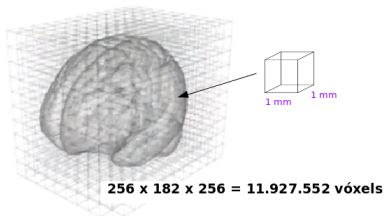
Functional MRI (fMRI) reveals brain function.



Workflow



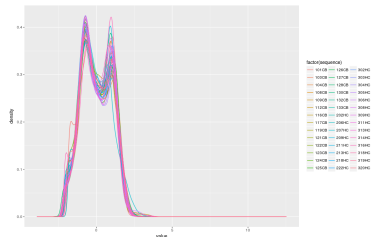
Visualización de los datos



Vóxel = La unidad básica de medida de una resonancia magnética.

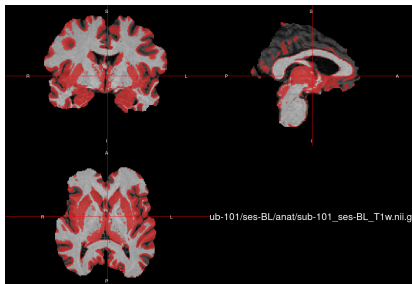
Registro y normalización

- **Registro no rígido:** transformación de las imágenes a un sistema de coordenadas estándar
- **Normalización (Z-score)**



Segmentación

- **Segmentación:** Identificación y extracción de los distintos tipos de tejidos, mediante una identificación de la escala de grises.
- **Tipos de tejido:** materia gris, materia blanca y fluido cerebro espinal.



Creación muestras



	CB	HC
Entrenamiento	12	16
Validación	8	6

Cuadro: Distribución individuos según muestras

Clasificación de individuos

Análisis estadísticos (Volumen total)

- Regresión logística
- Árbol de clasificación (CART)

Machine Learning (Información completa)

- SVM
- KNN

Información completa

Datos	Método	Accuracy	Especificidad	Sensibilidad
VT ¹	Regresión logística	50.00	25.00	83.33
VT	CART	57.14	62.50	50.00
IC ²	K-nn (5)	78.57	83.33	75.00
IC	SVM (lineal)	71.43	83.33	62.50

Cuadro: Resumen de clasificación de los datos de evaluación

¹VT = Volumen total

²IC = Información completa

Conclusiones

- Se ha realizado un resumen de las principales técnicas de pre-procesado para sMRI
- La mejor técnica de clasificación en grupos, en términos de predicción absoluta (*accuracy*), ha sido el algoritmo k-nearest neighbor (5 vecinos). Seguido de cerca por SVM
- Aparentemente los algoritmos utilizados tienen más dificultad (en estos datos) de predecir correctamente los consumidores de cannabis que los controles a excepción de la regresión logística que sucede lo contrario.
- El uso de la información completa (todos los vóxeles) produce mejores clasificaciones que teniendo en cuenta únicamente el volumen total de materia gris.

Limitaciones

- El tamaño de las muestras es bastante limitado por lo que hace que el estudio tenga poca potencia, menos reproducibilidad y mas falsos negativos
- En cada paso se han tomado decisiones relativamente subjetivas acerca de los métodos. Sería interesante saber como cambian los resultados si se tomaran otras decisiones.

Referencias



Laura Koenders, Janna Cousijn (2016) , Grey Matter Changes Associated with Heavy Cannabis Use: A Longitudinal sMRI Study *PLOS ONE* 11(5):e0152482.



Joanna Jacobus and Susan F Taper (2014) Effects of cannabis on the adolescent brain. *Current pharmaceutical design* 20(13):2186–93.



Nick Tustison Brian B. Avants and Gang Song, (2011). Advanced Normalization Tools (ANTs)



Krzysztof J. Gorgolewski, Tibor Auer (2016) The brain imaging data structure, a format for organizing and describing outputs of neuroimaging experiments *Scientific Data* 3:160044, jun 2016.



John HHopkins University. Introduction to Neurohacking In R.

