

Mentoría Data Science aplicado a BCI

Consignas TP2: Continuación Análisis y Visualización y Curación de Datos

Dominio de frecuencia.

Para quienes no estén interiorizados en los conceptos de dominio de frecuencia, espectro de frecuencia o espectrogramas, les dejamos la siguiente introducción, orientada al tema y con python. <https://mark-kramer.github.io/Case-Studies-Python/03.html#the-spectrogram>.

En el repositorio de la mentoría les dejamos un notebook de guía complementario al post anterior, con herramientas para usar en este TP (https://github.com/joaco18/MentoriaBCI/blob/main/codigo_notebooks/Introduccion%20al%20Estudio%20en%20Frecuencia.ipynb). No necesariamente todo lo que hay debe ser usado.

Es importante aprovechar la interdisciplinariedad de los grupos para compartir el conocimiento entre los integrantes. Estamos siempre disponibles para consultas, pero creemos importante que interactúen entre ustedes. Quienes sepan de estos temas, introduzcan a sus compañeros en ellos.

Para todos los ejercicios usar un sujeto diferente al usado en el notebook de ejemplos.

Antes de empezar: Familiarizarse con las nociones de filtrado en frecuencia y aplicarlas a lo largo de los siguientes ejercicios cuando lo crea necesario.

- A) Nivel Segmento/Estado: Seleccione los datos correspondientes a un paciente y un canal, y para él defina un conjunto de datos para cada estado presente en el dataset. Para cada uno de ellos estudie los siguientes elementos y luego compárelos.
- a) A partir de la frecuencia de muestreo previamente determinada y usando el teorema de Nyquist ($f_{NQ} = f_s/2$). ¿Qué frecuencia máxima pudo ser registrada en los datos disponibles?
 - b) Compute y grafique la densidad espectral de potencia para cada segmento. Elija las escalas apropiadas para su visualización (qué magnitudes están graficando? corresponde el uso de decibeles?).
 - c) Determine la/las frecuencias con mayor presencia en cada caso. ¿Qué puede concluir con esta nueva información? ¿Encuentra diferencias en este aspecto entre los tramos 99 y los tramos 1-2? ¿Hay frecuencias no asociadas con el fenómeno a detectar?
 - d) ¿Es útil el concepto de outliers o conviene usar otros criterios? ¿Se pueden descartar señales?
 - e) Comparar en el dominio del tiempo las señales con y sin filtrado de frecuencias indeseadas. ¿Qué diferencias encuentra?

B) Nivel Paciente - un canal: Seleccione los datos correspondientes a un paciente y un canal de adquisición y para ese caso estudie los siguientes elementos:

- a) Para cada intervalo de adquisición de cada estado, determine la frecuencia de mayor presencia relevante al problema estudiado. Ej: Un canal tiene 10 intervalos 2, para cada uno de ellos la frecuencia máxima oscila alrededor de 12.5.
- b) Realice un análisis estadístico de dichas frecuencias a lo largo de todo el conjunto de intervalos para cada clase. Determine su distribución y su resumen estadístico según considere apropiado. Ej: Siguiendo el ejemplo anterior, cómo es la distribución de esas frecuencias alrededor de 12.5 en el intervalo 2 y de la misma manera para los otros.
- c) Determine si existe una diferencia estadísticamente significativa, entre los valores centrales de frecuencias para los estados existentes.
- d) ¿Son variables independientes el estado registrado de la señal y la frecuencia mayoritaria presente en el registro? Use herramientas cuantitativas y cualitativas para justificar su respuesta.
- e) Compute y grafique el espectrograma del registro completo. Probar con el registro filtrado y sin filtrar y comentar las diferencias. Ver link informativo, si quedan dudas conceptuales no duden en preguntar.
- f) Resuma las principales conclusiones de este nivel de análisis.

C) Nivel Paciente - multi canal (optativo): Seleccione los datos correspondientes a un paciente y para ese caso estudie los siguientes elementos:

- a) Compute el espectrograma para cada canal, y compárelos. ¿A simple vista, existe alguna correlación?
- b) A partir de la respuesta anterior. ¿Considera relevante trabajar con todos los canales disponibles o podría quedarse con un subconjunto?
- c) Resuma las principales conclusiones de este nivel de análisis.

D) Nivel Multi-Paciente (optativo).

- a) A partir de las conclusiones extraídas de los niveles de análisis anteriores. Decida cuáles son los aspectos más importantes a analizar de los registros de un paciente y compárelos entre pacientes. ¿Encuentra diferencias significativas entre pacientes?

E) Comparando con el Trabajo Práctico anterior.

- a) ¿Cambia la complejidad del análisis según el dominio en el que se estudien los datos?
- b) Pensando en un problema de clasificación supervisada, ¿resulta más cómodo trabajar los datos en uno de los dos dominios?