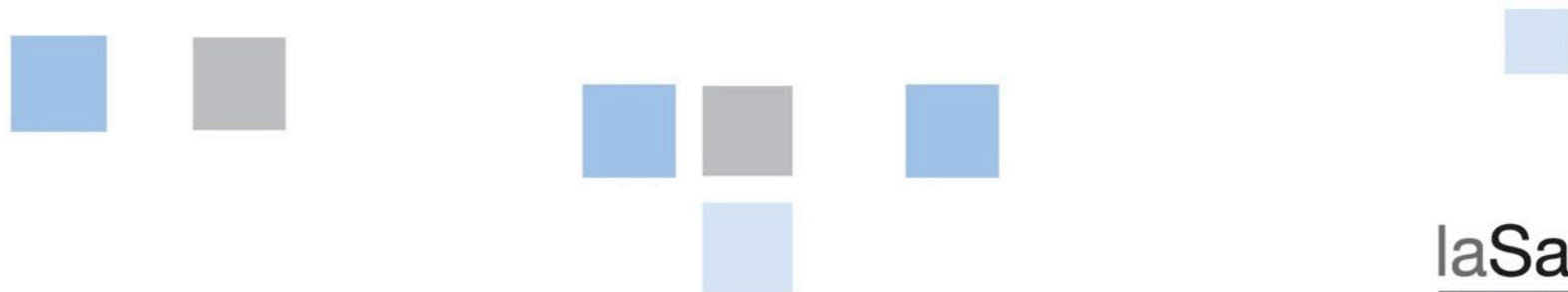


Práctica de minería de datos

Clasificación de embarcaciones según su sonido

Curso 2024-25

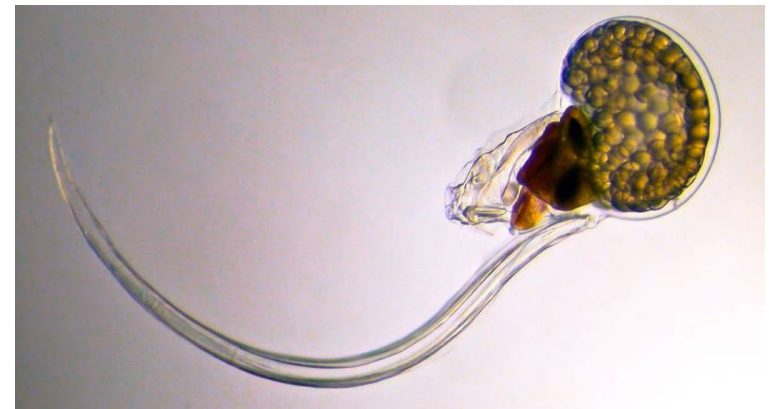


Contexto

El proyecto DeuteroNoise tiene como objetivo principal abordar los desafíos del ruido submarino generado por actividades humanas, especialmente las relacionadas con el tráfico marítimo, para evaluar su impacto sobre un grupo de invertebrados marinos llamados **deuterostomos** (concretamente, los Oikopleura Dioica).



<https://www.deuteronoise.eu/>

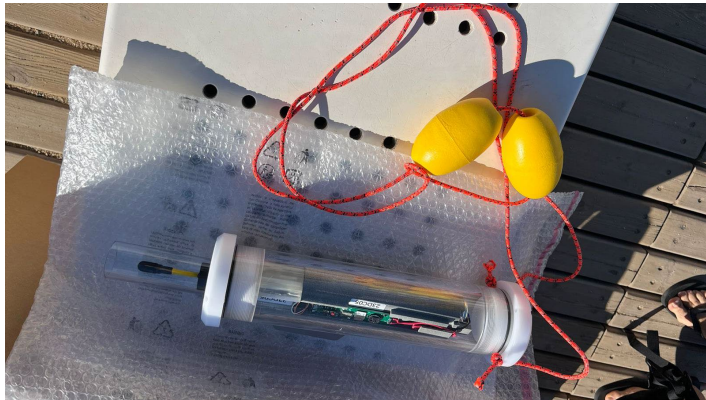


Nuestra tarea: clasificación de embarcaciones

Una de las principales fuentes de contaminación acústica submarina antropogénica viene dada por el ruido de barcos.

Actualmente, para capturar el ruido de barcos se hace de la siguiente manera:

1- Se tira un hidrófono (micrófono que graba bajo el agua) al mar.



Nuestra tarea: clasificación de embarcaciones

- 1- Se tira un hidrófono (micrófono que graba bajo el agua) al mar.
- 2- Se deja grabando durante horas o días...
- 3- Se recoge el hidrófono.



Nuestra tarea: clasificación de embarcaciones

- 1- Se tira un hidrófono (micrófono que graba bajo el agua) al mar.
- 2- Se deja grabando durante horas o días...
- 3- Se recoge el hidrófono.
- 4- Se cruzan los datos del hidrófono con los datos de localización AIS que mandan los barcos.



Nuestra tarea: clasificación de embarcaciones

Este proceso podría automatizarse si el propio sensor dispusiera de un sistema de Machine Learning capaz de detectar y clasificar barcos de forma automática.

Y... de esto vamos a encargarnos nosotros.

Objetivos de la práctica

Analizar un problema del mundo real y proponer una solución de Machine Learning *end-to-end*.

1. Definir vuestro pipeline y los diferentes módulos que lo van a componer.
2. Explorar los datos y, si es necesario, incrementar el corpus.
3. Aplicar técnicas de *feature engineering*.
4. Aplicar distintos modelos y reportar los resultados.
5. Aplicar técnicas de validación apropiadas.

Conjunto de datos

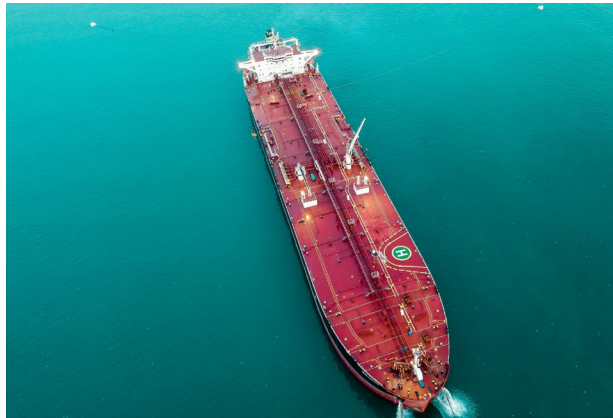
Ficheros de audio con ruido de barcos de cuatro categorías distintas:



Cargo



Passenger ship



Tanker (petrolero)

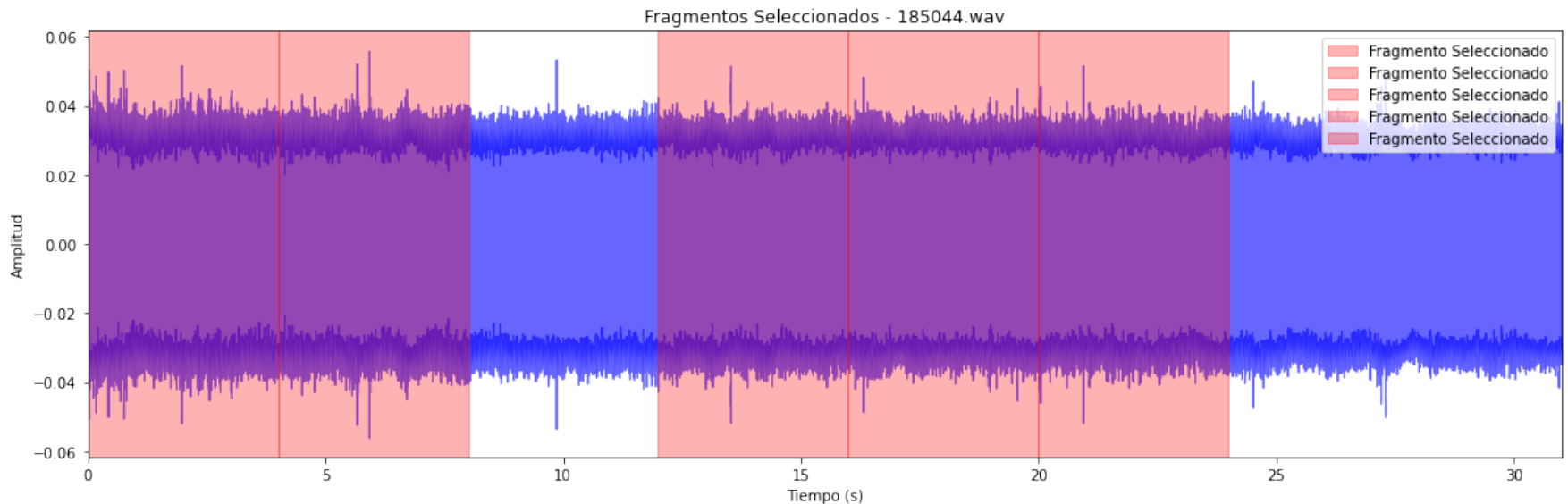


Tug (remolcador)

Conjunto de datos

Disponemos de dos conjuntos de datos:

- Dataset grande: con ~20GB de datos de barcos.
- Dataset pequeño: con ~1GB de datos de barcos.
 - Este segundo dataset es un subconjunto del primero. Contiene fragmentos de 4 segundos de como máximo 5 fragmentos de un audio original (se han escogido los que tienen más energía).



¿Qué hay en el dataset?



Ficheros de audio de 4 categorías de barcos.

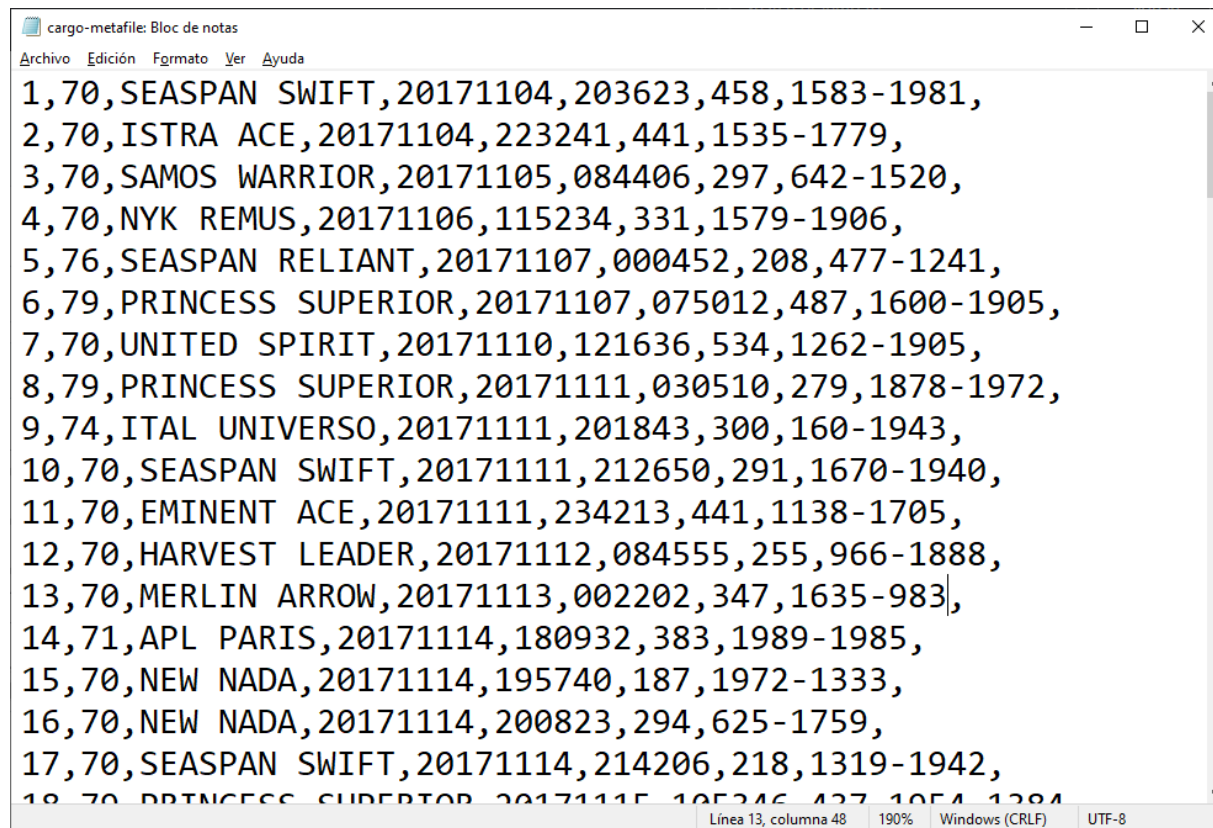


Ficheros de metadatos para cada una de las categorías.

Ejemplo de fichero de metadatos

Significado de las columnas:

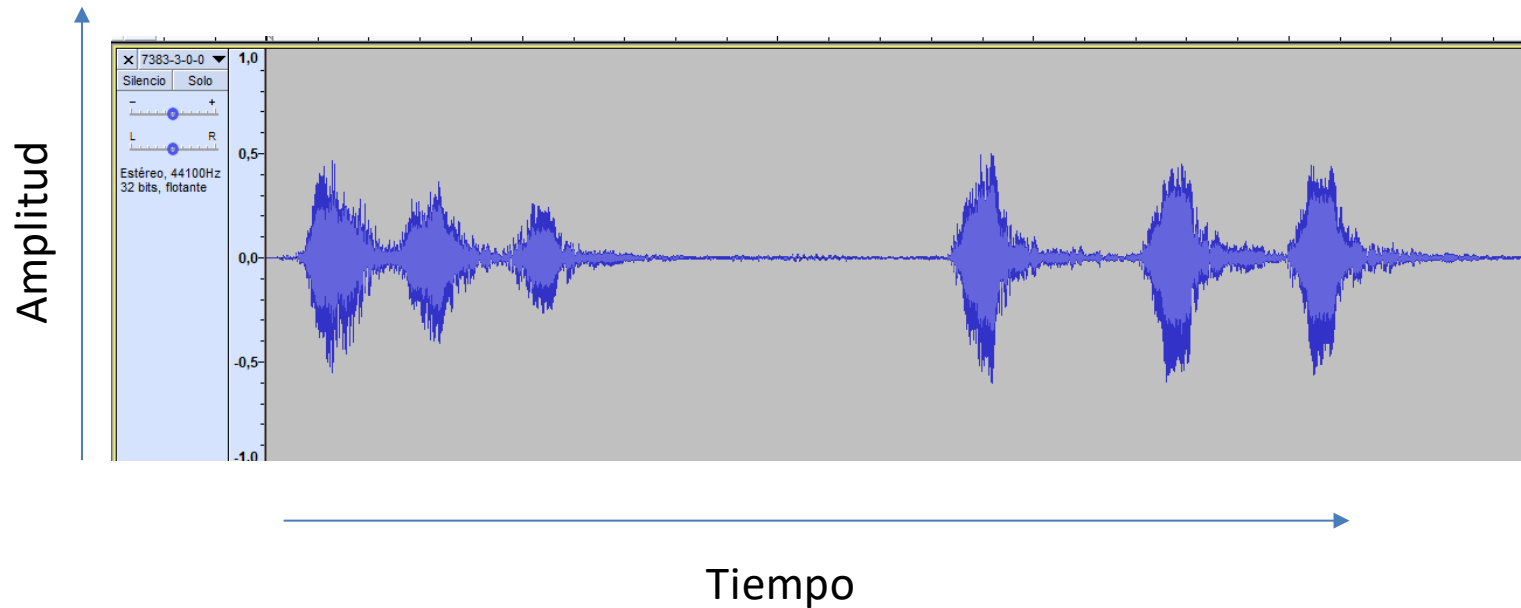
Un índice, ID de la clase, Nombre del barco, día de la grabación, duración del audio, distancia del barco al sensor



```
cargos-metafile: Bloc de notas
Archivo  Edición  Formato  Ver  Ayuda
1,70,SEASPAN SWIFT,20171104,203623,458,1583-1981,
2,70,ISTRA ACE,20171104,223241,441,1535-1779,
3,70,SAMOS WARRIOR,20171105,084406,297,642-1520,
4,70,NYK REMUS,20171106,115234,331,1579-1906,
5,76,SEASPAN RELIANT,20171107,000452,208,477-1241,
6,79,PRINCESS SUPERIOR,20171107,075012,487,1600-1905,
7,70,UNITED SPIRIT,20171110,121636,534,1262-1905,
8,79,PRINCESS SUPERIOR,20171111,030510,279,1878-1972,
9,74,ITAL UNIVERSO,20171111,201843,300,160-1943,
10,70,SEASPAN SWIFT,20171111,212650,291,1670-1940,
11,70,EMINENT ACE,20171111,234213,441,1138-1705,
12,70,HARVEST LEADER,20171112,084555,255,966-1888,
13,70,MERLIN ARROW,20171113,002202,347,1635-983,
14,71,APL PARIS,20171114,180932,383,1989-1985,
15,70,NEW NADA,20171114,195740,187,1972-1333,
16,70,NEW NADA,20171114,200823,294,625-1759,
17,70,SEASPAN SWIFT,20171114,214206,218,1319-1942,
18,70,PRINCESS SUPERIOR,20171115,105246,427,1054-1284
```

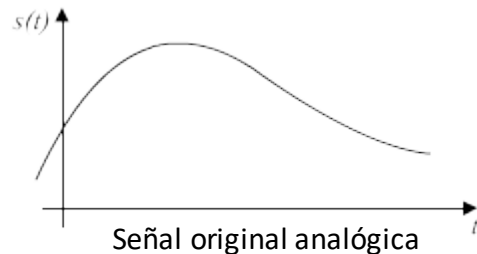
Señales de audio

■ Forma de onda:



Señales de audio

- Frecuencia de muestreo: con qué frecuencia se ha obtenido una muestra digital.
- Se suele denominar sampling frequency (f_s).



Valores típicos en audio:

44.100 Hz

o

22.050 Hz

En el caso de nuestro dataset, la $f_s = 32.000$ Hz

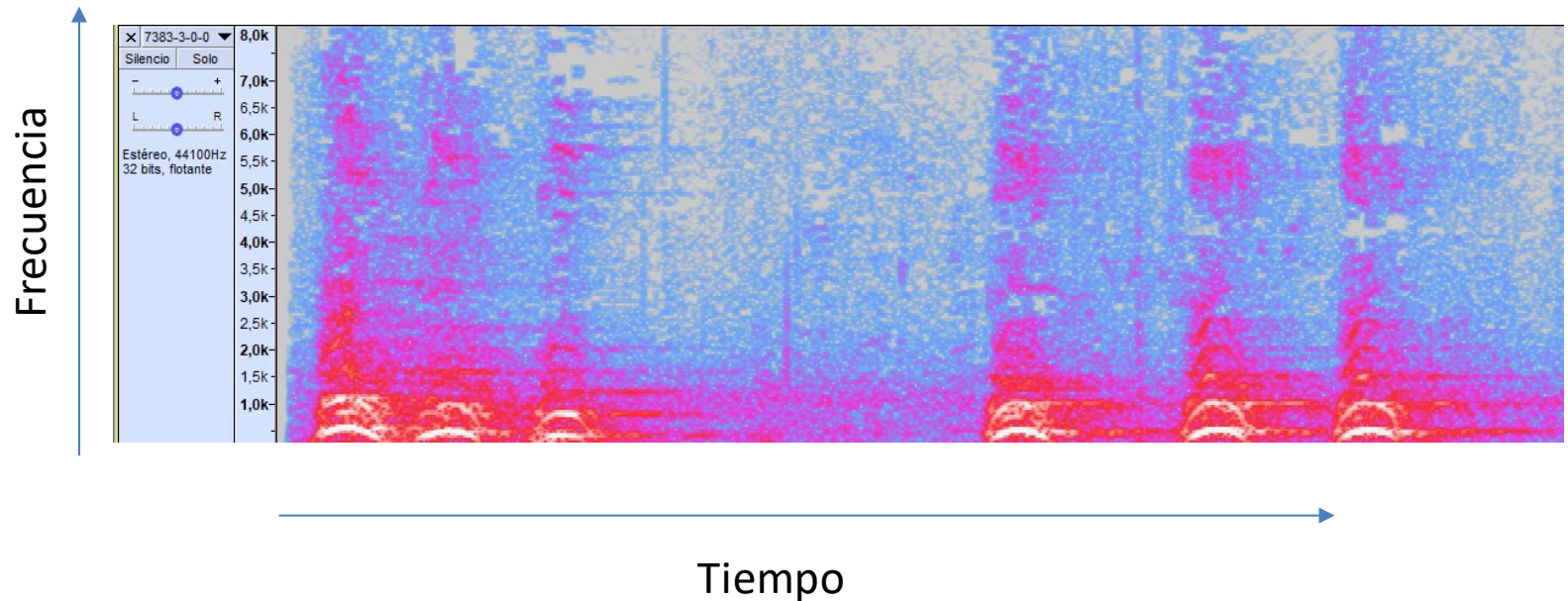
Features en audio

- **Características temporales:** Son fáciles de obtener y su interpretación física es muy sencilla. Ejemplos: energía del señal, Zero Crossing Rate (ZCR), amplitud máxima, energía mínima, etc.
- **Características espectrales:** Se obtienen a partir de la transformación del señal temporal a señal frecuencial utilizando la Transformada de Fourier. Ejemplos: Frecuencia fundamental, componentes frecuenciales, centroide espectral, flujo espectral, densidad espectral, MFCC, etc. Se utilizan para identificar notas, ritmo, melodía, etc.
- **Características espectrotemporales:** tienen en cuenta tanto la frecuencia del señal como su evolución temporal. Ejemplo: espectrograma, espectrograma mel.

Señales de audio

■ Espectrograma:

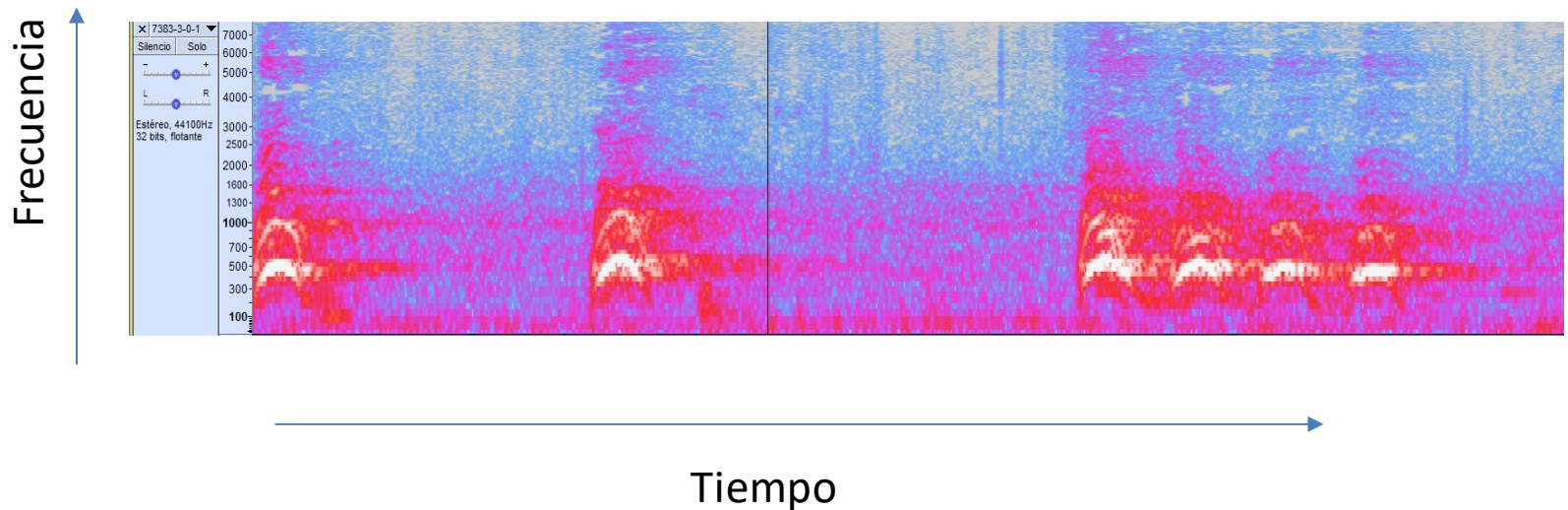
Color = intensidad



Señales de audio

- Espectrograma mel:
- Se adapta a la percepción auditiva de los humanos. Las frecuencias más relevantes son las que escuchamos mejor.

Color = intensidad



Librerías python

- Librosa: paquete para análisis de música y audio. Os permitirá por ejemplo leer los ficheros de audio y obtener características.
- Matplotlib: paquete para la visualización. Es una alternativa para calcular el espectrograma de la función.
- Numpy: librería numérica.
- Torchvision/Torchaudio: otra alternativa para obtener características de audio. Nos permitirá también transformar el espectrograma para adecuarlo a las medidas que necesitemos.
- H5PY: paquete para gestionar ficheros HDF5 (formato de datos jerárquicos que permite gestionar y manipular grandes cantidades de datos. En un solo fichero podemos incluir varios datasets).