

Propuesta preliminar: Detección de especies por medio de audio Aprendizaje Profundo

Dante Bermúdez Marbán

19 de marzo del 2020

1. Acerca de la propuesta

La presencia de animales endémicas de la selva es un buen indicador que nos habla del impacto que tiene el cambio climático en la destrucción del hábitat. Debido a que es más fácil escuchar a dichas especies que verlas, *Rainforest Connection* (RFCx) ha proporcionado a Kaggle, un conjunto de datos que consta de audios con sonidos de ranas y pájaros de la selva con el fin de crear un sistema que permita identificar la presencia de dichas especies.

2. Datos

Los datos se encuentran en [Kaggle](#), y consta principalmente de archivos de audio donde se escuchan las especies. Se pueden distinguir entre dos tipos de audios: verdaderos positivos y falsos positivos.

3. Objetivo

El objetivo en concreto es, para cada archivo de audio, determinar probabilidades de que las especies en estudio, estén presentes en dicho audio.

4. Posibles metodologías

Una manera de abordar el problema, es por medio de un enfoque de clasificación multiclase, donde cada clase corresponde a cierta especie.

Para el procesamiento de audios, se pueden usar espectrogramas o los Coeficientes Cepstrales en las Frecuencias de Mel, conocidos por sus siglas en inglés, MFCC. Un módulo de Python que puede resultar de utilidad es [librosa](#), el cual tiene procedimientos para el cálculo de dichos coeficientes, entre otras cosas. Por recomendación, también se contemplará el uso de [speechbrain](#), un módulo basado en PyTorch para tareas relacionadas con audio, entre las cuales, la de *speech recognition* puede resultar de interés.

Se pueden probar varias arquitecturas de redes para este problema, como las convoluciones y las recurrentes (como una GRU, por ejemplo).