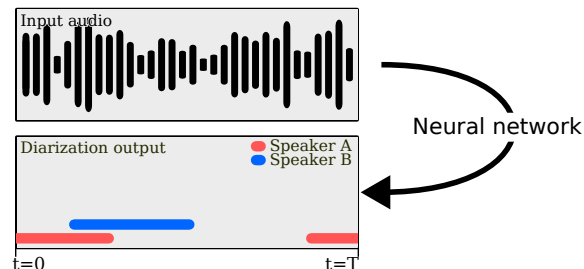




Projet de fin d'étude // Master internships Active Learning for Speaker Diarization

Keywords: deep learning, speaker diarization, active learning

Speaker diarization is the task of partitioning an audio stream into homogeneous temporal segments according to the identity of the speaker, it is commonly summarized as the task of “who speaks when”.



Active learning is a process through which a system gets better at its task by asking an oracle (e.g. a human annotator) some “questions” about a small part of the data. The answers can then be integrated into a classical fully-supervised learning scheme. Once learning on this new data doesn’t yield any more improvements, we can iteratively repeat the active learning process.

Deep learning methods usually require huge amounts of *annotated* data, and active learning is a way to mitigate that. Instead of blindly annotating all the data, we devise a smart selection criterion that helps select what annotations would bring the biggest improvement to the model, with the lowest annotation cost.

However, little work has been done on active learning applied to speaker diarization, and most of it is concerned with the speaker identities and not the temporal segments. In this internship, we propose to:

- Familiarize with the open-source diarization library **pyannote.audio**¹
- Devise, implement and test active learning schemes for the speaker diarization segmentation model of pyannote.audio

In practice, this internship will be based on (and contributed to) pyannote.audio open source toolkit and experiments will make good use of Jean Zay supercomputer ².

Please send your application (including CV, grades and references):

- to herve.bredin@irit.fr and alexis.plaquet@irit.fr
- with subject “Internship about active learning for speaker diarization”

Location: France, Toulouse, University of Toulouse III Paul Sabatier, IRIT, SAMoVA team

Date and duration: 5 to 6 months, possibly leading to a PhD

Allowance: around 600 euros monthly

Qualifications: Knowledge of python and deep learning

¹<https://github.com/pyannote/pyannote-audio>

²<http://www.idris.fr/eng/jean-zay/cpu/jean-zay-cpu-hw-eng.html>