University of Cyprus

HUMAN-CENTERED INTELLIGENT USER INTERFACES - MAI648

Marios Belk

2022
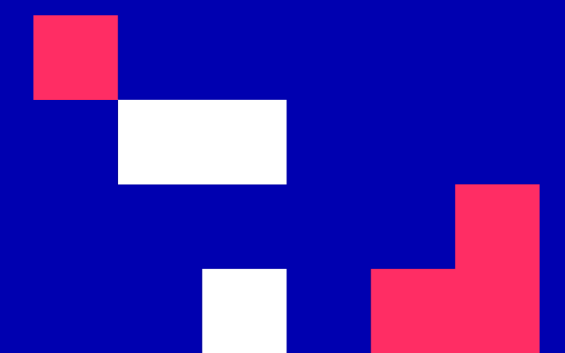
LAB 9

Voice-based User Identification

CONTENTS

- Laboratory Description

- State-of-the-art Voice-based Identification Libraries

- About Python Voice Recognition

- License

- Installation

- Main Features of the SpeechBrain Library

- Example Codes

LAB 9

Laboratory Description

- In this laboratory you will familiarize yourself with using existing state-of-the-art libraries for voice-based user identification

LAB 9

State-of-the-art Voice-based Identification Libraries

- **Kaldi**: a toolkit for speech recognition written in C++ and licensed under the Apache License v2.0. Kaldi is intended for use by speech recognition researchers
  - Python wrappers available
  - https://kaldi-asr.org

- **VoiceSens**: an open-source voice biometric solution
  - Developed in Python
  - Uses Watson Speech to Text (speech recognition)
  - https://github.com/bedangSen/VoiceSens

LAB 9
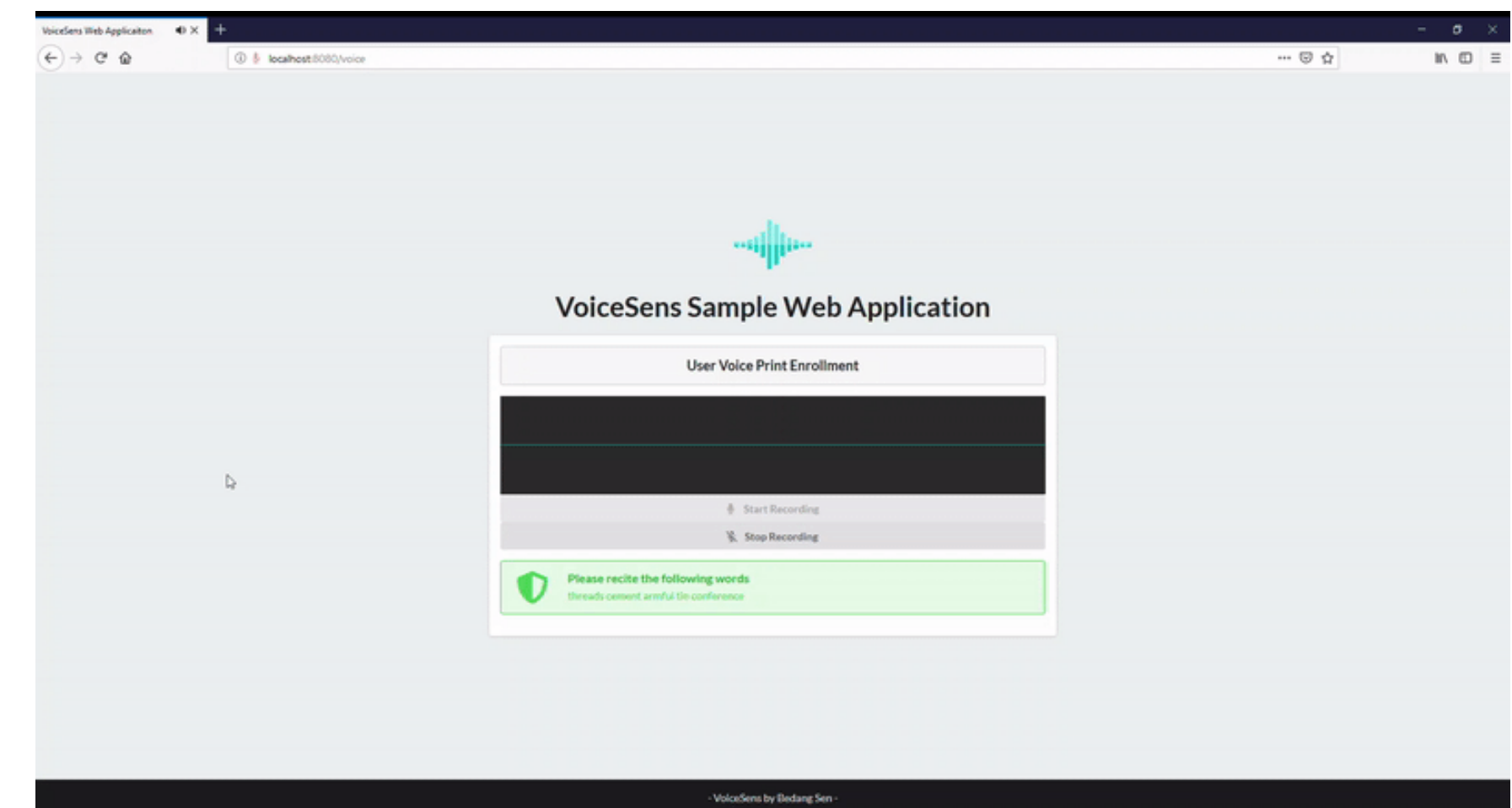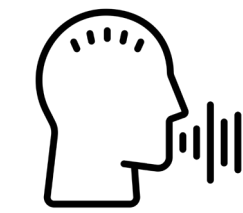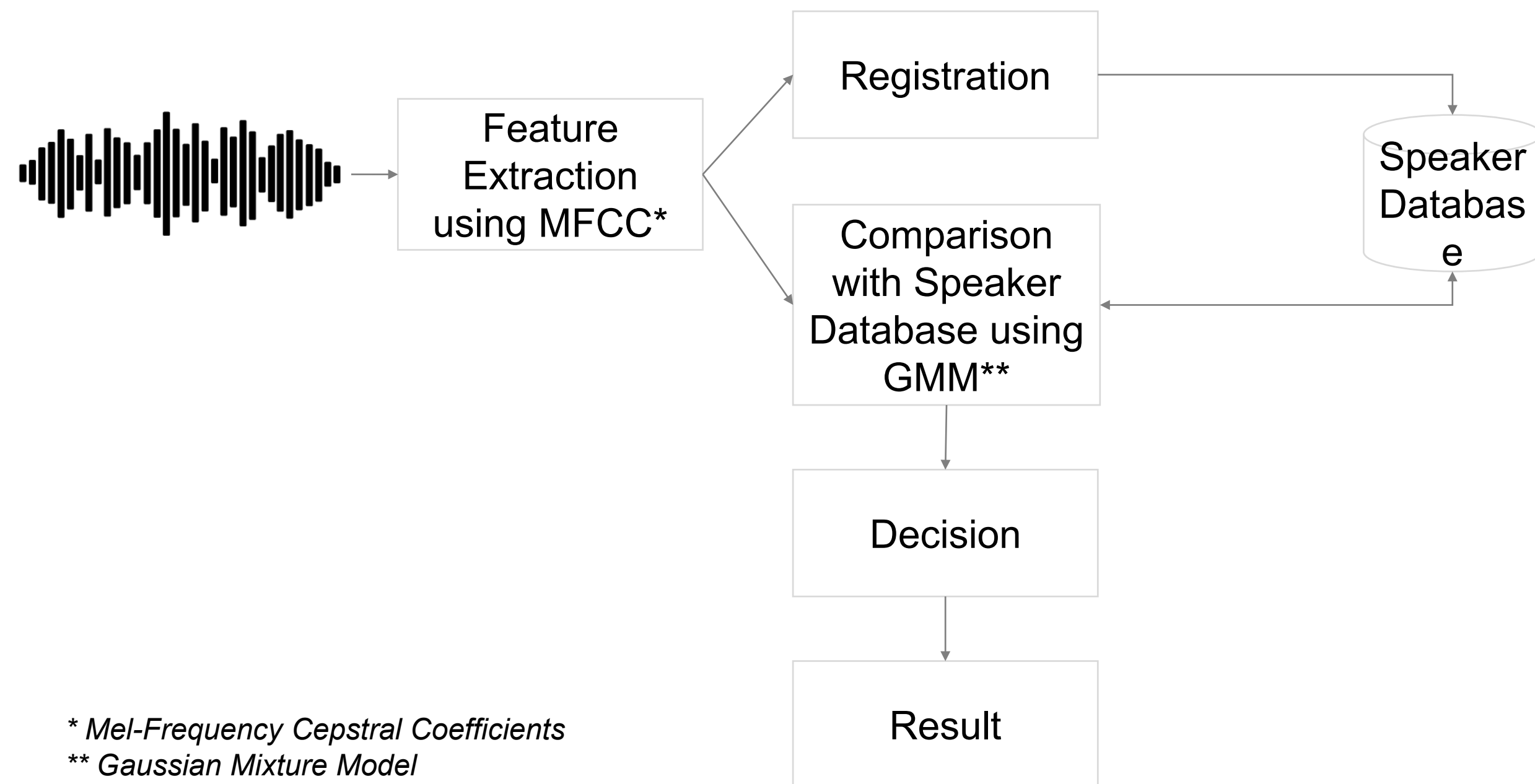
About SpeechBrain

▪ You will follow the steps as documented in the following PyTorch-based speech toolkit - https://github.com/speechbrain/speechbrain, called Speechbrain

LAB 9

About SpeechBrain

- SpeechBrain is an open-source and all-in-one conversational AI toolkit. It is designed to be simple, extremely flexible, and user-friendly

- Competitive or state-of-the-art performance is obtained in various domains - https://speechbrain.github.io

LAB 9

License

- speechbrain/speechbrain is licensed under the

- **Apache License 2.0**

- A permissive license whose main conditions require preservation of copyright and license notices. Contributors provide an express grant of patent rights. Licensed works, modifications, and larger works may be distributed under different terms and without source code.

Co-financed by the European Union
Connecting Europe Facility

8

This Master is run under the context of Action
No 2020-EU-IA-0087, co-financed by the EU CEF Telecom
under GA nr. INEA/CEF/ICT/A2020/2267423

LAB 9

Installation

- pip3 install pytorch

- pip3 install torchaudio

- pip3 install PySoundFile

- Speaker Recognition   Voxceleb   ECAPA-TDNN

- https://huggingface.co/speechbrain/spkrec-ecapa-voxceleb

Co-financed by the European Union
Connecting Europe Facility

9

This Master is run under the context of Action
No 2020-EU-IA-0087, co-financed by the EU CEF Telecom
under GA nr. INEA/CEF/ICT/A2020/2267423

Main Features of the SpeechBrain Library

- Speech recognition

- Feature extraction and augmentation

- Speech enhancement and separation

- Speaker recognition and identification

- Language Identification

LAB 9

Example Codes

```python
import torchaudio
from speechbrain.pretrained import EncoderClassifier

classifier = EncoderClassifier.from_hparams(source="speechbrain/spkrec-ecapa-voxceleb")
signal, fs =torchaudio.load('s1.wav')
embeddings = classifier.encode_batch(signal)
print(embeddings)
```

LAB 9

Using a speech recognition model (trained on LibriSpeech) to transcribe an audio recording

```
from speechbrain.pretrained import EncoderDecoderASR

asr_model =
EncoderDecoderASR.from_hparams(source="speechbrain/asr-crdnn-rnnlm-
librispeech",
savedir="pretrained_models/asr-crdnn-rnnlm-librispeech")
asr_model.transcribe_file('speechbrain/asr-crdnn-rnnlm-
librispeech/example.wav')
```

Source: Mirco Ravanelli, Loren Lugosch (2021). Introducing SpeechBrain: A general-purpose PyTorch speech processing toolkit - https://mila.quebec/en/article/introducing-speechbrain-a-general-purpose-pytorch-speech-processing-toolkit

Speaker Verification

```
from speechbrain.pretrained import SpeakerRecognition

verification =
SpeakerRecognition.from_hparams(source="speechbrain/spkrec-ecapa-vox
celeb", savedir="pretrained_models/spkrec-ecapa-voxceleb")
score, prediction =
verification.verify_files("speechbrain/spkrec-ecapa-voxceleb/example1.wav",
"speechbrain/spkrec-ecapa-voxceleb/example2.flac")
```

Source: Mirco Ravanelli, Loren Lugosch (2021). Introducing SpeechBrain: A general-purpose PyTorch speech processing toolkit - https://mila.quebec/en/article/introducing-speechbrain-a-general-purpose-pytorch-speech-processing-toolkit

Co-financed by the European Union
Connecting Europe Facility

Master programmes in Artificial
Intelligence 4 Careers in Europe

LAB 9

Speech Separation

```
from speechbrain.pretrained import SepformerSeparation as separator
import torchaudio


model =
separator.from_hparams(source="speechbrain/sepformer-wsj02mix",
savedir='pretrained_models/sepformer-wsj02mix')


# for custom file, change path
est_sources =
model.separate_file(path='speechbrain/sepformer-wsj02mix/test_mixture.wav')
```

Source: Mirco Ravanelli, Loren Lugosch (2021). Introducing SpeechBrain: A general-purpose PyTorch speech processing toolkit - https://mila.quebec/en/article/introducing-speechbrain-a-general-purpose-pytorch-speech-processing-toolkit

Speech Separation

```
torchaudio.save("source1hat.wav", est_sources[:, :,
0].detach().cpu(), 8000)
torchaudio.save("source2hat.wav", est_sources[:, :,
1].detach().cpu(), 8000)
```

Source: Mirco Ravanelli, Loren Lugosch (2021). Introducing SpeechBrain: A general-purpose PyTorch speech processing toolkit - https://mila.quebec/en/article/introducing-speechbrain-a-general-purpose-pytorch-speech-processing-toolkit

LAB 8

Sources

- Erasmus+ TRUSTID Project – https://trustid-project.eu

- https://github.com/speechbrain/speechbrain

- https://speechbrain.github.io

- Mirco Ravanelli, Loren Lugosch (2021). Introducing SpeechBrain: A general-purpose PyTorch speech processing toolkit - https://mila.quebec/en/article/introducing-speechbrain-a-general-purpose-pytorch-speech-processing-toolkit

Co-financed by the European Union
Connecting Europe Facility

16

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

MAI4CAREU

Master programmes in Artificial
Intelligence 4 Careers in Europe

Thank you.