

# SUR

People detection based on voice recordings and headshot images

## Development setup

To install the required dependencies, run the following command:

```
make venv
```

Also, download the dataset from URL and store it into `data/` directory.

If you want to use audio augmentation, download Room Impulse Response and Noise Database and extract it into `RIRS_NOISES/` directory.

## Usage

To train and evaluate Gaussian Mixture Model (GMM) for speaker recognition, run the following commands:

```
# Train your model
```

```
python audio/gmm.py train gmm_model.npz
```

```
# Evaluate your model
```

```
# WARNING: never load models from untrusted sources, loading of the model is not secure against erroneous  
# or maliciously constructed data (uses pickle.load under the hood)
```

```
python audio/gmm.py eval gmm_model.npz
```

## Audio training evaluation

Evaluation of the GMM models during training was performed using `audio/peekin.py` and `plotting.py` helper scripts. See following figure for performance of the GMM models on test and validation datasets with respect to the number of components and iterations used.

The model we selected for further use is located at `models/gmm_audio_24_27.npz` (24 components, 27 iterations).

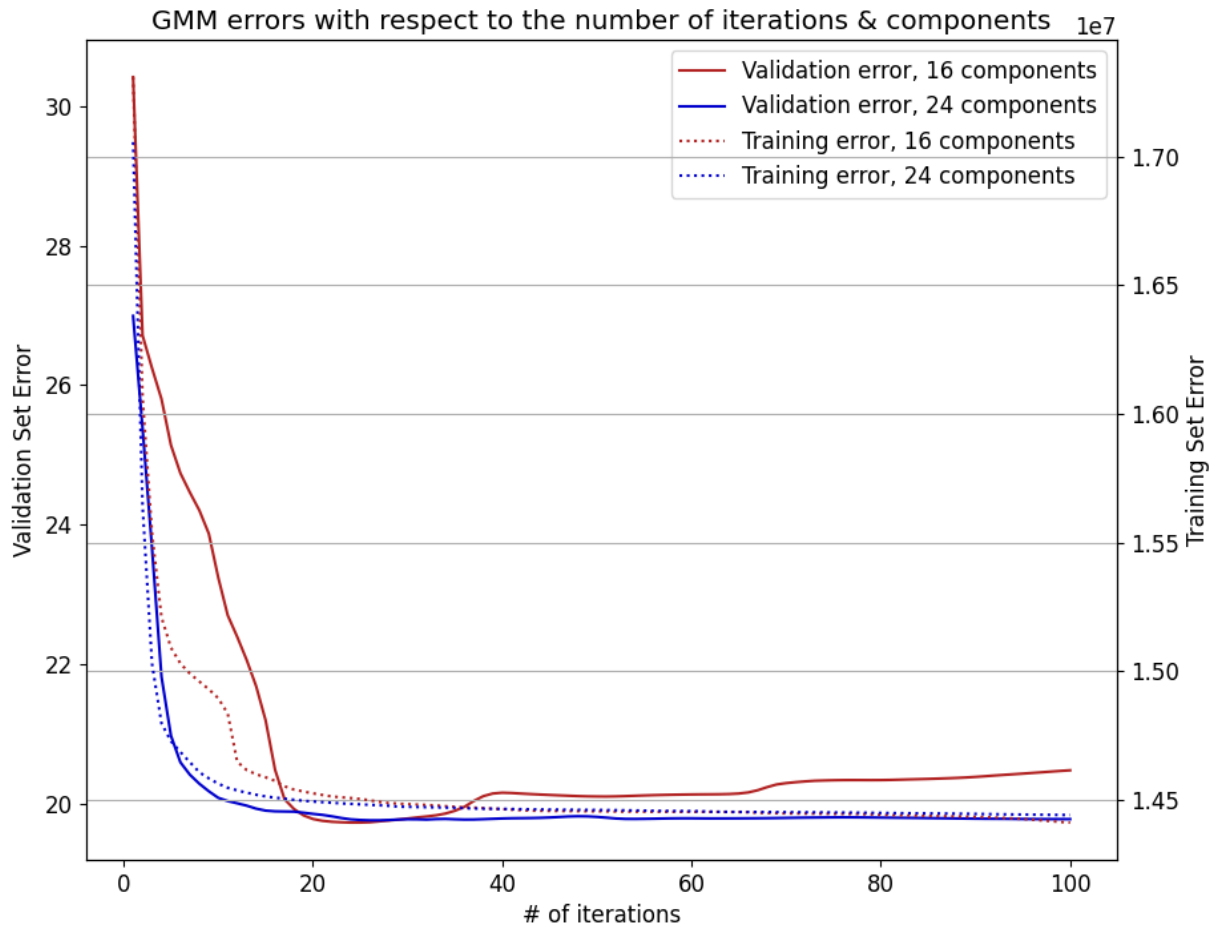


Figure 1: GMM performance