Practical Exercise - Task 4
Speech Technology - COM4511/6511

Thomas Hain

25th February 2019

1 Introduction

The aim of this task is to understand sound classification using Gaussian Mixture Models (GMMs). GMMs will be used to classify the vowel sounds from Task 3.

For this task you have two options - either to write the Gaussian mixture model code yourself (whcih would have immense learning benefit), or to use use "scikit-learn", a Machine Learning toolkit for Python. The latter has the advantage that you will be able to concentrate on speech issues.

So if you choose the second option, you need to install "scikit-learn" if you haven't done so.

`http://scikit-learn.org/stable/install.html`

2 Infrastructure

Download the helper script from MOLE. We will need the same vowel sounds from Task 3 and the "`speechtech`" package.

For more details on the `scikit-learn` GMM implementation have a look at

`http://scikit-learn.org/stable/modules/mixture.html`

3 Task

1. A separate GMM is trained for each vowel. Classification is done at first for for each frame independently. Hence frame classification accuracy should be measured.
2. A function `train_gmm` based on `scikit-learn` is provided.
3. With help from the scikit-learn website, you can fill the missing parts in the function "`train_gmm`" and "`eval_gmm`" with your own code.
4. Increase the number of mixture components and see if it improves the accuracy. Are there other ways to improve ? Are there limits and why ?
5. Note in this task the training set is used for evaluation. Is this a good idea ?