

# Audio analysis and feature extraction

David Štych  
Aleksandra Jamróz

November 21, 2022

## Introduction

## Feature extraction

On top of the baseline implementation we added another three features. The five features we used are:

1. *Average value of the entropy of the energy of the audio signal's frames*
  - Included in the baseline code
2. *Maximum value of the entropy of the energy of the audio signal's frames*
  - Included in the baseline code
3. *Average value of the spectral entropy of the audio signal's frames*
  - Implemented in the provided code
4. *Average value of the spectral flux of the audio signal's frames*
  - Implemented in the provided code
5. *Average value of the zero crossing rates of the audio signal's frames*
  - Implemented in the provided code

We have also tried experimenting with other features. We have tried adding *maximum value of the zero crossing rate of the audio signal's frames* and *average value of the spectral contrast of the audio signal's frames*. However, in both cases a detrimental effect was observed, so we decided not to include them in our model.

## Results

We kept the default split ration in the dataset (70% training, 30% testing). Firstly, we kept the default parameters (frame length, number of subframes and hop length) and just experimented with the features. We tracked the progress and recorded the results in the table below.

Features	AUC
Baseline	0.486
Baseline Mean of spectral entropies	0.554
Baseline Mean of spectral entropies Mean of spectral flux	0.577
<b>Baseline</b> <b>Mean of spectral entropies</b> <b>Mean of spectral flux</b> <b>Mean of zero crossing rate</b>	<b>0.6</b>
Baseline Mean of spectral entropies Mean of spectral flux Mean of zero crossing rate Max of zero crossing rate	0.588
Baseline Mean of spectral entropies Mean of spectral flux Mean of zero crossing rate Mean of spectral contrast	0.583

After we have decided which features to use, an automatic algorithm was implemented to tune the parameters of the feature engineering. Sets of three parameters are defined in a list and a function goes through the list and tries to train and test the model. AUC is calculated and stored in another list. After completion the set of parameters with the highest AUC is selected. Effect of the parameter change can be seen in the table below.

Best parameters found were:

- Frame length: 1024
- Number of subframes: 8
- Hop length: 256

Frame length	Number of subframes	Hop length	AUC
512	10	128	0.6004
512	10	256	0.5944
512	8	128	0.6050
512	8	256	0.5541
1024	10	128	0.5399
1024	10	256	0.5794
1024	8	128	0.6409
<b>1024</b>	<b>8</b>	<b>256</b>	<b>0.6424</b>

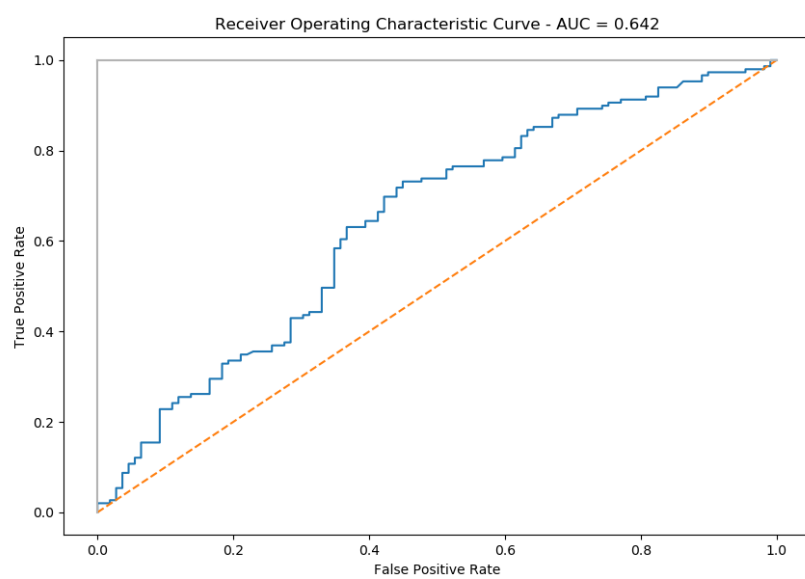


Figure 1: Best result achieved