**Learning Strategies for Voice Disorder Detection**

Preprocessing-

1. Silence at the start and end of the audio recordings is removed (if exists).
2. Down sampling to 16 kHz.
3. Each file is segmented into multiple 500 ms long snippets, with a 400 ms overlap of subsequent snippets.
4. Dataset augmentation – each sample is pitch-shifted by 8 half-semitones up and 8 half-semitones down.

Input Representation-

1. **Feature based representation with MFCC – for SVM and AE models**

Each snippet is divided into multiple blocks for a Short-time Fourier transform (STFT) with a block size of 512 and a hop size of 128 samples, respectively. Then, for each block, 20 MFCCs are extracted to form a 20 \* 63 input matrix and their mean values and standard deviations over blocks are calculated, resulting in a 40-dimensional feature vector per snippet.

1. **Mel-spectrogram – for CNN models**

SVM detection using all global scalars features

Accuracy Score of Vowel 'a' is - 94.03% with FAR=14.81%

Accuracy Score of Vowel 'i' is - 81.48% with FAR=33.33%

Accuracy Score of Vowel 'u' is - 77.78% with FAR=92.59%

**Accuracy\_score:**

SVM detection using MFCC feature only

Max Accuracy Score of Vowel 'a' is - 69.59% with 12 Coefficients

Max Accuracy Score of Vowel 'i' is - 89.18% with 12 Coefficients

Max Accuracy Score of Vowel 'u' is - 94.43% with 18 Coefficients

**Balanced\_accuracy\_score (linear kernel):**

SVM detection using MFCC feature only

Max Accuracy Score of Vowel 'a' is - 68.38% with 20 Coefficients

Max Accuracy Score of Vowel 'i' is - 82.16% with 12 Coefficients

Max Accuracy Score of Vowel 'u' is - 91.69% with 20 Coefficients

Max Accuracy Score of Vowels 'aiu' is - 81.94% with 16 Coefficients

**Balanced\_accuracy\_score (rbf kernel):**

SVM detection using MFCC feature only

Max Accuracy Score of Vowel 'a' is - 74.71% with 18 Coefficients

Max Accuracy Score of Vowel 'i' is - 80.72% with 10 Coefficients

Max Accuracy Score of Vowel 'u' is - 90.08% with 20 Coefficients

Max Accuracy Score of Vowels 'aiu' is - 81.4% with 12 Coefficients

**Balanced\_accuracy\_score (poly kernel):**

SVM detection using MFCC feature only

Max Accuracy Score of Vowel 'a' is - 51.12% with 10 Coefficients

Max Accuracy Score of Vowel 'i' is - 52.92% with 11 Coefficients

Max Accuracy Score of Vowel 'u' is - 58.84% with 10 Coefficients

Max Accuracy Score of Vowels 'aiu' is - 71.27% with 10 Coefficients

**Balanced\_accuracy\_score (sigmoid kernel):**

SVM detection using MFCC feature only

Max Accuracy Score of Vowel 'a' is - 72.96% with 15 Coefficients

Max Accuracy Score of Vowel 'i' is - 79.34% with 10 Coefficients

Max Accuracy Score of Vowel 'u' is - 85.72% with 19 Coefficients

Max Accuracy Score of Vowels 'aiu' is - 79.78% with 12 Coefficients

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Max Balanced Accuracy Score** | | | | |
| **Kernel Type** | **Linear** | **Rbf** | **Poly** | **Sigmoid** |
| Vowel 'a' | 68.38% | 74.71% | 51.12% | 72.96% |
| Vowel 'i' | 82.16% | 80.72% | 52.92% | 79.34% |
| Vowel 'u' | 91.69% | 90.08% | 58.84% | 85.72% |
| Vowels 'aiu' | 81.94% | 81.4% | 71.27% | 79.78% |

########## Vowel 'a' results ##########

Confidence test for vowel 'a'':

Confidence Interval of 90% likelihood is [0.334,0.382]

Confidence Interval of 95% likelihood is [0.33,0.387]

Confidence Interval of 98% likelihood is [0.324,0.392]

Confidence Interval of 99% likelihood is [0.32,0.396]

########## Vowel 'i' results ##########

Confidence test for vowel 'i'':

Confidence Interval of 90% likelihood is [0.217,0.256]

Confidence Interval of 95% likelihood is [0.213,0.26]

Confidence Interval of 98% likelihood is [0.208,0.264]

Confidence Interval of 99% likelihood is [0.205,0.267]

########## Vowel 'u' results ##########

Confidence test for vowel 'u'':

Confidence Interval of 90% likelihood is [0.049,0.072]

Confidence Interval of 95% likelihood is [0.047,0.074]

Confidence Interval of 98% likelihood is [0.045,0.077]

Confidence Interval of 99% likelihood is [0.043,0.079]

########## Vowels 'aiu' results ##########

Confidence test for vowels 'iau':

Confidence Interval of 90% likelihood is [0.189,0.211]

Confidence Interval of 95% likelihood is [0.187,0.214]

Confidence Interval of 98% likelihood is [0.185,0.216]

Confidence Interval of 99% likelihood is [0.183,0.218]

SVM detection using MFCC feature only

Max Balanced Accuracy Score of Vowel 'a' is - 66.77% , with Far = 46.23% and 2 Coefficients

Max Balanced Accuracy Score of Vowel 'i' is - 80.88% , with Far = 29.74% and 2 Coefficients

Max Balanced Accuracy Score of Vowel 'u' is - 85.13% , with Far = 25.98% and 2 Coefficients

Max Balanced Accuracy Score of Vowels 'aiu' is - 74.39% , with Far = 29.12% and 2 Coefficients

SVM detection using MFCC feature only - balanced

Max Balanced Accuracy Score of Vowel 'a' is - 53.7% , with Far = 55.77% and 18 Coefficients

Max Balanced Accuracy Score of Vowel 'i' is - 89.13% , with Far = 1.6% and 18 Coefficients

Max Balanced Accuracy Score of Vowel 'u' is - 95.35% , with Far = 5.37% and 10 Coefficients

Max Balanced Accuracy Score of Vowels 'aiu' is - 88.06% , with Far = 1.07% and 18 Coefficients