

Max Marmer
Carmel Ron
Odelia Hochman
Efrat Cohen

Deep Learning-proposal

In this project we will build a speaker recognizer using deep learning that will identify the speaker gender. (female/ male).

Databases:

The databases consist of around 1150 utterances.

The databases include US English male (bdl) and female (slt) speakers (both experinced voice talent) as well as other accented speakers.

The distributions include 16KHz waveform and simultaneous EGG signals and include Festival CLUNITS based voices. Complete runnable Festival Voices are included with the database distributions, as examples though better voices can be made by improving labelling.

Link to dataset: http://festvox.org/cmu_arctic/

we downloaded the database to our computer:

```
ic_a0589  arctic_a0547  arctic_a0505  arctic_a0463  arctic_a0421  arctic_a0379  arctic_a0337  arctic_a0295  arctic_a0253  arctic_a0211  arctic_a0169  arctic_a0127  arctic_a0085  arctic_a0043  arctic_a0001
ic_a0590  arctic_a0548  arctic_a0506  arctic_a0464  arctic_a0423  arctic_a0381  arctic_a0339  arctic_a0296  arctic_a0254  arctic_a0212  arctic_a0170  arctic_a0128  arctic_a0086  arctic_a0044  arctic_a0002
ic_a0591  arctic_a0549  arctic_a0507  arctic_a0465  arctic_a0423  arctic_a0381  arctic_a0339  arctic_a0297  arctic_a0255  arctic_a0213  arctic_a0171  arctic_a0129  arctic_a0087  arctic_a0045  arctic_a0003
ic_a0592  arctic_a0550  arctic_a0508  arctic_a0466  arctic_a0425  arctic_a0383  arctic_a0341  arctic_a0299  arctic_a0257  arctic_a0215  arctic_a0173  arctic_a0131  arctic_a0089  arctic_a0046  arctic_a0004
c_a0002  arctic_a0552  arctic_a0510  arctic_a0468  arctic_a0427  arctic_a0384  arctic_a0342  arctic_a0300  arctic_a0258  arctic_a0216  arctic_a0174  arctic_a0132  arctic_a0090  arctic_a0048  arctic_a0006
c_a0003  arctic_a0553  arctic_a0511  arctic_a0469  arctic_a0427  arctic_a0385  arctic_a0343  arctic_a0301  arctic_a0259  arctic_a0217  arctic_a0175  arctic_a0133  arctic_a0091  arctic_a0049  arctic_a0007
c_a0004  arctic_a0554  arctic_a0512  arctic_a0470  arctic_a0428  arctic_a0386  arctic_a0344  arctic_a0302  arctic_a0260  arctic_a0218  arctic_a0176  arctic_a0134  arctic_a0092  arctic_a0050  arctic_a0008
c_a0005  arctic_a0555  arctic_a0513  arctic_a0471  arctic_a0429  arctic_a0387  arctic_a0345  arctic_a0303  arctic_a0261  arctic_a0219  arctic_a0177  arctic_a0135  arctic_a0093  arctic_a0051  arctic_a0009
c_a0006  arctic_a0556  arctic_a0514  arctic_a0472  arctic_a0430  arctic_a0388  arctic_a0346  arctic_a0304  arctic_a0262  arctic_a0220  arctic_a0178  arctic_a0136  arctic_a0094  arctic_a0052  arctic_a0010
c_a0007  arctic_a0557  arctic_a0515  arctic_a0473  arctic_a0431  arctic_a0389  arctic_a0347  arctic_a0305  arctic_a0263  arctic_a0221  arctic_a0179  arctic_a0137  arctic_a0095  arctic_a0053  arctic_a0011
c_a0008  arctic_a0558  arctic_a0516  arctic_a0474  arctic_a0432  arctic_a0390  arctic_a0348  arctic_a0306  arctic_a0264  arctic_a0222  arctic_a0180  arctic_a0138  arctic_a0096  arctic_a0054  arctic_a0012
c_a0009  arctic_a0559  arctic_a0517  arctic_a0475  arctic_a0433  arctic_a0391  arctic_a0349  arctic_a0307  arctic_a0265  arctic_a0223  arctic_a0181  arctic_a0139  arctic_a0097  arctic_a0055  arctic_a0013
c_a0010  arctic_a0560  arctic_a0518  arctic_a0476  arctic_a0434  arctic_a0392  arctic_a0350  arctic_a0308  arctic_a0266  arctic_a0224  arctic_a0182  arctic_a0140  arctic_a0098  arctic_a0056  arctic_a0014
c_a0011  arctic_a0561  arctic_a0519  arctic_a0477  arctic_a0435  arctic_a0393  arctic_a0351  arctic_a0309  arctic_a0267  arctic_a0225  arctic_a0183  arctic_a0141  arctic_a0099  arctic_a0057  arctic_a0015
c_a0012  arctic_a0562  arctic_a0520  arctic_a0478  arctic_a0436  arctic_a0394  arctic_a0352  arctic_a0310  arctic_a0268  arctic_a0226  arctic_a0184  arctic_a0142  arctic_a0100  arctic_a0058  arctic_a0016
c_a0013  arctic_a0563  arctic_a0521  arctic_a0479  arctic_a0437  arctic_a0395  arctic_a0353  arctic_a0311  arctic_a0269  arctic_a0227  arctic_a0185  arctic_a0143  arctic_a0101  arctic_a0059  arctic_a0017
c_a0014  arctic_a0564  arctic_a0522  arctic_a0480  arctic_a0438  arctic_a0396  arctic_a0354  arctic_a0312  arctic_a0270  arctic_a0228  arctic_a0186  arctic_a0144  arctic_a0102  arctic_a0060  arctic_a0018
c_a0015  arctic_a0565  arctic_a0523  arctic_a0481  arctic_a0439  arctic_a0397  arctic_a0355  arctic_a0313  arctic_a0271  arctic_a0229  arctic_a0187  arctic_a0145  arctic_a0103  arctic_a0061  arctic_a0019
c_a0016  arctic_a0566  arctic_a0524  arctic_a0482  arctic_a0440  arctic_a0398  arctic_a0356  arctic_a0314  arctic_a0272  arctic_a0230  arctic_a0188  arctic_a0146  arctic_a0104  arctic_a0062  arctic_a0020
c_a0017  arctic_a0567  arctic_a0525  arctic_a0483  arctic_a0441  arctic_a0399  arctic_a0357  arctic_a0315  arctic_a0273  arctic_a0231  arctic_a0189  arctic_a0147  arctic_a0105  arctic_a0063  arctic_a0021
c_a0018  arctic_a0568  arctic_a0526  arctic_a0484  arctic_a0442  arctic_a0400  arctic_a0358  arctic_a0316  arctic_a0274  arctic_a0232  arctic_a0190  arctic_a0148  arctic_a0106  arctic_a0064  arctic_a0022
c_a0019  arctic_a0569  arctic_a0527  arctic_a0485  arctic_a0443  arctic_a0401  arctic_a0359  arctic_a0317  arctic_a0275  arctic_a0233  arctic_a0191  arctic_a0149  arctic_a0107  arctic_a0065  arctic_a0023
c_a0020  arctic_a0570  arctic_a0528  arctic_a0486  arctic_a0444  arctic_a0402  arctic_a0360  arctic_a0318  arctic_a0276  arctic_a0234  arctic_a0192  arctic_a0150  arctic_a0108  arctic_a0066  arctic_a0024
c_a0021  arctic_a0571  arctic_a0529  arctic_a0487  arctic_a0445  arctic_a0403  arctic_a0361  arctic_a0319  arctic_a0277  arctic_a0235  arctic_a0193  arctic_a0151  arctic_a0109  arctic_a0067  arctic_a0025
c_a0022  arctic_a0572  arctic_a0530  arctic_a0488  arctic_a0446  arctic_a0404  arctic_a0362  arctic_a0320  arctic_a0278  arctic_a0236  arctic_a0194  arctic_a0152  arctic_a0110  arctic_a0068  arctic_a0026
c_a0023  arctic_a0573  arctic_a0531  arctic_a0489  arctic_a0447  arctic_a0405  arctic_a0363  arctic_a0321  arctic_a0279  arctic_a0237  arctic_a0195  arctic_a0153  arctic_a0111  arctic_a0069  arctic_a0027
c_a0024  arctic_a0574  arctic_a0532  arctic_a0490  arctic_a0448  arctic_a0406  arctic_a0364  arctic_a0322  arctic_a0280  arctic_a0238  arctic_a0196  arctic_a0154  arctic_a0112  arctic_a0070  arctic_a0028
c_a0025  arctic_a0575  arctic_a0533  arctic_a0491  arctic_a0449  arctic_a0407  arctic_a0365  arctic_a0323  arctic_a0281  arctic_a0239  arctic_a0197  arctic_a0155  arctic_a0113  arctic_a0071  arctic_a0029
c_a0026  arctic_a0576  arctic_a0534  arctic_a0492  arctic_a0450  arctic_a0408  arctic_a0366  arctic_a0324  arctic_a0282  arctic_a0240  arctic_a0198  arctic_a0156  arctic_a0115  arctic_a0072  arctic_a0030
c_a0027  arctic_a0577  arctic_a0535  arctic_a0493  arctic_a0451  arctic_a0409  arctic_a0367  arctic_a0325  arctic_a0283  arctic_a0241  arctic_a0199  arctic_a0157  arctic_a0117  arctic_a0073  arctic_a0031
c_a0028  arctic_a0578  arctic_a0536  arctic_a0494  arctic_a0452  arctic_a0410  arctic_a0368  arctic_a0326  arctic_a0284  arctic_a0242  arctic_a0200  arctic_a0159  arctic_a0116  arctic_a0074  arctic_a0032
c_a0029  arctic_a0579  arctic_a0537  arctic_a0495  arctic_a0453  arctic_a0411  arctic_a0369  arctic_a0327  arctic_a0285  arctic_a0243  arctic_a0201  arctic_a0159  arctic_a0117  arctic_a0075  arctic_a0033
c_a0030  arctic_a0580  arctic_a0538  arctic_a0496  arctic_a0454  arctic_a0412  arctic_a0370  arctic_a0328  arctic_a0286  arctic_a0244  arctic_a0202  arctic_a0160  arctic_a0118  arctic_a0076  arctic_a0034
c_a0031  arctic_a0581  arctic_a0539  arctic_a0497  arctic_a0455  arctic_a0413  arctic_a0371  arctic_a0329  arctic_a0287  arctic_a0245  arctic_a0203  arctic_a0161  arctic_a0119  arctic_a0077  arctic_a0035
c_a0032  arctic_a0582  arctic_a0540  arctic_a0498  arctic_a0456  arctic_a0414  arctic_a0372  arctic_a0330  arctic_a0288  arctic_a0246  arctic_a0204  arctic_a0162  arctic_a0120  arctic_a0078  arctic_a0036
c_a0033  arctic_a0583  arctic_a0541  arctic_a0499  arctic_a0457  arctic_a0415  arctic_a0373  arctic_a0331  arctic_a0289  arctic_a0247  arctic_a0205  arctic_a0163  arctic_a0121  arctic_a0079  arctic_a0037
c_a0034  arctic_a0584  arctic_a0542  arctic_a0500  arctic_a0458  arctic_a0416  arctic_a0374  arctic_a0332  arctic_a0290  arctic_a0248  arctic_a0206  arctic_a0164  arctic_a0122  arctic_a0080  arctic_a0038
c_a0035  arctic_a0585  arctic_a0543  arctic_a0501  arctic_a0459  arctic_a0417  arctic_a0375  arctic_a0333  arctic_a0291  arctic_a0249  arctic_a0207  arctic_a0165  arctic_a0123  arctic_a0081  arctic_a0039
c_a0036  arctic_a0586  arctic_a0544  arctic_a0502  arctic_a0460  arctic_a0418  arctic_a0376  arctic_a0334  arctic_a0292  arctic_a0250  arctic_a0208  arctic_a0166  arctic_a0124  arctic_a0082  arctic_a0040
c_a0037  arctic_a0587  arctic_a0545  arctic_a0503  arctic_a0461  arctic_a0419  arctic_a0377  arctic_a0335  arctic_a0293  arctic_a0251  arctic_a0209  arctic_a0167  arctic_a0125  arctic_a0083  arctic_a0041
c_a0038  arctic_a0588  arctic_a0546  arctic_a0504  arctic_a0462  arctic_a0420  arctic_a0378  arctic_a0336  arctic_a0294  arctic_a0252  arctic_a0210  arctic_a0168  arctic_a0126  arctic_a0084  arctic_a0042
```

We will use MFCC features on the signal to obtain them from this website:

https://github.com/jameslyons/python_speech_features

The features we will use:

Parameter	Description
signal	the audio signal from which to compute features. Should be an N*1 array
samplerate	the samplerate of the signal we are working with.
winlen	the length of the analysis window in seconds. Default is 0.025s (25 milliseconds)
winstep	the step between successive windows in seconds. Default is 0.01s (10 milliseconds)
numcep	the number of cepstrum to return, default 13
nfilt	the number of filters in the filterbank, default 26.
nfft	the FFT size. Default is 512
lowfreq	lowest band edge of mel filters. In Hz, default is 0
highfreq	highest band edge of mel filters. In Hz, default is samplerate/2
preemph	apply preemphasis filter with preemph as coefficient. 0 is no filter. Default is 0.97
ceplifter	apply a lifter to final cepstral coefficients. 0 is no lifter. Default is 22
appendEnergy	if this is true, the zeroth cepstral coefficient is replaced with the log of the total frame energy.
returns	A numpy array of size (NUMFRAMES by numcep) containing features. Each row holds 1 feature vector.