Mood Detect

Chih Wei Tung, Shanshan Zhao, Weixuan Jiang, Zhonghao Guo

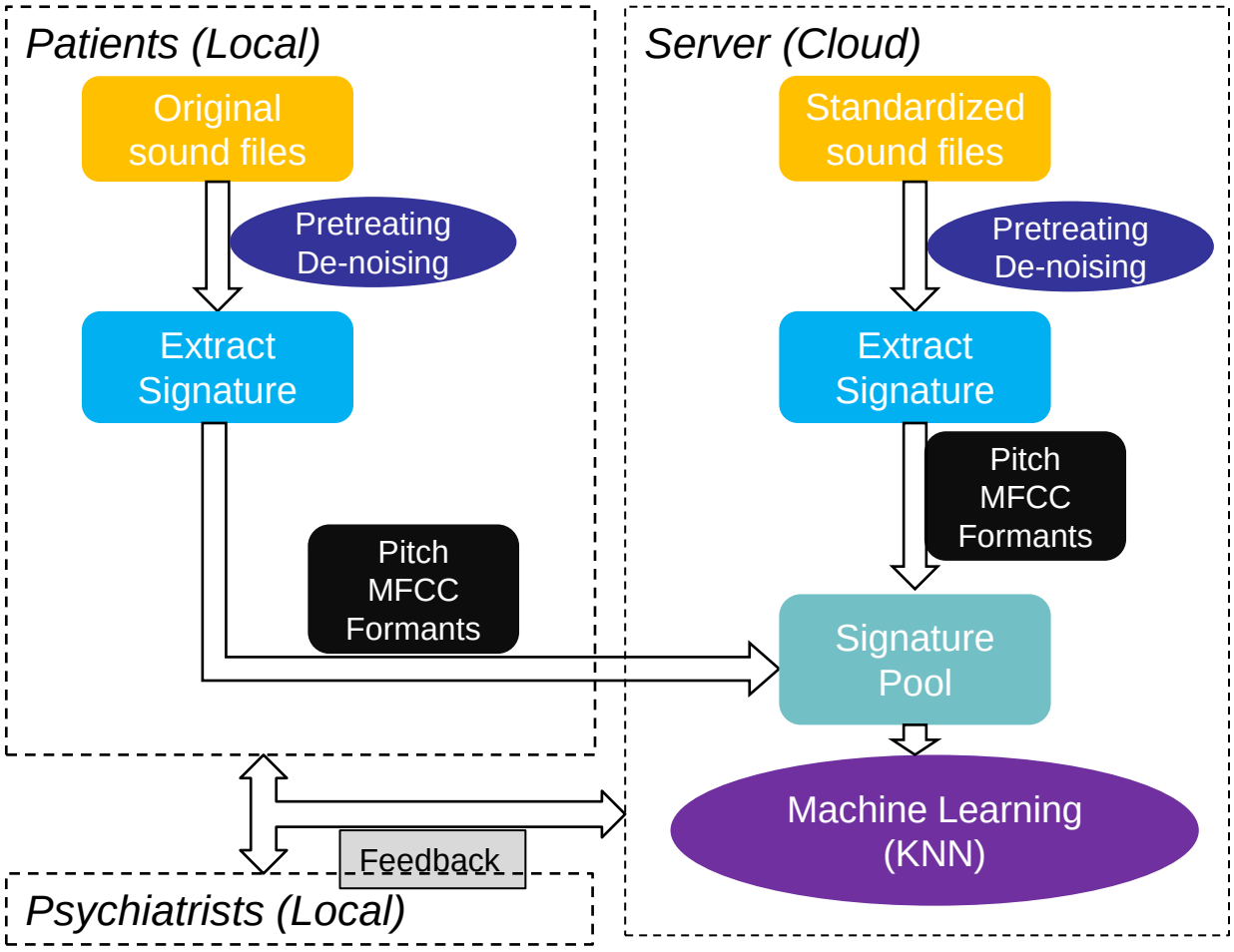
Team #2

Github Link: https://github.com/tungchihwei/EC601-Mood-Detect

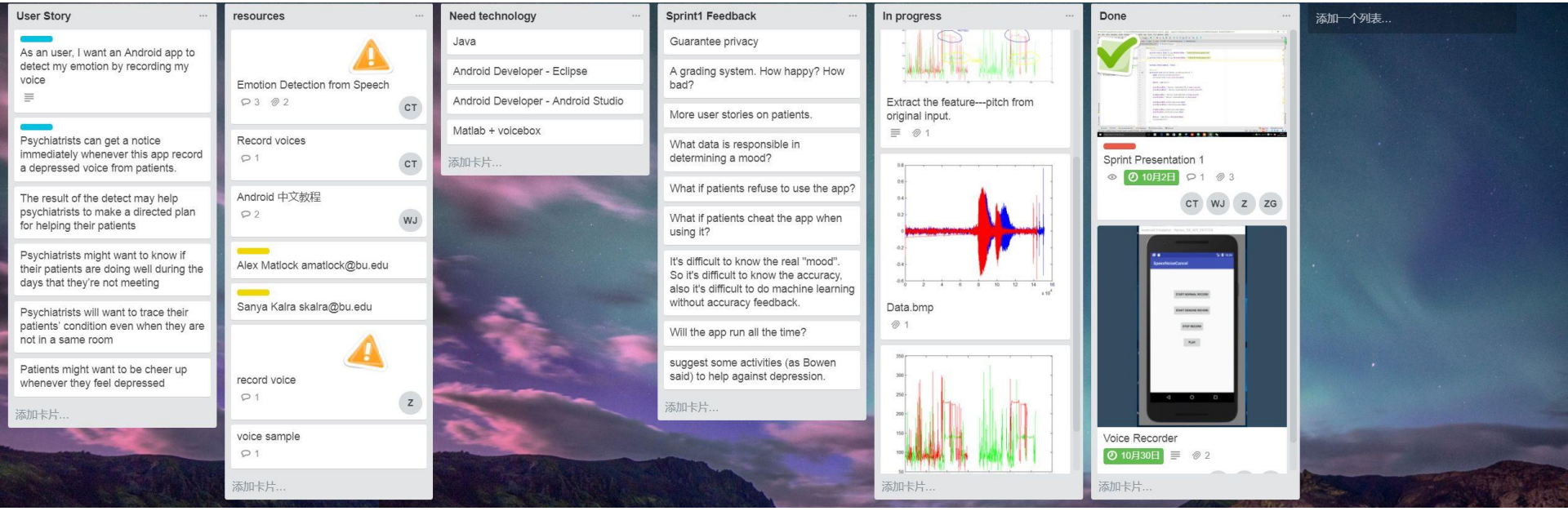
Trello Link: https://trello.com/b/QZmTZ7UV/ec-601-final-project

BOSTON UNIVERSITY

System Diagram



Trello Diagram



Accomplishments #1

Theoretical Background

- ✓ *Pitch*, *Mel Frequency Cepstral Coefficients(MFCCs)* and *Formants* are extracted from human speech as features influence the recognition of emotions. [1][2][3]
- Pitch is extracted from the speech waveform using *RAPT algorithm* implemented in the *VOICEBOX* toolbox.[4]
- MFCC parameters are calculated by *STFT* and *DCT* algorithm.[5]
- Using *Linear Predictive Coding*(LPC) to model formants.[6]

- ✓ *K-Means Clustering* and *Support Vector Machines*(SVM) [7]are adopted to classify opposing emotions.

All Speakers						
Experiment	Features	Distance Measure	Centroid	Iterations	Recognition Accuracy	Variance
despair-elation	MFCC	L1 norm	UDC	100	75.76%	1.74%
happy-sadness	MFCC	L1 norm	UDC	1	77.91%	14.34%
interest-boredom	Pitch	L1 norm	UDC	100	71.21%	2.48%
shame-pride	MFCC	L1 norm	UDC	1	73.15%	3.23%
hot anger-elation	MFCC	L1 norm	UDC	1	69.70%	10.75%
cold anger-sadness	MFCC	L1 norm	UDC	1	75.66%	3.35%
Male Speakers						
Experiment	Features	Distance Measure	Centroid	Iterations	Recognition Accuracy	Variance
despair-elation	MFCC & Pitch	Correlation	UDC	1	87.80%	0.14%
happy-sadness	MFCC	L1 norm	UDC	1	88.80%	3.66%
interest-boredom	MFCC & Pitch	Cosine	Random	100	81.20%	6.36%
shame-pride	MFCC & Pitch	Correlation	UDC	1	74.24%	15.53%
hot anger-elation	MFCC	L1 norm	UDC	1	65.89%	14.95%
cold anger-sadness	MFCC	L1 norm	UDC	1	88.43%	9.78%
Female Speakers						
Experiment	Features	Distance Measure	Centroid	Iterations	Recognition Accuracy	Variance
despair-elation	MFCC	L1 norm	UDC	1	80.42%	9.66%
happy-sadness	MFCC	L1 norm	UDC	1	72.80%	15.24%
interest-boredom	MFCC	L1 norm	UDC	1	70.62%	18.06%
shame-pride	MFCC	L1 norm	UDC	1	81.18%	19.79%
hot anger-elation	MFCC	L1 norm	UDC	1	77.16%	4.37%
cold anger-sadness	MFCC	Correlation	UDC	1	72.04%	15.00%

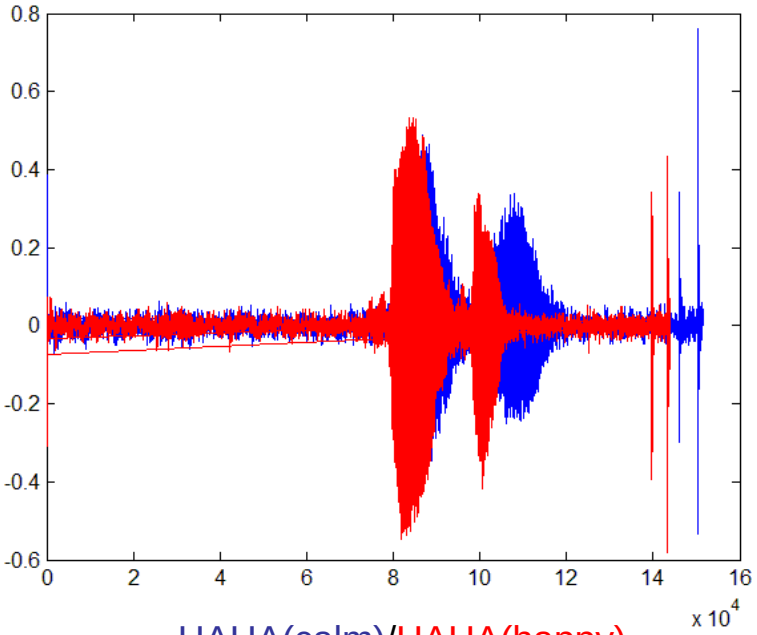
Table 2: Highest Recognition Accuracies using K-means Clustering

Accomplishments #2

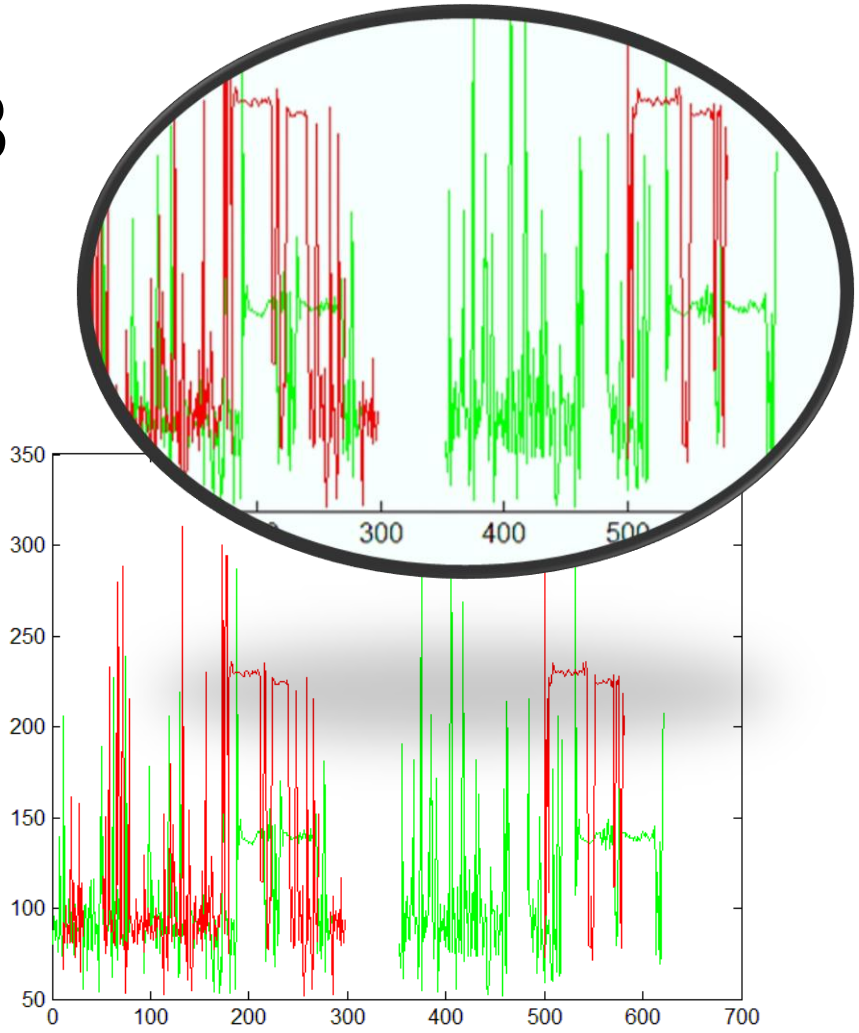
- De-noise Recorder
  - ✓ Build an demo to achieve voice recording
  - ✓ Noise elimination

Accomplishments #3

- Pitch Extraction



HAHA(calm)/HAHA(happy)



Pitch(calm)/Pitch(happy)

Pitch.m

```
% Copyright 2017 Zhonghao Guo gzh1994@bu.edu
% Pitch Extraction by Voicebox in MATLAB

clear all
clc
filename1 = 'd:/calm.wav';
filename2 = 'd:/happy.wav';
[Data1,Fs1]=wavread(filename1);
[Data2,Fs2]=wavread(filename2);
x1=[1:151423,1];
x2=[1:143807,1];
plot(x1,Data1,'g',x2,Data2,'r');
[fx1,tt1] = fxrapt(Data1,Fs1,'u',50);
[fx2,tt2] = fxrapt(Data2,Fs2,'u',50);
x3=[1:684,1];
x4=[1:649,1];
figure(2)
plot(x3,fx1,'g',x4,fx2,'r');

fx1_adjust = fx1(~isnan(fx1));
fx2_adjust = fx2(~isnan(fx2));
fx1_mean = mean(fx1_adjust);
fx2_mean = mean(fx2_adjust);
fx1_max = max(fx1_adjust);
fx2_max = max(fx2_adjust);
fx1_min = min(fx1_adjust);
fx2_min = min(fx2_adjust);
fx1_variance = fx1_max-fx1_min;
fx2_variance = fx2_max-fx2_min;
fx1_midian = median(fx1_adjust);
fx2_midian = median(fx2_adjust);

fx1_derivative = diff(fx1_adjust);
fx2_derivative = diff(fx2_adjust);

fx1_energy_avg = mean(Data1);
fx2_energy_avg = mean(Data2);
speaking_rate1 = length(fx1)/length(fx1_adjust);
```

- For the pitch vector and its derivative:
  - *Mean*
  - *Median*
  - *Maximum*
  - *Minimum*
  - *Variance*

- Average energy of speech

- Speaking rate
  - *Inverse of the average length of the voiced part of the utterance*

Hence, the pitch feature vector is 12-dimensional

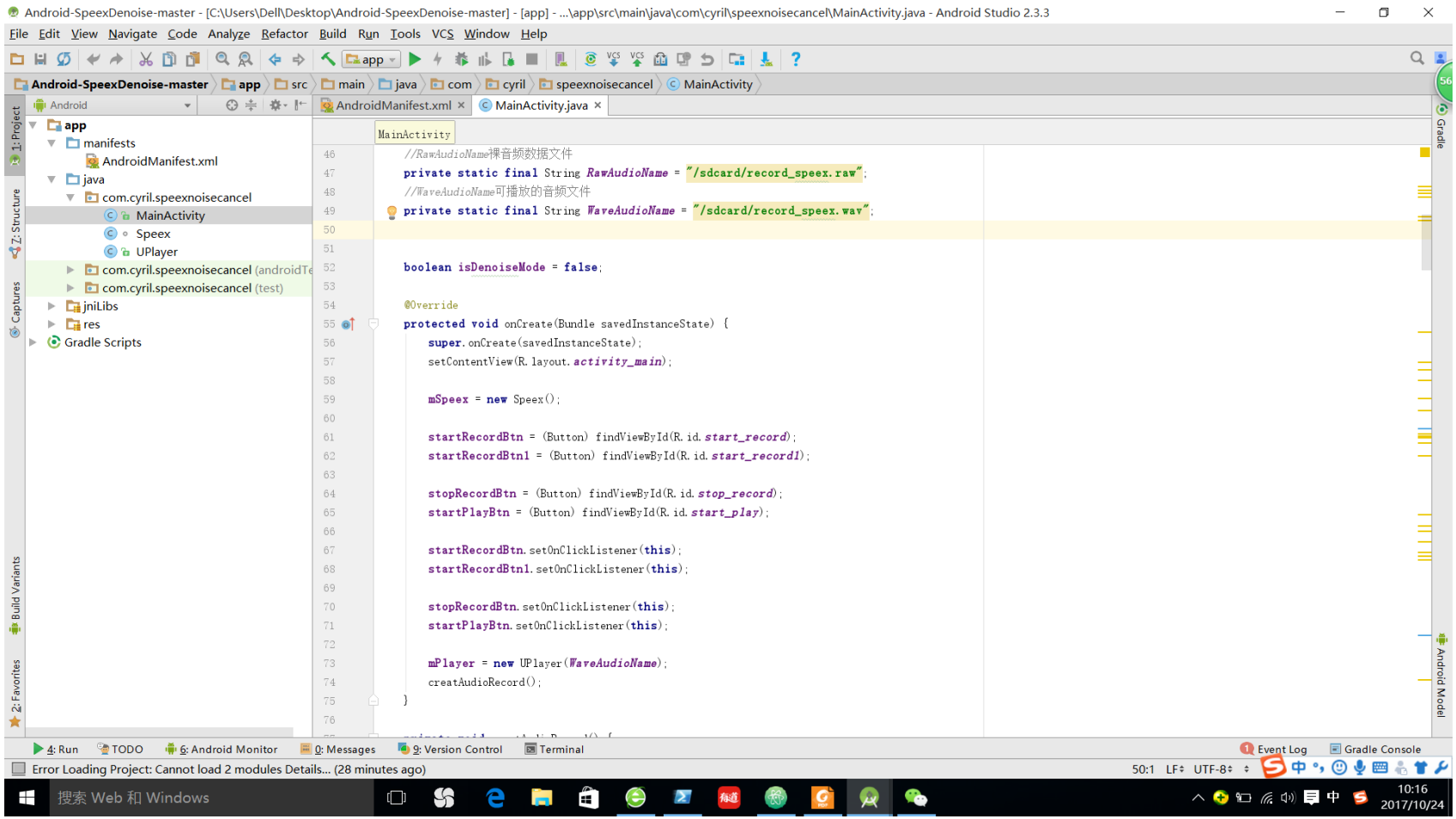
Accomplishments #4

- Emotional Prosody Speech and Transcripts[8]

URL:https://catalog.ldc.upenn.edu/LDC2002S28

- The recordings consist of professional actors reading a series of semantically neutral utterances (dates and numbers) spanning fourteen distinct emotional categories.
- The content and emotion of each word and phrase is recorded in the matching transcript.

DEMO


Android-SpeexDenoise-master - [C:\Users\Dell\Desktop\Android-SpeexDenoise-master] - [app] - ...\app\src\main\java\com\cyril\speexnoisecancel\MainActivity.java - Android Studio 2.3.3
File Edit View Navigate Code Analyze Refactor Build Run Tools VCS Window Help
Android-SpeexDenoise-master app src main java com cyril speexnoisecancel MainActivity
AndroidManifest.xml MainActivity.java
app
manifests
AndroidManifest.xml
java
com.cyril.speexnoisecancel
MainActivity
Speex
UPlayer
com.cyril.speexnoisecancel (androidTe
com.cyril.speexnoisecancel (test)
jniLibs
res
Gradle Scripts
//RawAudioName裸音频数据文件
private static final String RawAudioName = "/sdcard/record_speex.raw";
//WaveAudioName可播放的音频文件
private static final String WaveAudioName = "/sdcard/record_speex.wav";
boolean isDenoiseMode = false;
@Override
protected void onCreate(Bundle savedInstanceState) {
super.onCreate(savedInstanceState);
setContentView(R.layout.activity_main);
mSpeex = new Speex();
startRecordBtn = (Button) findViewById(R.id.start_record);
startRecordBtn1 = (Button) findViewById(R.id.start_record1);
stopRecordBtn = (Button) findViewById(R.id.stop_record);
startPlayBtn = (Button) findViewById(R.id.start_play);
startRecordBtn.setOnClickListener(this);
startRecordBtn1.setOnClickListener(this);
stopRecordBtn.setOnClickListener(this);
startPlayBtn.setOnClickListener(this);
mPlayer = new UPlayer(WaveAudioName);
creatAudioRecord();
}
4: Run TODO 6: Android Monitor 0: Messages 9: Version Control Terminal
Event Log Gradle Console
Error Loading Project: Cannot load 2 modules Details... (28 minutes ago)
50:1 LF UTF-8
搜索 Web 和 Windows
10:16
2017/10/24

DEMO

Normal Recorder

Denoise Recorder

Challenges

- User story——What if customers “cheat” the app when recording voice or even refuse to use it when feeling down?
- Privacy Issue——If monitoring constantly, will the app offense user’s privacy?
- Feedback Fraud——How to make sure the feedback from patients and psychiatrists are real? We may not get an objective training data set from subjective feedback.
- Technical Challenges——How to transplant the RAPT toolbox from MATLAB to JAVA?

Next Sprint Plan

- Extract signatures from original sound samples
- Create multi-dimensional feature vectors from signatures acquired
- Set up a dataspace to storage statistics
- Upload all feature files to dataspace for Machine Learning

Reference

- [1] Shah Hewlett, Emotion Detection From Speech, CS229, 2007.
- [2] Garima Vyas, Malay Kishore Dutta, Automatic Mood Detection of Indian music Using MFCCs and K-means Algorithm, 2014.
- [3] Sharifa Alghowinem, Roland Goecke, Michael Wagner, Julien Epps, Michael Breakspear, gordon Parker, From Joyous to Clinically Depressed: Mood Detection Using Spontaneous Speech, 2012.
- [4] http://www.ee.ic.ac.uk/hp/staff/dmb/voicebox/voicebox.html
- [5] http://www.ee.columbia.edu/~dpwe/resources/matlab/rastamat/
- [6] L.R.Rabiner and B.H.Juang. “Fundamentals of Speech Recognition”, Upper Saddle River; NJ: Prentice-Hall, 1993
- [7] http://ida.first.fraunhofer.de/~anton/software.html
- [8] https://catalog.ldc.upenn.edu/LDC2002S28

THANK YOU