

Singing Data Labeling Tool
Milestone 6 Progress Evaluation Report

Group members:

Nandith Narayan nnarayan2018@my.fit.edu

Avinash Persaud apersaud2018@my.fit.edu

Carlos Cepeda ccepeda2018@my.fit.edu

Advisor:

Dr. William Shoaff wds@fit.edu

Client:

Caleb Matthew Long, Appalachian State University

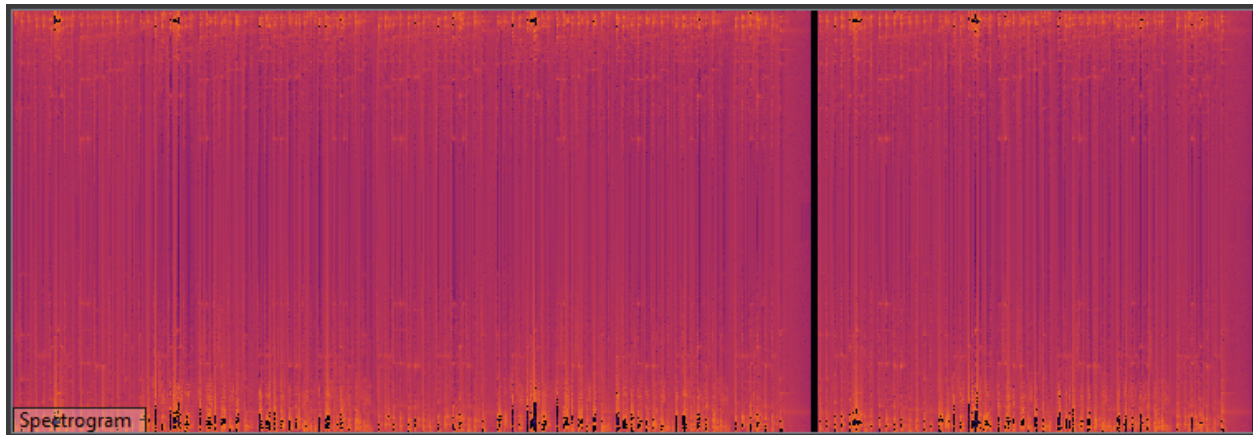
Task	Completion %	Nandith Narayan	Avinash Persaud	Carlos Cepeda	To Do
Integrated Machine Learning Models	100%	100%			
Spectrogram Update	90%		90%		Need to switch to asynchronous processing
Added the ability to edit phoneme values.	100%	80%	20%		
Audio Playback			10%		Implementation pending
Implement Remaining Track Types	100%		100%		
Improved UI	100%	90%	10%		
Generate Output file.	100%	80%	20%		
Note UI	0%				

Task Discussion:

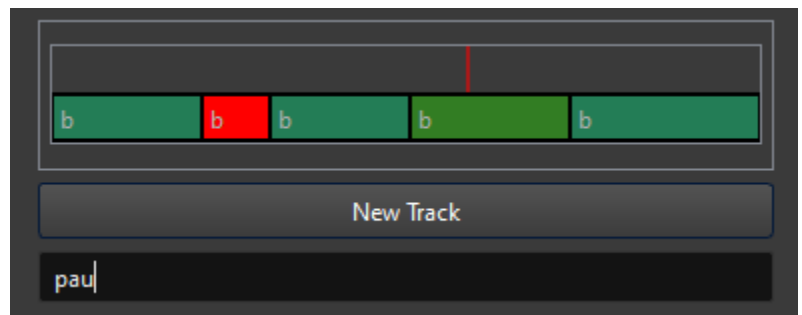
Task 1: Integrating the machine learning models into the application was a tough task. We explored multiple methods of integrating our models into the application. We ended up writing a python script to run our models from the application. We used an `exec()` function call to run the python interpreter from the application on a different process. The main application sends the path of the audio file as well as the timestamps of every phoneme to the python script. The python script then does the following:

- 1) Load the audio data.
- 2) Slice the audio data into sections according to the given timestamps.
- 3) Compute the spectrogram of the given section of audio.
- 4) Preprocess this spectrogram by rescaling, normalizing, and duplicating the spectrograms. By repeating features, we found that we could use smaller networks.
- 5) Run the first model to decide whether a given phoneme is a vowel, consonant, or pause.
- 6) Run the hierarchy of models based on the results of the previous step.
- 7) Output the resulting classifications back to the main application.

Task 2: We switched from a hand implemented fft to `fftw3`. This has significantly improved spectrogram generation time. This increase of spectrogram generation speed cannot be understated; We went from taking minutes to generate the spectrogram of a 30 second long audio file to fractions of a second to generate the same spectrogram.

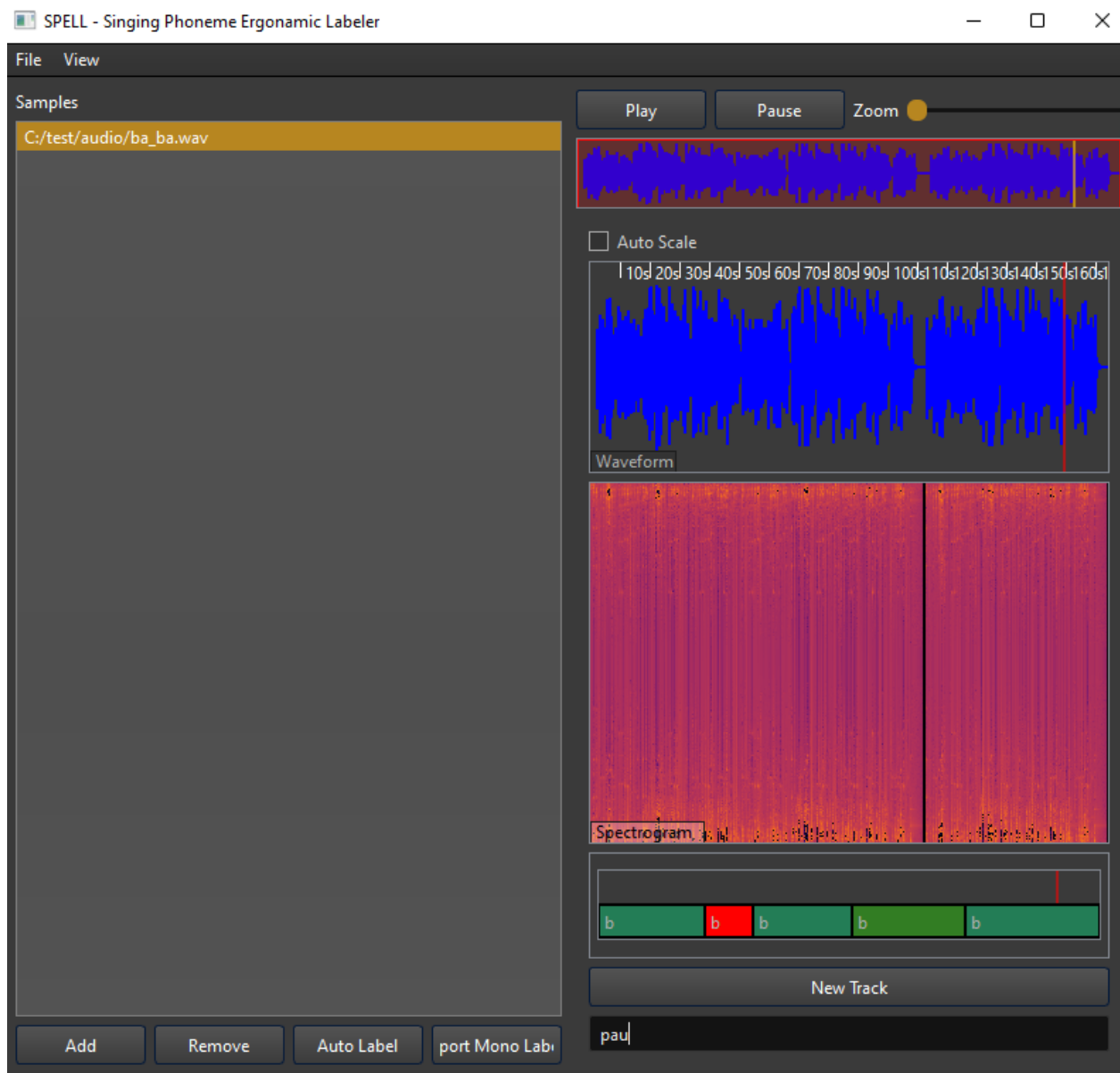


Task 3: We added the ability to directly type in the value of any phoneme. Now, the user can select a phoneme to edit by using the arrow keys. We also added a text input field below the timeline view. This allows the user to type in an arbitrary value for the selected phoneme. Upon pressing enter, the selected phoneme is modified with the value that the user entered.



Task 4: Using port audio, playback of the selected and/or visible portion of the wav file will occur.

Task 5: We polished the UI in a number of ways. We added buttons to perform various tasks like exporting the labeled data and running our machine learning models. We added a highlight to the phoneme that is currently selected. We added a text input field to allow the user to edit phonemes. We added shortcuts for the user; The user can change which phoneme is selected using the arrow keys. The user can cancel a phoneme move/edit operation by right clicking. The user can click on the top section of the timeline view to select it. The user is presented with a file save/open dialogue box when such operations take place.



(Note: the widgets are collapsible)

Task 6: We added the ability to export all the phonemes that the user has labeled as mono-labels. Upon clicking the export button, the user is prompted to select a file location, and then the labeled data is exported into a '.lab' mono-label file.

Client Feedback

Discussed with the client over discord occasionally.

Date of meeting with Faculty Advisor: **4/18/22**

Faculty Advisor feedback on milestone tasks:

- Evaluation by Faculty Advisor
- Faculty Advisor: detach and return this page to Dr. Chan (HC 214) or email the scores to pkc@cs.fit.edu
- Score (0-10) for each member: circle a score (or circle two adjacent scores for .25 or write down a real number between 0 and 10)

Carlos Cepeda	0	1	2	3	4	5	5.5	6	6.5	7	7.5	8	8.5	9	9.5	10
Avinash Persaud	0	1	2	3	4	5	5.5	6	6.5	7	7.5	8	8.5	9	9.5	10
Nandith Narayan	0	1	2	3	4	5	5.5	6	6.5	7	7.5	8	8.5	9	9.5	10

■ Faculty Advisor Signature: _____ Date: _____