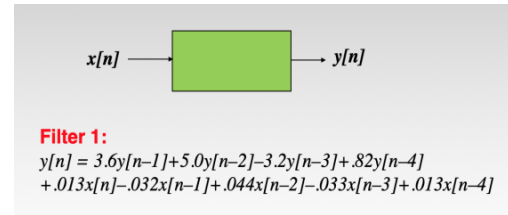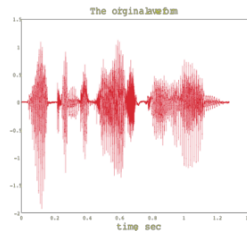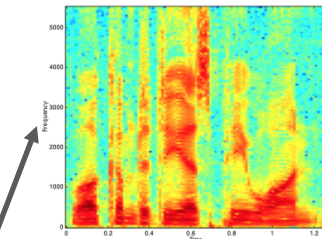# Earworm Design

Team C6:
Wenting Chang
Nolan Hiehle
Anja Kalaba

Application Area

Vocal Input → Song Identification

Sing into a microphone a chunk of a selected piece that is contained in a database and have it correctly identified.

Filter 1:
$$y[n] = 3.6y[n-1]+5.0y[n-2]-3.2y[n-3]+.82y[n-4]$$
$$+.013x[n]-.032x[n-1]+.044x[n-2]-.033x[n-3]+.013x[n-4]$$

Solution Approach

- Multi-step algorithm inspired by
  - **A Comparative Evaluation of Search Techniques for Query-by-Humming Using the MUSART Testbed**
  - Roger B. Dannenberg, William P. Birmingham, Bryan Pardo, Ning Hu, Colin Meek, George Tzanetakis
- Implemented in C -- fast
- Add a data visualization component -- explanation for user
- Potential added robustness -- Convolutional Neural Network membership to cross verify
- Package this as an app
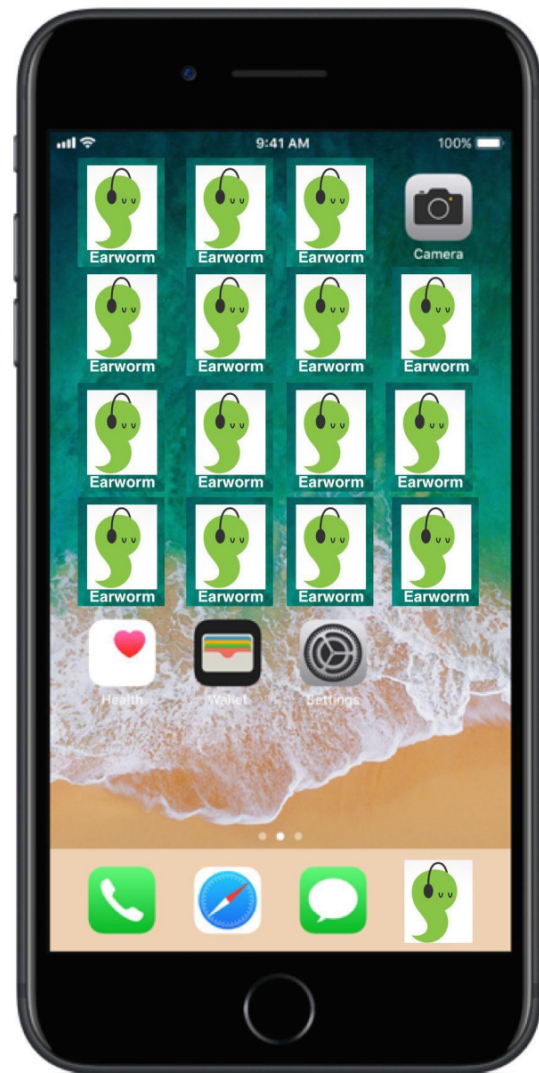
System Specification
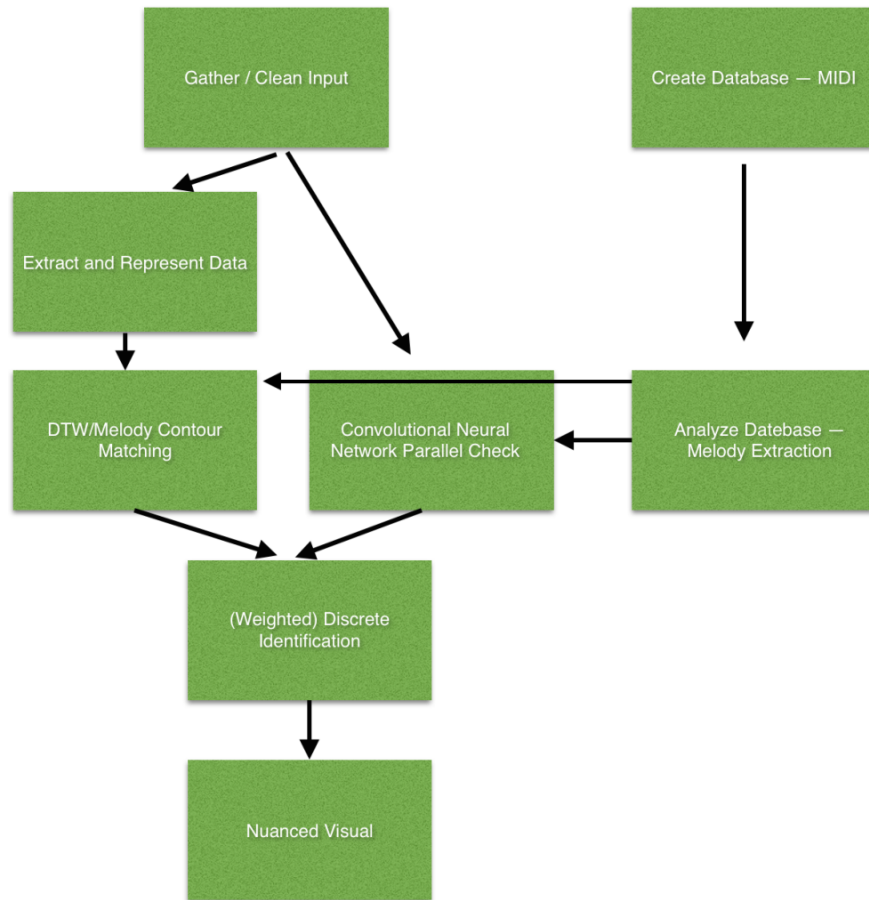
System Specification

- Gather Input

Audio Input
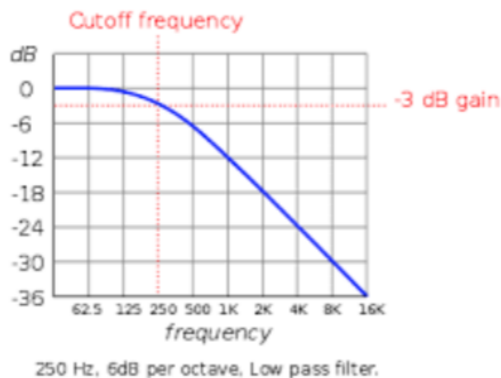
Filtering

250 Hz, 6dB per octave, Low pass filter.

Cutoff frequency

-3 dB gain

Pitch/Note Extraction

Data Representation

	Absolute Pitch:	Relative Pitch:	IOI:	IOI Ratio:	Log IOI Ratio:

Absolute Pitch: 67   69   71   67

Relative Pitch: 2   2   −4

IOI: 1   0.5   0.5   1

IOI Ratio: 0.5   1   2

Log IOI Ratio: −1   0   1
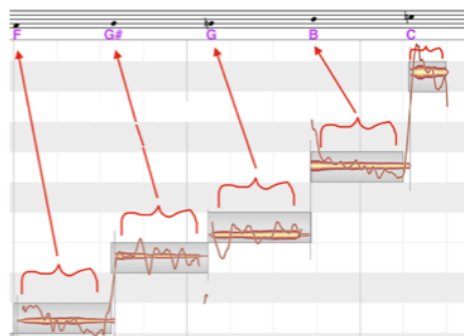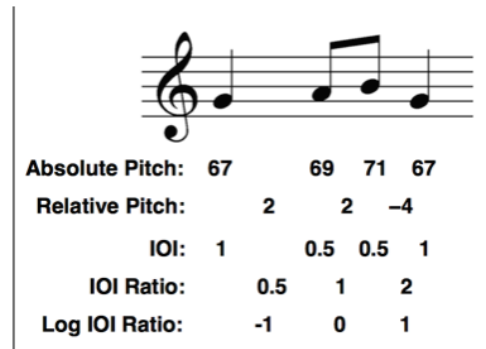
$<<2, -1>, <2, 0>, <-4, 1>>.$

System Specification

- Gather backend and combine with input in analysis to give overall match and visualization

MIDI

DTW/Melody Contour

Mel-spectrogram

Convolutional Neural Net

Convolutions

Max-pooling

Convolutions + Max-pooling

Fully-connected

Output

SONG MATCH

DATA VISUALIZATION
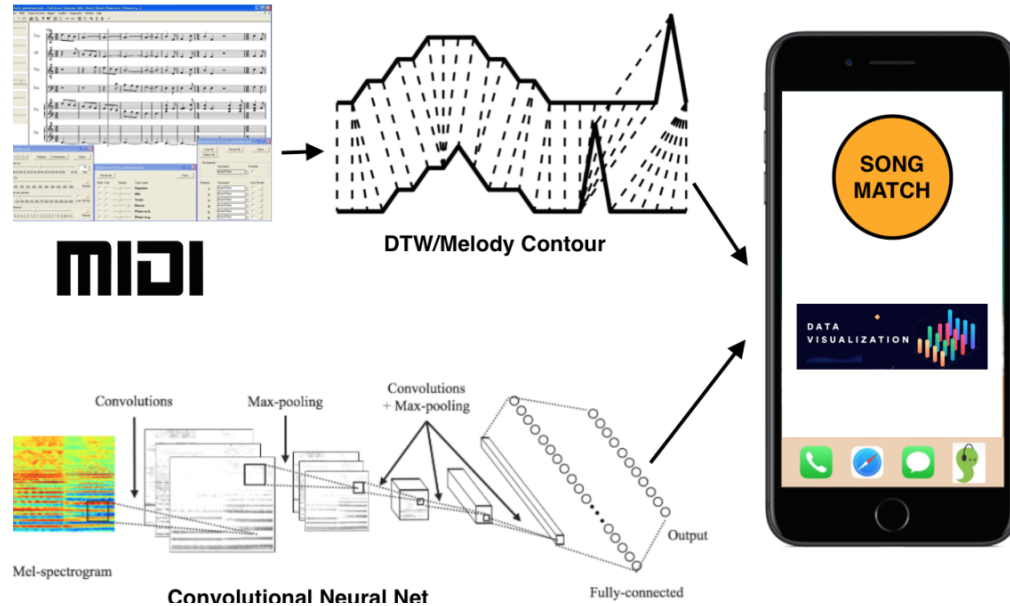
Implementation Plan

- Query By Humming Implementation
- Us:
  - Audio → pitch conversion / processing → note determination
  - Data representation
  - Melodic Contour / DTW
  - Data Visualization
  - (potential) Convolutional Neural Network
- Borrowed:
  - Database creation → MIDI Files
  - MIDI File processing → Melody extraction - *ThemeExtractor* (Meek & Birmingham, 2001)

Metrics and Validation

- **Accuracy** across dimensions: singer, song (song type), time frame, etc.

	Singer_1	Singer_2	Singer_3	Singer_k
Song_1	★	★	✘	✘
Song_2	★	✘	★	✘
Song_3	★	✘	✘	✘
Song_n	★	★	★	★

- Average **time** for processing and identification

Metrics and Validation

- Test Outputs:
  - Discrete Yes/No on identification
  - Nuance in Data Visualization

- Risk Correction ≈ Analysis Explanation
  - Data Visualization:
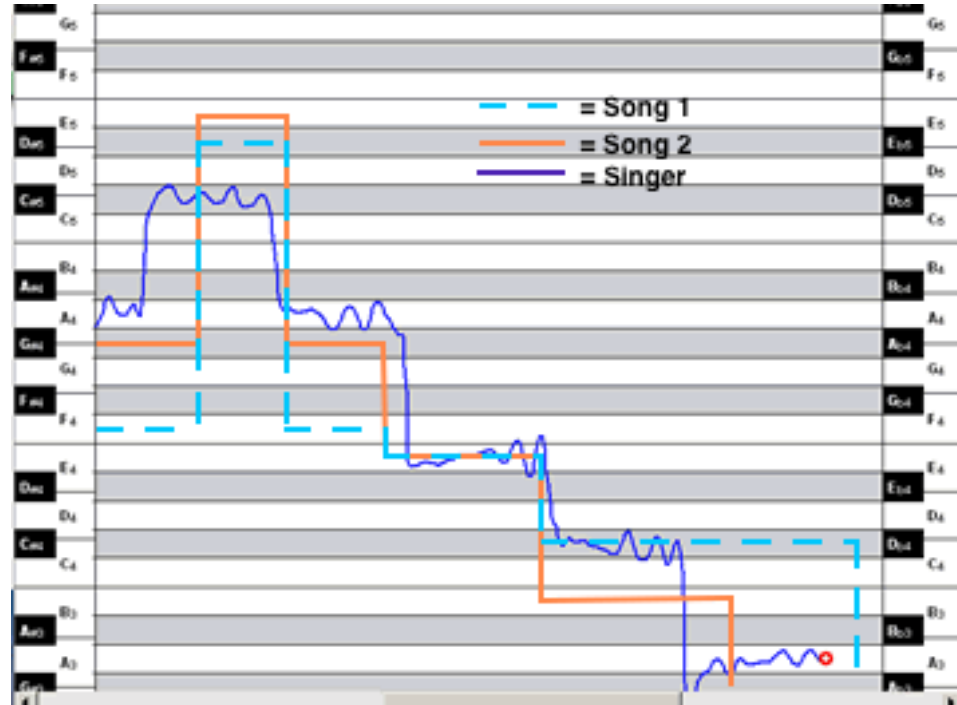
Project Management

All
- Read papers and figure out technique for analysis
- Design data format
- Testing

Anja (Signals)
- Audio processing
    - Polyphonic pitch tracking

Wenting (App + ML)
- Make app
- Work on algorithm with Nolan

Nolan (ML + Matching Algorithm)
- Train library and matching algorithm

Project Management

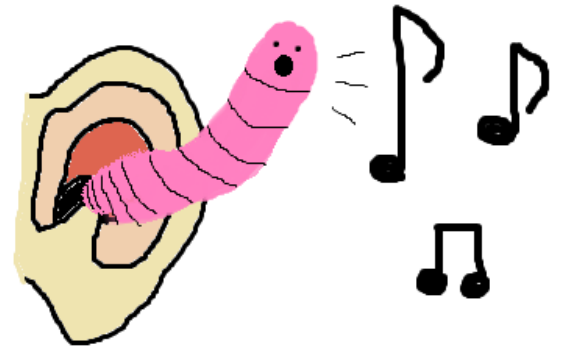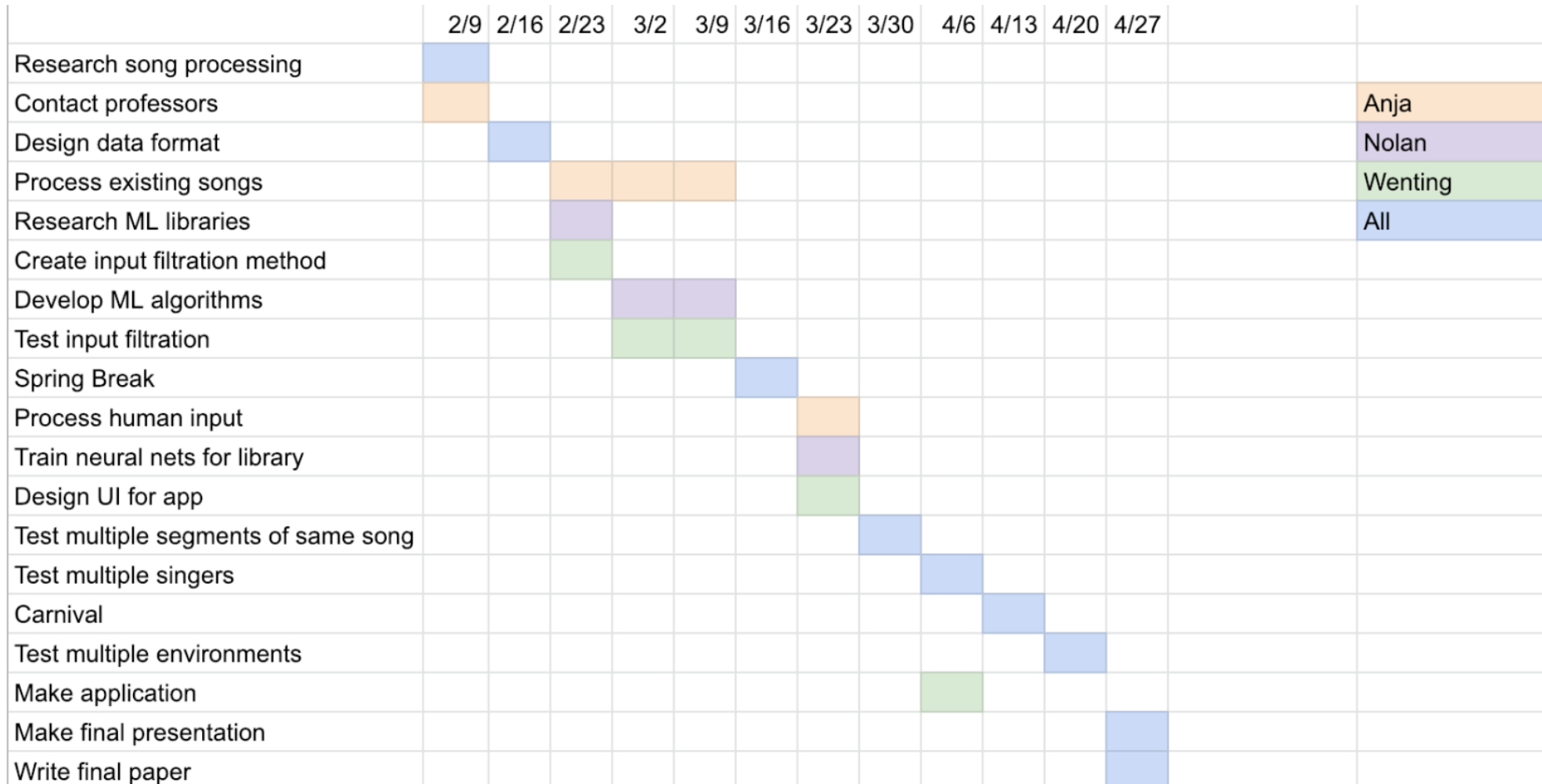| Task	2/9	2/16	2/23	3/2	3/9	3/16	3/23	3/30	4/6	4/13	4/20	4/27
Research song processing	All											
Contact professors	Anja											
Design data format		Nolan										
Process existing songs			Anja	Anja	Anja							
Research ML libraries			Nolan									
Create input filtration method			Wenting									
Develop ML algorithms				Nolan	Nolan							
Test input filtration				Wenting	Wenting							
Spring Break						All						
Process human input							Anja					
Train neural nets for library							Nolan					
Design UI for app							Wenting					
Test multiple segments of same song								All				
Test multiple singers									All			
Carnival										All		
Test multiple environments											All	
Make application									Wenting			
Make final presentation											All	
Write final paper												All

Legend:
- Anja (orange)
- Nolan (purple)
- Wenting (green)
- All (blue)

Conclusion

- Why?
- Marketable?
- Synthesis of research fields for entertainment?