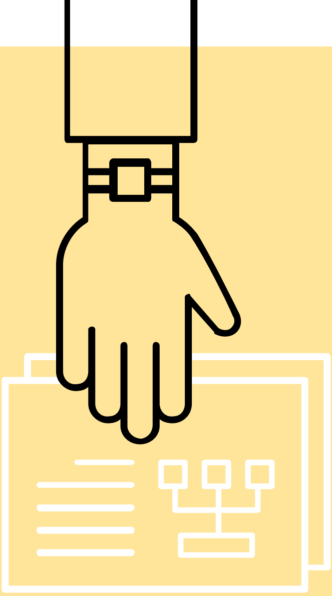
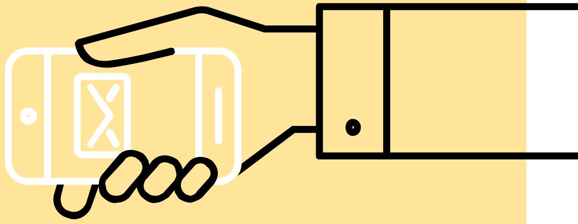




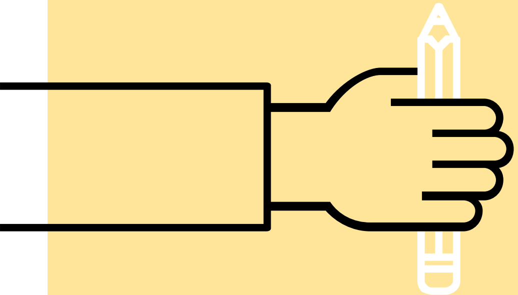
Sprint #0 End

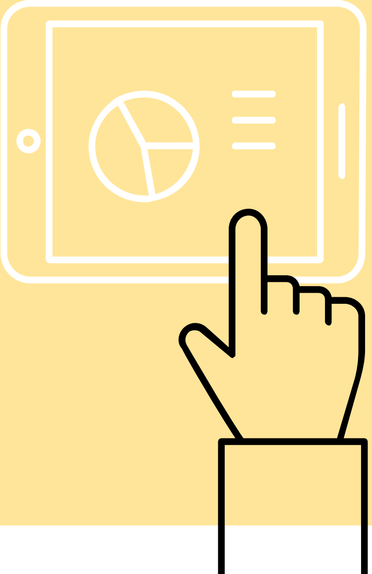
Instrument Recognition Software

Members:

Aleks, Aner, Axel, Cuong

Joe, and Thomas

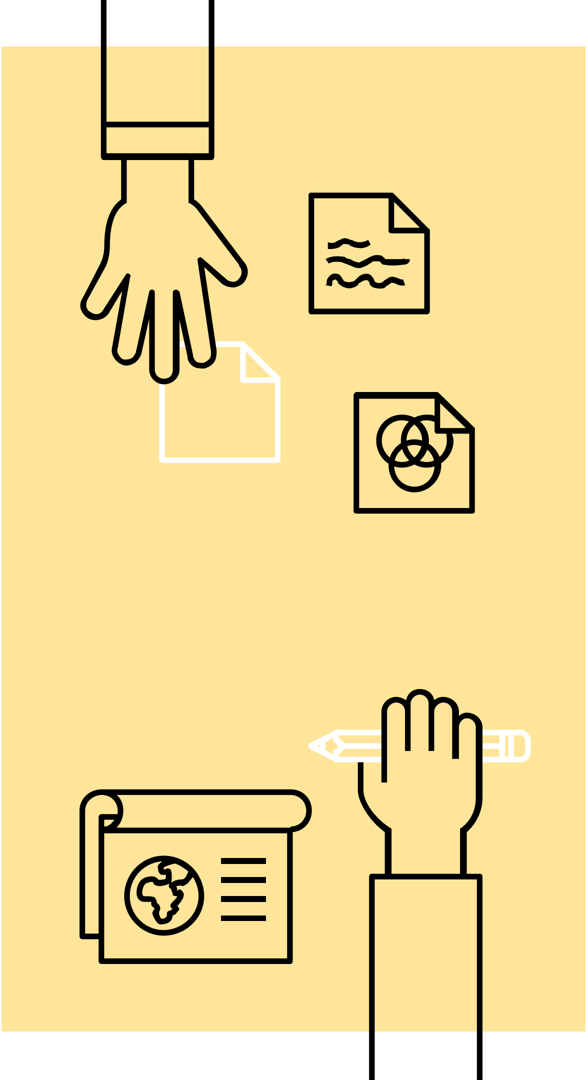




Product overview

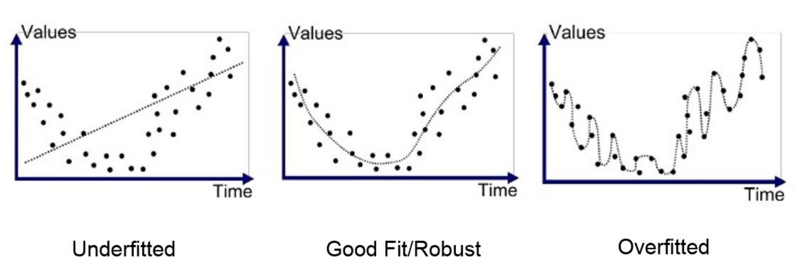
Mission Statement:

The primary goal of the Instrument Recognition Software team is to develop a program that can identify specific instruments within polyphonic audio samples. This currently is not possible, as music must be identified manually or through broader categorization. We wish to implement and evolve machine learning techniques to accomplish our goal and explore a new realm of applications for our new technology.

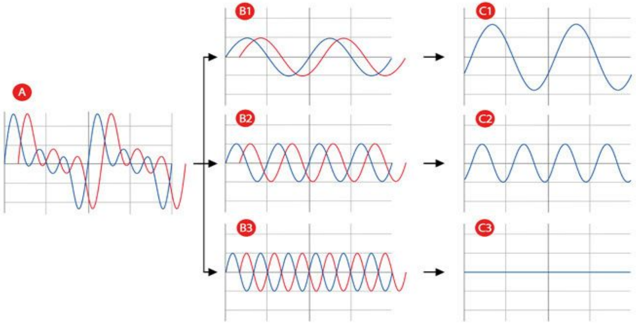


Challenges

- None of us have a background in Machine Learning
- Machine Learning is rarely 100% accurate

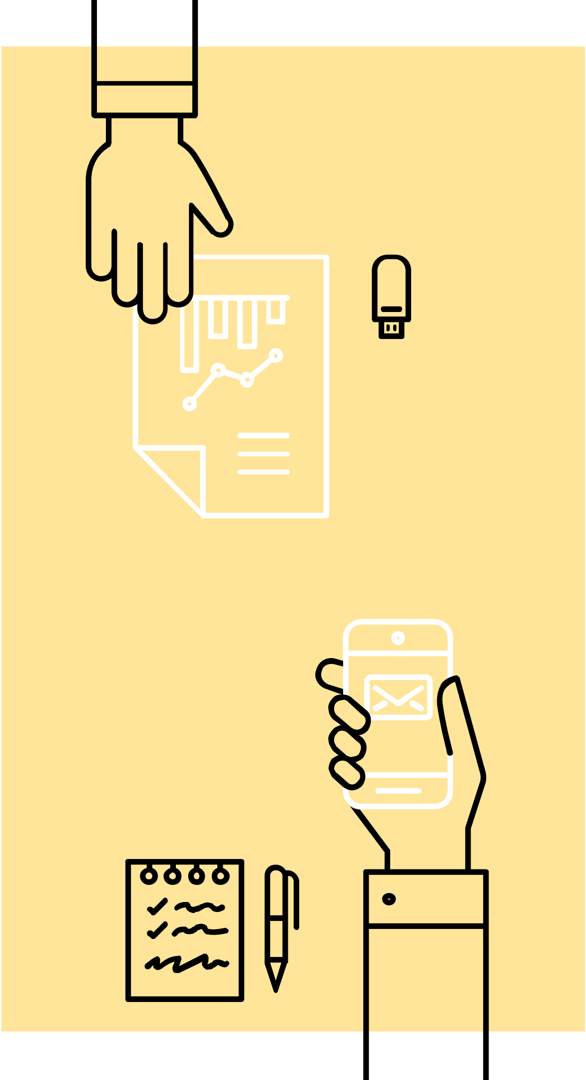


- Instruments are challenging to identify
- Attempting to separate sound sources from a jumbled mess



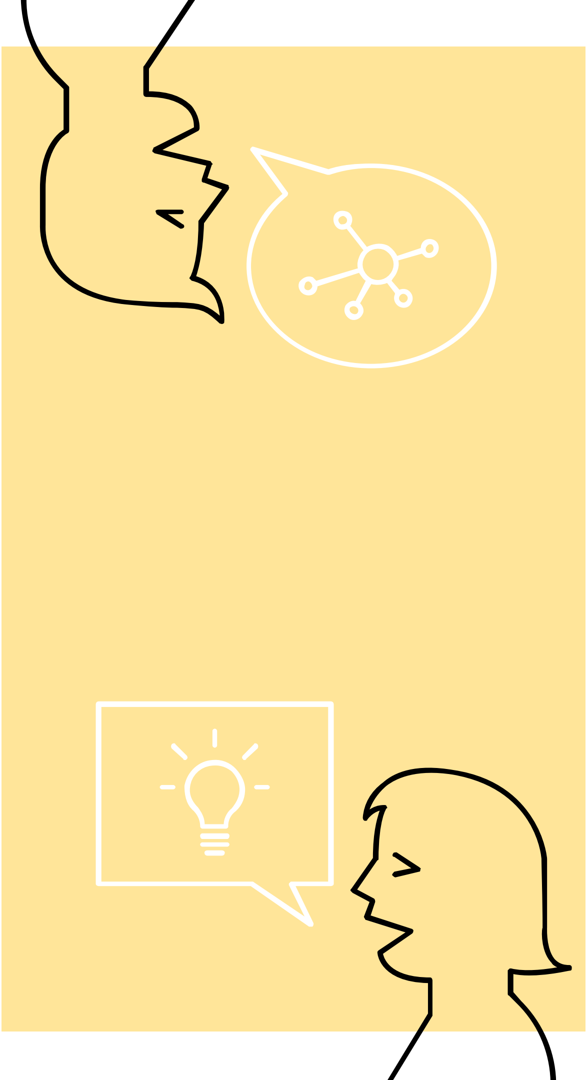
Instrument Recognition Research

- Can identify based on image but not by sound
- 82% accuracy in identifying instruments in polyphonic music (unknown polyphony count)
- 59% accuracy in identifying six note (six instruments) polyphony



Sources of information

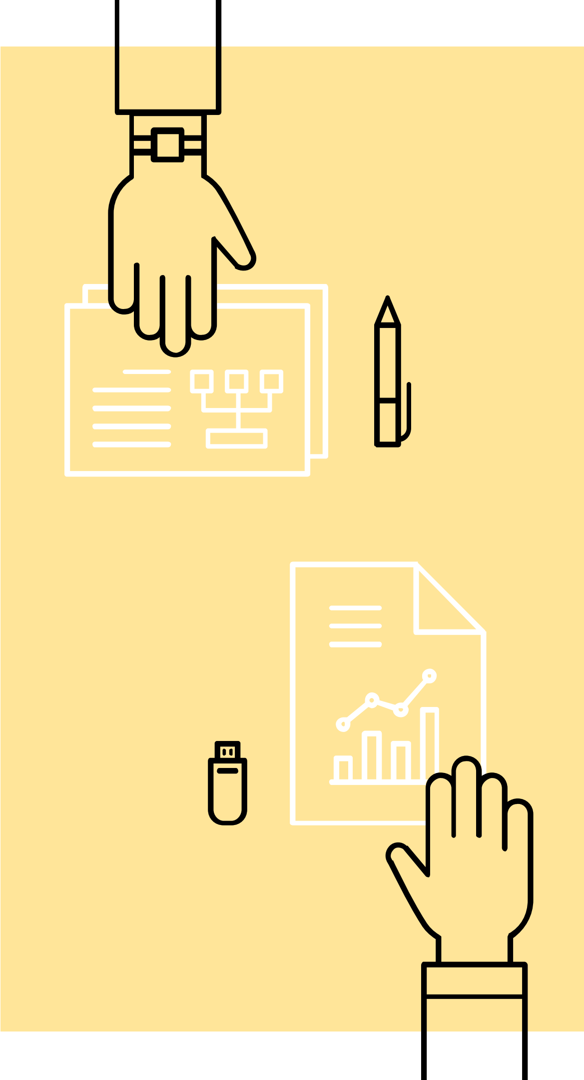
- MIT's PixelPlayer
- NY University research by Peter Li, Jiyuan Qian, and Tian Wang
- Tampere University of Technology research by Toni Heittola, Anssi Klapuri, and Tuomas Virtanen



Monophonic vs Polyphonic Music

Monophonic

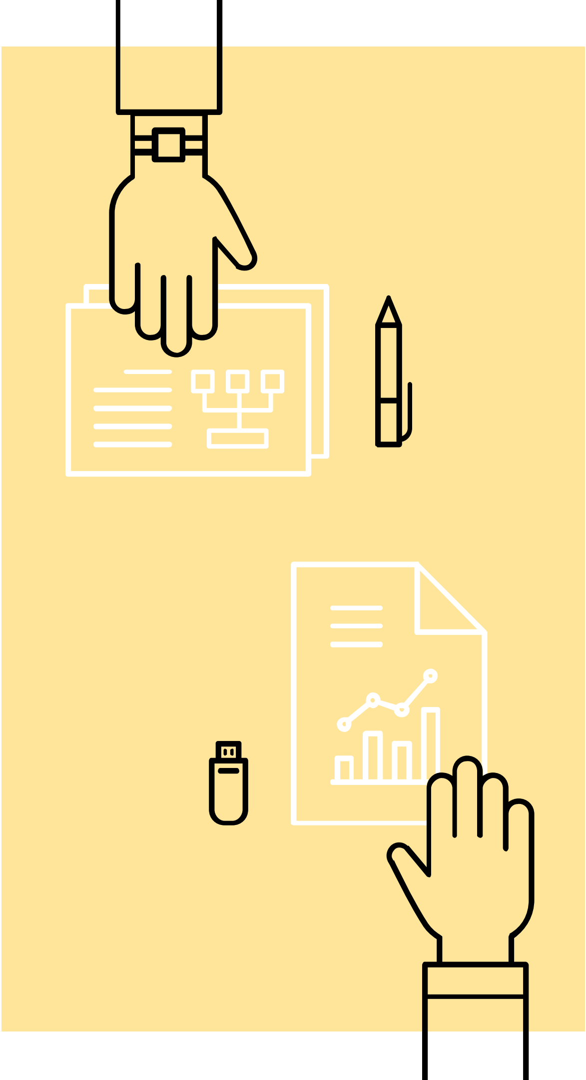
- Single note, but with high fidelity, resulting in a very accurate reproduction of the sound
- Simple
- Timbre features can be easily evaluated since there's only a single note
- Many samples exists, but still incomplete



Monophonic vs Polyphonic Music

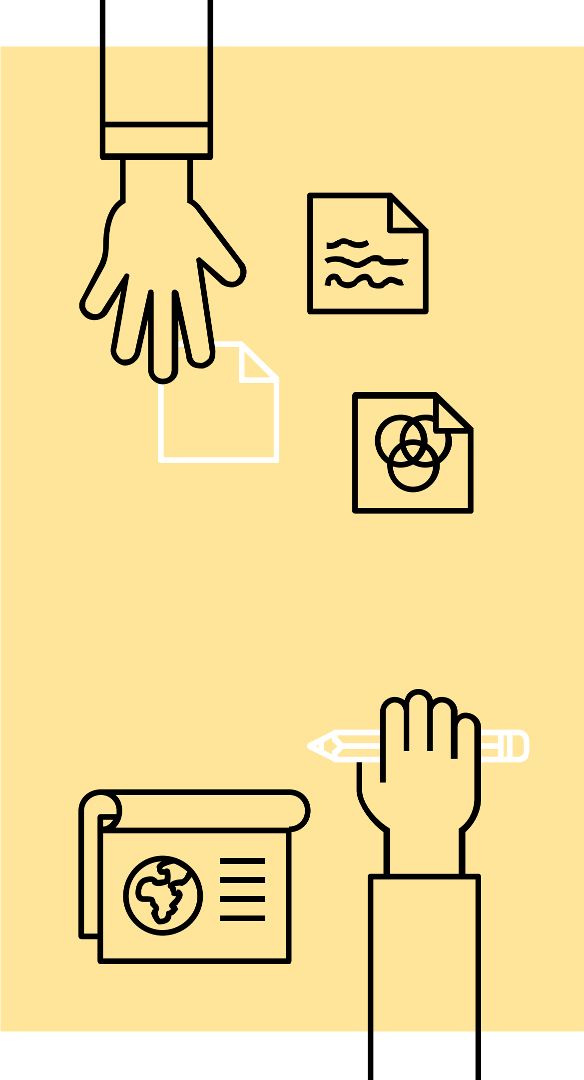
Polyphonic

- Overlapped sounds of different instruments
- Difficult to work with and hard to track. Source separation is needed
- More practical and useful since music is almost always polyphonic
- No good sample collections due to its complexity



Who is our customer and why would they need our product?

- Audiophiles – people that love to listen to music
- Music makers – people that want to find inspiration from others
- Music industry – companies like Sony/Spotify may find this research valuable



Primary Research/User Personas

Cliff Rodgers

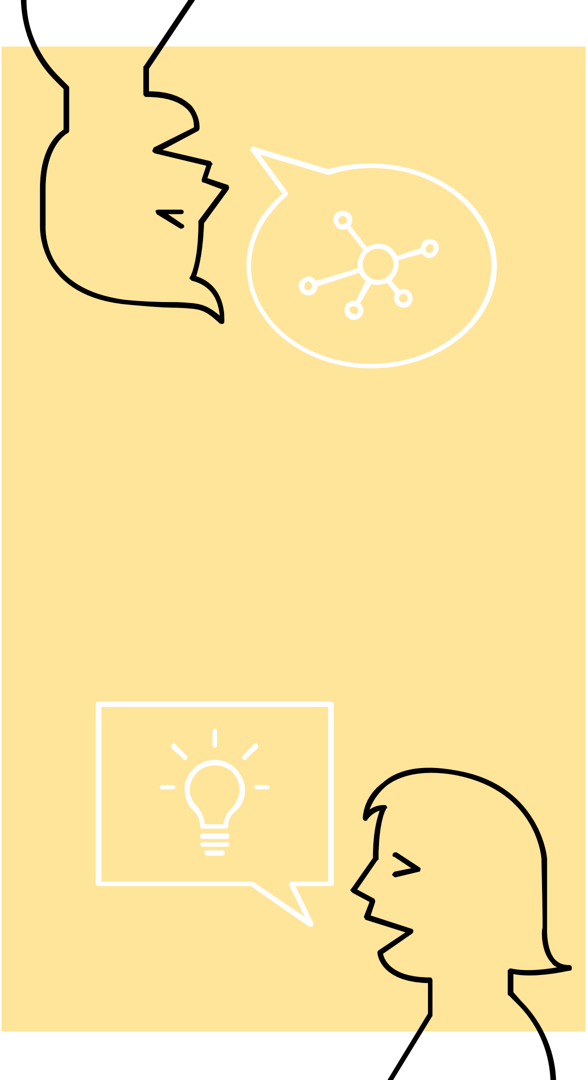
- Self employed DJ

Utilization

- Searching for similarly styled songs
- Smoother transitions between songs

New Features

- Tracking beats-per-minute



Primary Research/User Personas

David Billings - Sony Entertainment

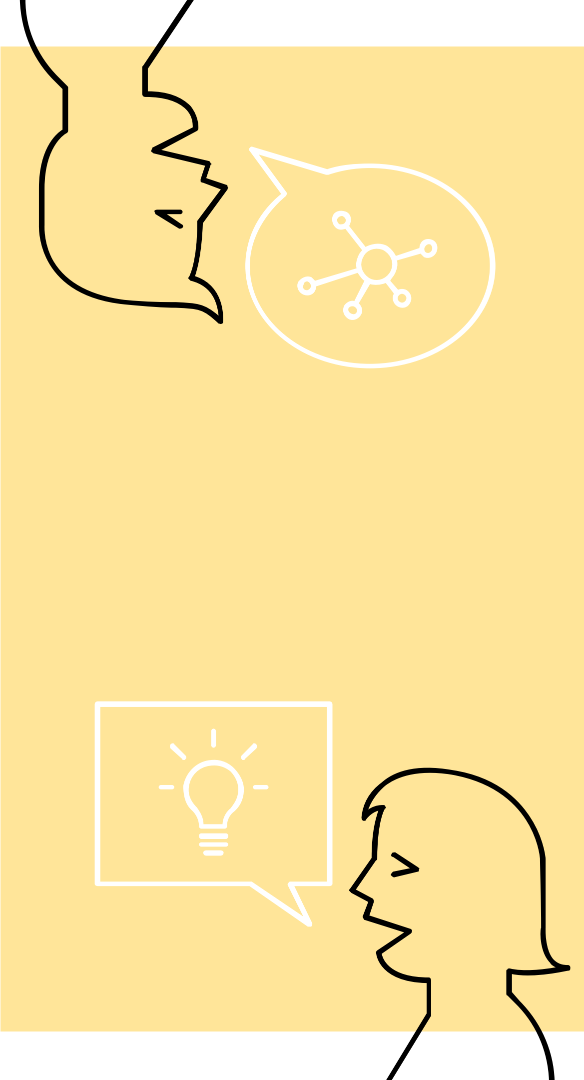
- Sound Engineer

Utilization

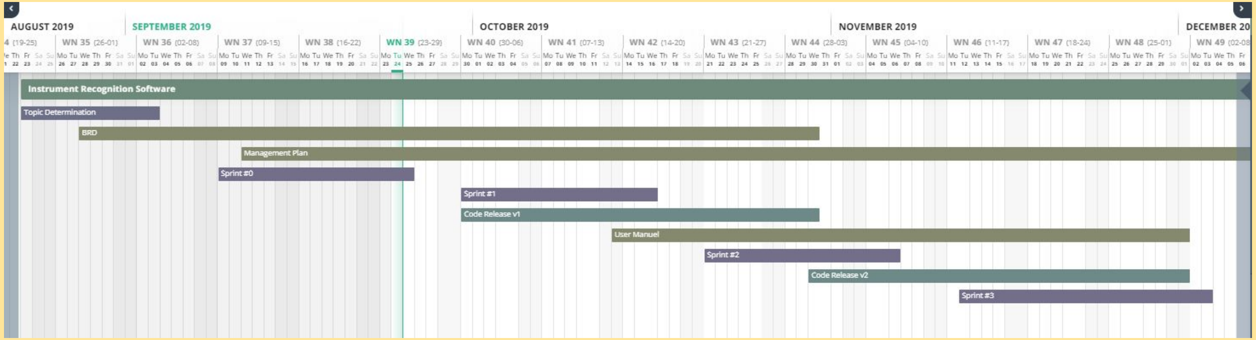
- Sorting sound libraries
- Interest in "dynamic" or "developer" version

New Features

- Ability to add instruments



Research and Development



Phase 0: Prepare

-Sprint 0

- Research
- Determine models
- Document and plan

Phase 1: Model

-Sprints 1 & 2

Demonstrate Models:

- Linear Regression
- Convolutional Neural Networking

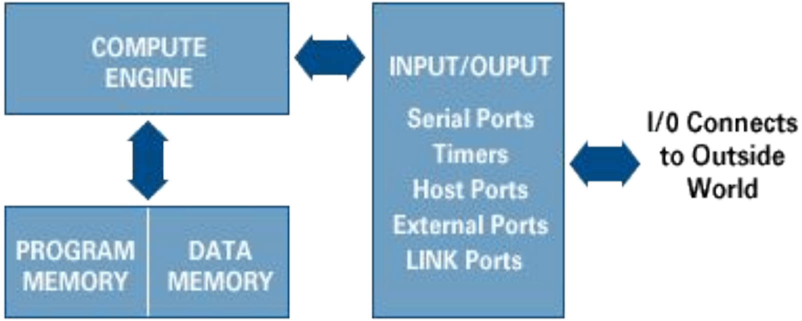
Phase 2: Implement

-Sprint 3

- Implement an algorithm
- Train for categorized instruments

Analog to Digital Audio Conversion

- During the recording phase, analog audio is input through a receiver, usually a microphone
- The analog signal is then converted into a digital signal by an Analog-to-Digital Converter (ADC) and passed to the Digital Signal Processor (DSP)
- The DSP performs the encoding and saves the file to the memory. Sound cards that do not have their own DSP use the CPU for processing.

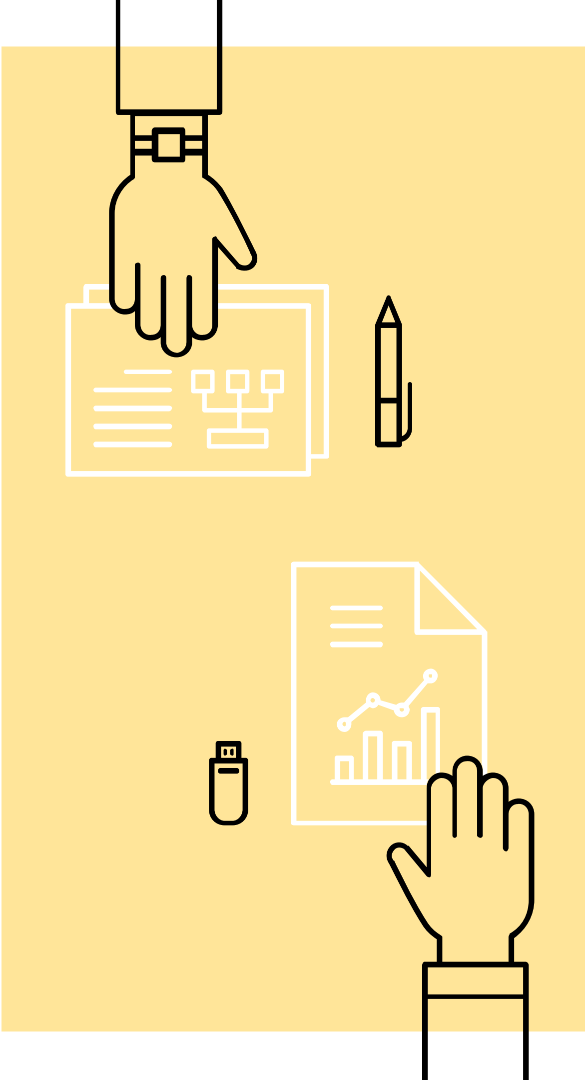


Files are we trying to read to recognize sound

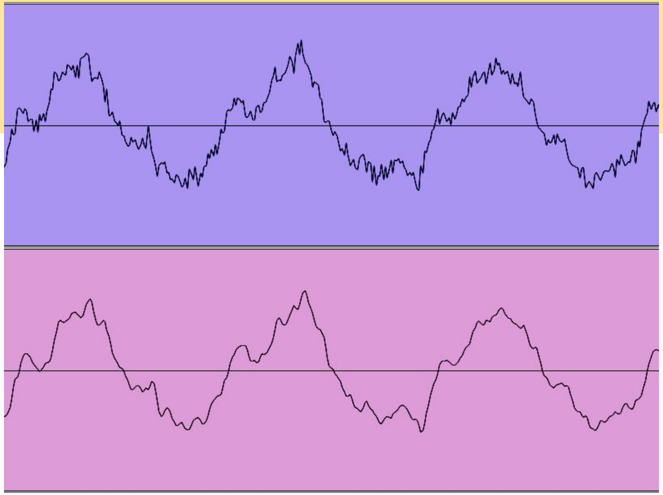
- WAV (Waveform Audio File) format
- MP3 (MPEG-1 Audio Layer 3) format
- WMA (Windows Media Audio) format

We choose to start with WAV since it is easier to separate the frequencies.

This will give us an easier time when we start to get a hands on coding.

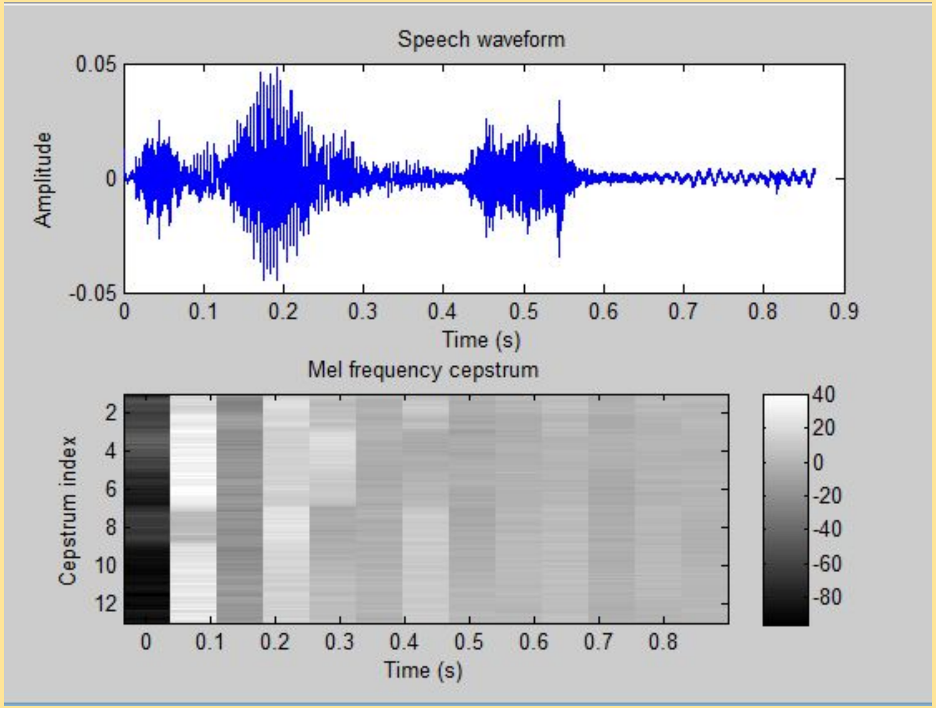


- WAV format can cover the full frequency that the human ear is able to hear.
- An MP3 file is compressed and has quality loss whereas a WAV file is lossless and uncompressed.
- An MP3 will never sound better than a WAV, no matter what kbps it's at as it is all still lossy



Features

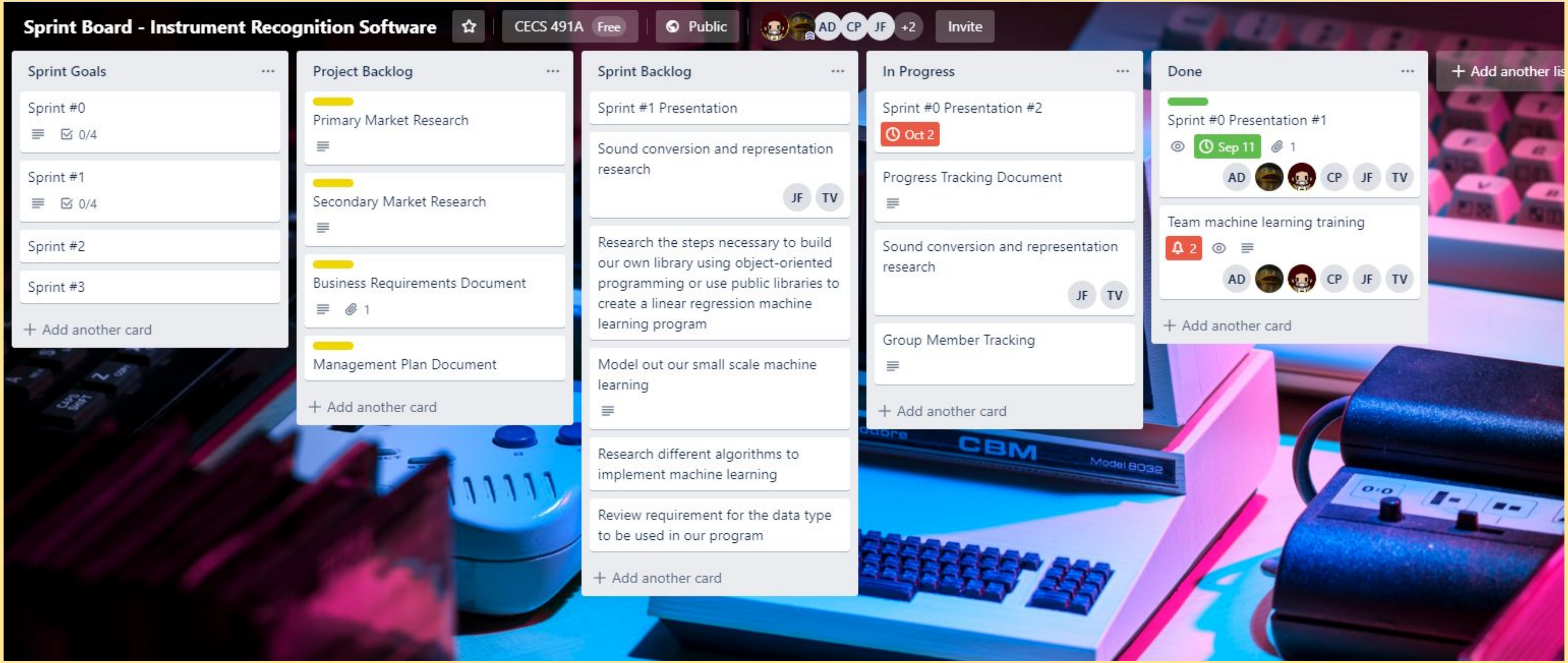
- Features are a summarization of the analysis
- In terms of music:
  - Mel-Frequency Cepstral Coefficient (MFCC) – measure of the timbre
  - Representation of keys, chords, harmonies, melodies, pitches, beats per minute, or rhythm



Sprint 0

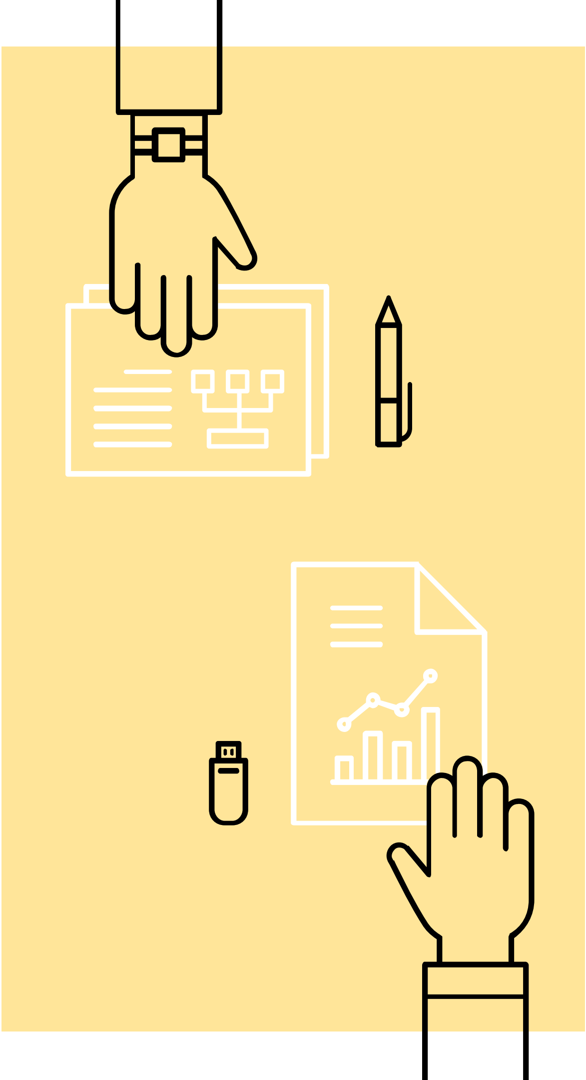
Date	Name	Category	Link	Summary	Conclusions
9/22/19	Aner	Dataset	https://www.kaggle.com/c/tenso	3.5 gb of training and testing dataset of people saying short words (yes, no, up, down, left, right, on, off, stop, go)	Useful for training albeit not for music
9/25/19	Joe	Linear Regression	https://www.dezyre.com/project	Implementation of a linear regression model on pre-existing data using a library	We should avoid libraries to better understand the actual methods used in ML
9/25/19	Joe	Library Creation	https://www.youtube.com/watch	In progress	
9/25/19	Joe	Resources	https://www.reddit.com/r/Pythor	In progress	
9/25/19	Aleks	Linear Regression	https://towardsdatascience.con	In Progress	
9/25/19	Cuong	Linear Regression	https://towardsdatascience.con	Talks about what the linear regression is and goes over the equation.	We can use the equation in our program if necessary. (y = a_0 + a_1 * x)
9/25/19	Cuong	Linear Regression	https://www.youtube.com/watch	Tutorial on how to work on the machine learning linear regression.	We can use this to help us code the algortihm.
9/25/19	Thomas	Bird sound recognition	https://www.youtube.com/watch	Identification of bird species based on sound made through machine learning and signal processing	
9/25/19	Thomas	Multivariate Regression	https://www.youtube.com/watch	Predicts the home pricing from analyzing data on previous home prices based on multiple variables	
9/25/19	Aner	Dataset	http://www.philharmonia.co.uk/	250 mb of data from different types of instruments	
9/30/19	Axel	Audio voice Processing	Audio Voice Processing Deap L	In progress	

Sprint 0


Sprint Board - Instrument Recognition Software
CECS 491A Free
Public
AD CP JF +2
Invite
Sprint Goals
Sprint #0
0/4
Sprint #1
0/4
Sprint #2
Sprint #3
Add another card
Project Backlog
Primary Market Research
Secondary Market Research
Business Requirements Document
1
Management Plan Document
Add another card
Sprint Backlog
Sprint #1 Presentation
Sound conversion and representation research
JF TV
Research the steps necessary to build our own library using object-oriented programming or use public libraries to create a linear regression machine learning program
Model out our small scale machine learning
Research different algorithms to implement machine learning
Review requirement for the data type to be used in our program
Add another card
In Progress
Sprint #0 Presentation #2
Oct 2
Progress Tracking Document
Sound conversion and representation research
JF TV
Group Member Tracking
Add another card
Done
Sprint #0 Presentation #1
Sep 11
1
AD CP JF TV
Team machine learning training
2
AD CP JF TV
Add another card
Add another lis

Code?

- Currently, we have not yet generated any code
- Focus on learning how machine learning works through testing pre-written codes
- Plans on generating a small scale linear regression code for Sprint #1



Machine Learning Libraries

Tensorflow – suggested ML library

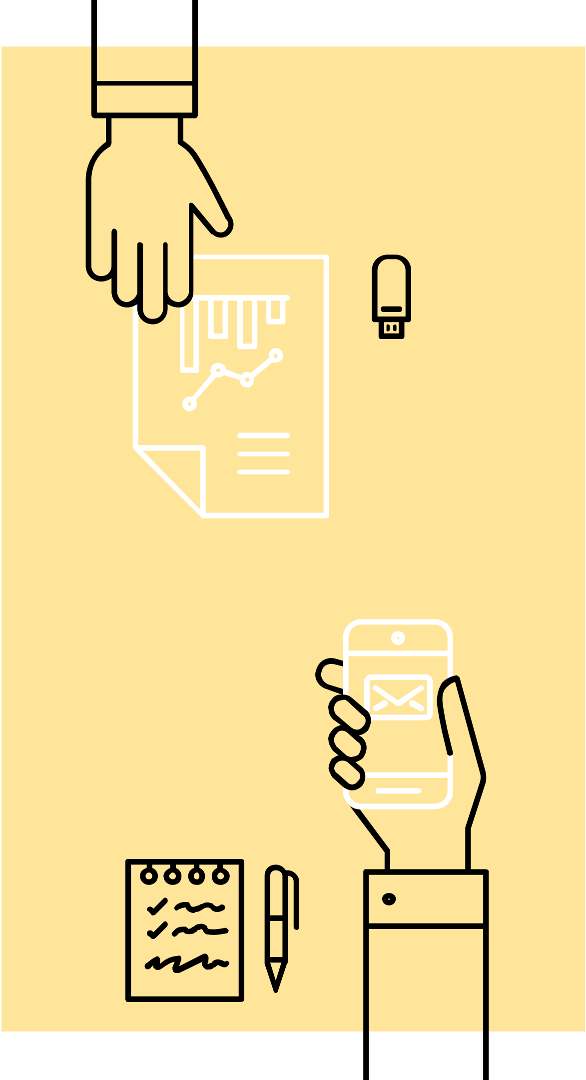
Keras – deep learning library for tensorflow

Librosa – feature extraction and data processing

Pandas – data structures & analysis library

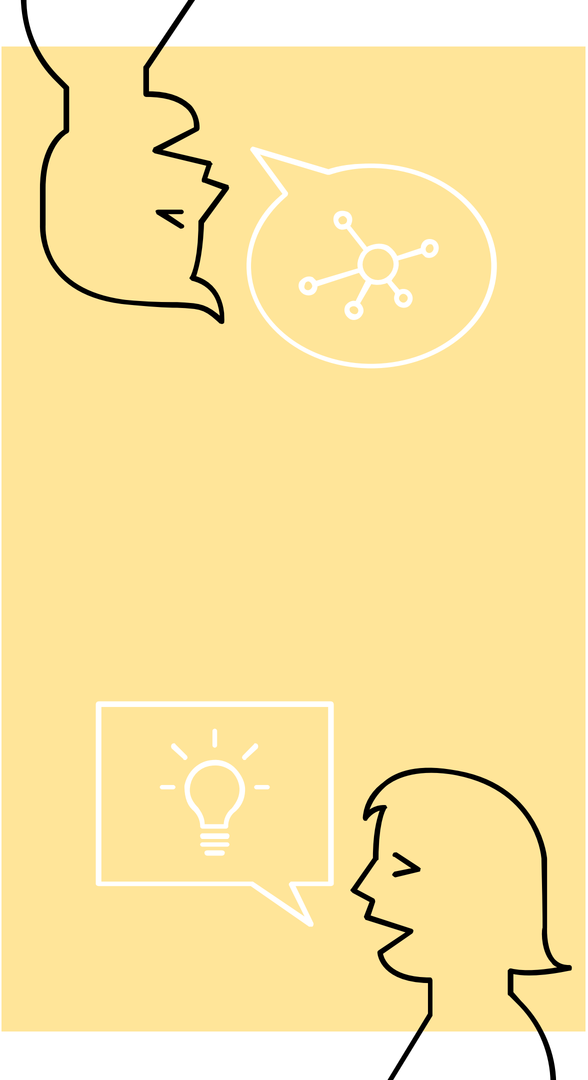
PyAudio – audio stream library

Scikit – modules for ML and data mining



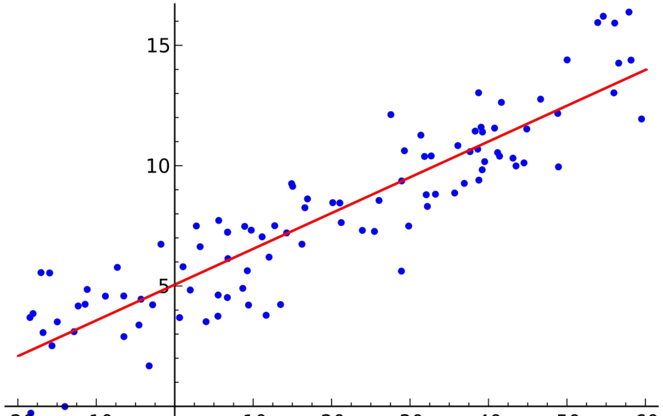
Tasks

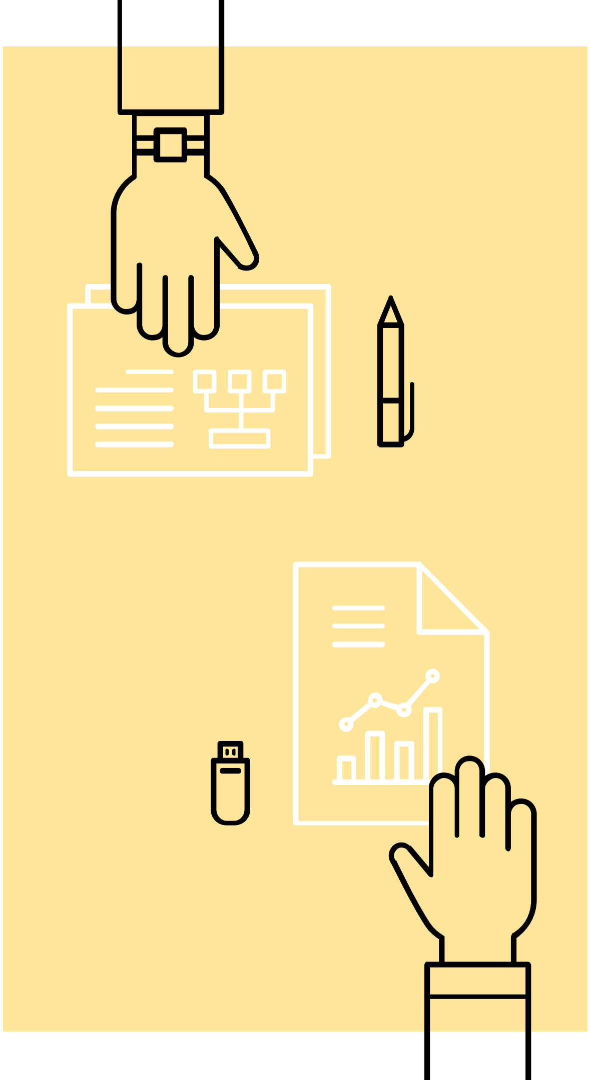
- Read research articles on machine learning
- Researched machine learning algorithms
- Found a basic machine learning project to learn

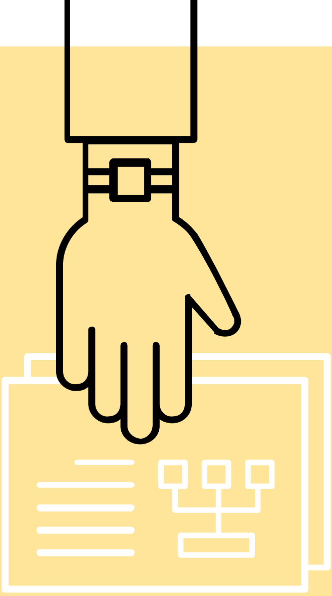


The future

- Plans for a small scale Linear Regression ML Project
- Continue research on sound engineering







Thank you for listening!