Sprint #1
Instrument Recognition Software

Members: Aleks, Aner, Axel, Cuong, Joe, and Thomas

BRD Update

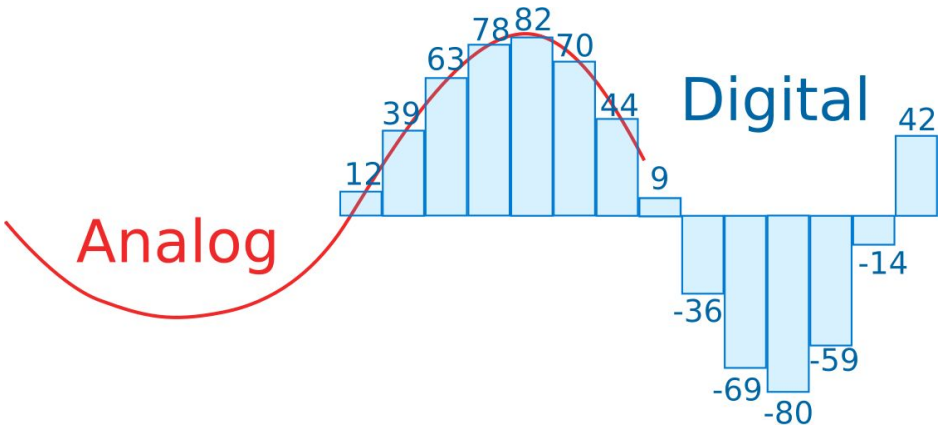
Currently no updates to the Business Requirements Document, this sprint is more implementation focused so we focused more on the research for sound recognition.

Management Plan Update

Currently no updates to the Management Plan but we our in the progress of updating the **Gantt Chart**, **Project Tracker Matrix**, **Task Management Velocity**, **Sprint Board**, and **Burndown Chart** inside of the Management Plan.

Digital signals

- In order to extract meaningful data for training our model, we must first transform the analog audio coming from the instruments into **digital signals** that our computers can easily understand.

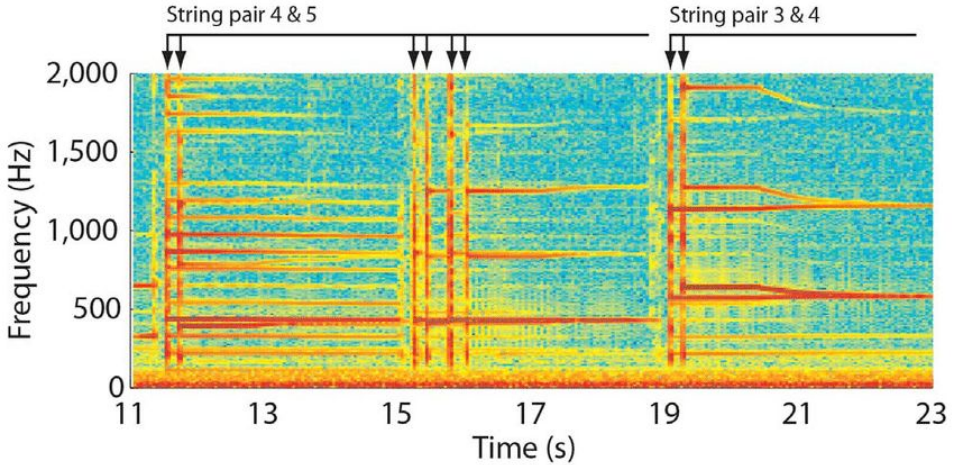


Digital signals

- In our case, we are planning on using a sizable data package provided by Philharmonia Orchestra
  - 250 megabytes worth of sound data in MP3 format
    - Since MP3 is lossy and compressed, it requires conversion to WAV which is lossless and uncompressed, resulting in a better quality
  - A total of 20 instruments from the Brass, Woodwind, Strings, and Percussion family
    - We are planning on comparing at least two instruments for now, as training for accuracy can take a long time.
    - We are planning on adding the other instruments over time.

Digital signals

- Digital signals can be represented in a Spectrogram, where **Frequency** is in the Y-axis and **Time** in the X-axis



Spectrogram of a guitar being tuned in real time by CI1 audio interface

Instrument differentiation

Timbre is the unique and distinctive sound that belongs to a specific instrument and helps the listener to distinguish it from any other instrument or source of sound.

- For instance, timbre is what makes a guitar sound like a guitar and not like a flute, a piano, or a violin.

Timbre is also what makes broad categories of musical instruments – such as “string instruments” or “wind instruments” – sound similar, but it is also what separates each instrument within those broad categories.

- For instance, although guitars, violins, ukuleles, and harps are all members of the string family, each of these individual instruments still has its own exclusive and defining sound.

Other potential features

While Timbre refers to the quality of sounds among different instruments, **Tone** can be used to refer to the quality and frequency of a sound as compared to itself.

- This sounds complicated, but what it really boils down to is the difference in frequencies of Bass (low sounds) and Trebble (high sounds).
- For example, imagine that you are listening to someone playing an electric guitar. The guitar's tone control is turned all the way to the lowest frequency setting so that it produces a very deep and low sound. Suddenly, the guitarist moves the tone control all the way to the highest setting so that the sound is very piercing and sharp. The timbre of this instrument is unchanged and they may repeat the same notes as before, but the tone has been altered dramatically.

Other potential features

Texture is sometimes used loosely to mean “the timbre of multiple instruments playing together”

The texture of music can fall into a number of distinct categories, including:

- Monophonic: This is an unaccompanied melody. Think of humming to yourself or a saxophone playing a wailing melody without any accompaniment.
- Homophonic: This means a melody which has an accompaniment. Most pop and rock songs are homophonic: the singer performs with a band supporting them.
- Polyphonic: This is where there is more than one melody, each basically independent of the others.

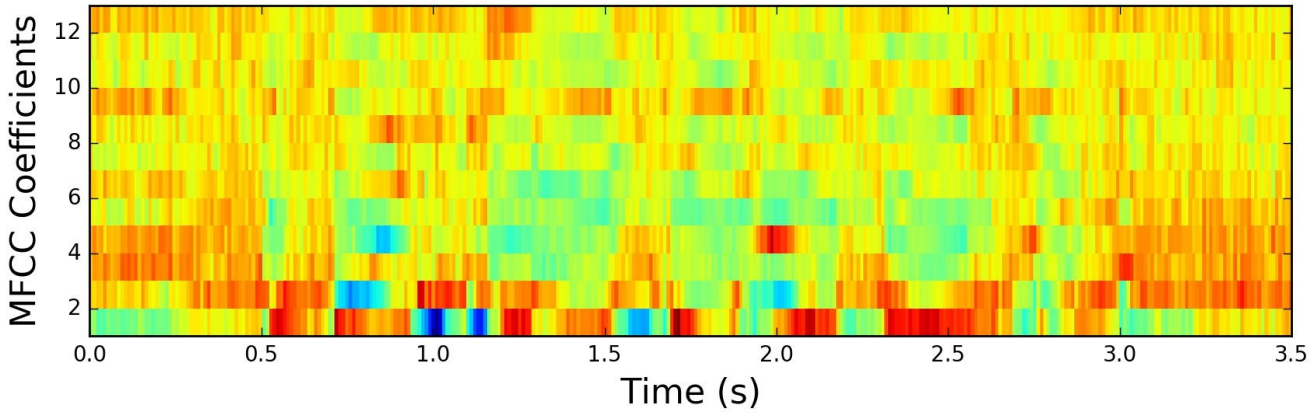
Other potential features

Other features that can be used:

- Pitch and Melody
- Harmony and Chords
- Rhythm
- Expression
- Form

Mel-Frequency Cepstral Coefficients

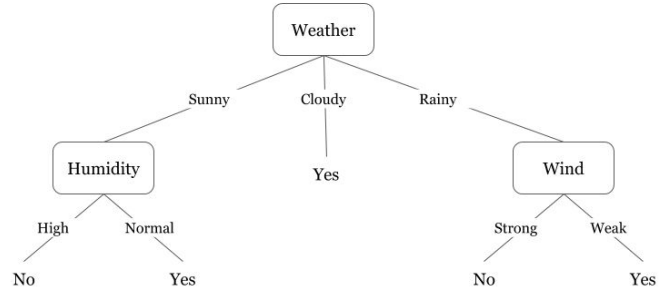
- MFCCs are a way of representing the spectral information in a sound.
- A number of studies used MFCCs to study the Timbre space.
- MFCC is expressed on a **time-domain scale** as shown below.

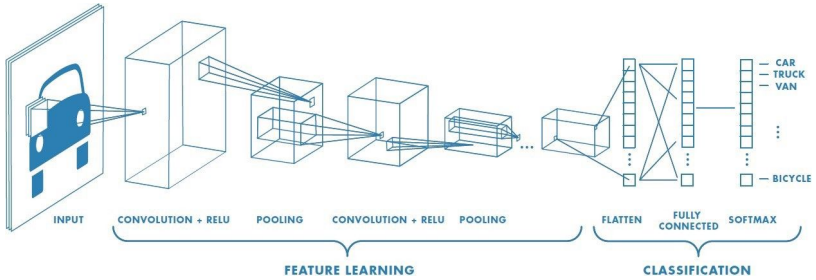


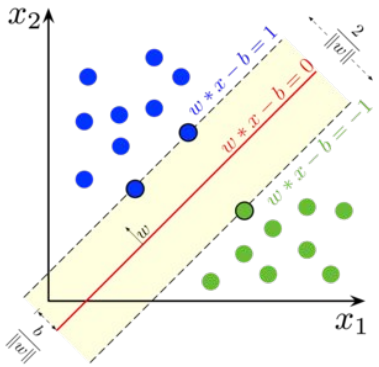
Candidate Models

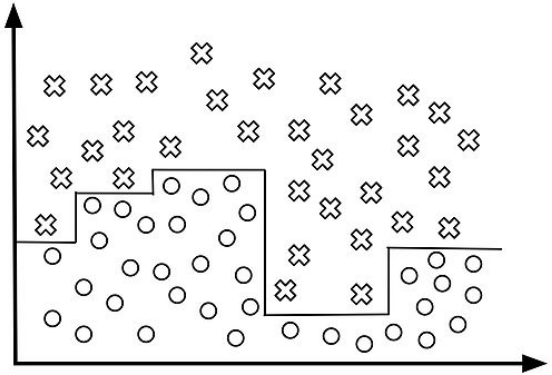
Label focused machine learning:

- Decision Tree
- Random Forest
- SVM
- Convolutional Neural Network









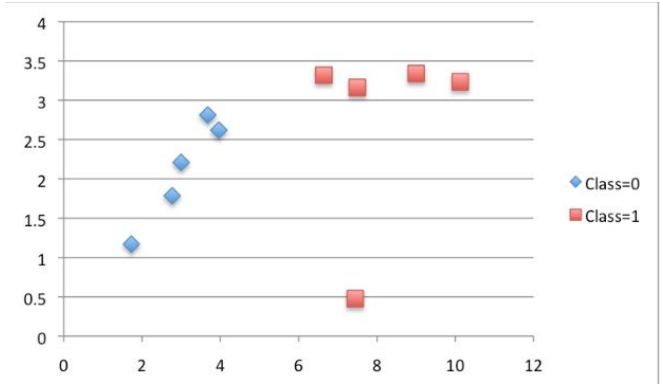
Code

Model Implementation: Decision Tree

- Label-based model for instrument recognition
- 1300~ data points tested
- Accuracy will be far from perfect in final product
- Transition into SVM / CNN

	X1	X2	Y
1	X1	X2	Y
2	2.771244718	1.784783929	0
3	1.728571309	1.169761413	0
4	3.678319846	2.81281357	0
5	3.961043357	2.61995032	0
6	2.999208922	2.209014212	0
7	7.497545867	3.162953546	1
8	9.00220326	3.339047188	1
9	7.444542326	0.476683375	1
10	10.12493903	3.234550982	1
11	6.642287351	3.319983761	1

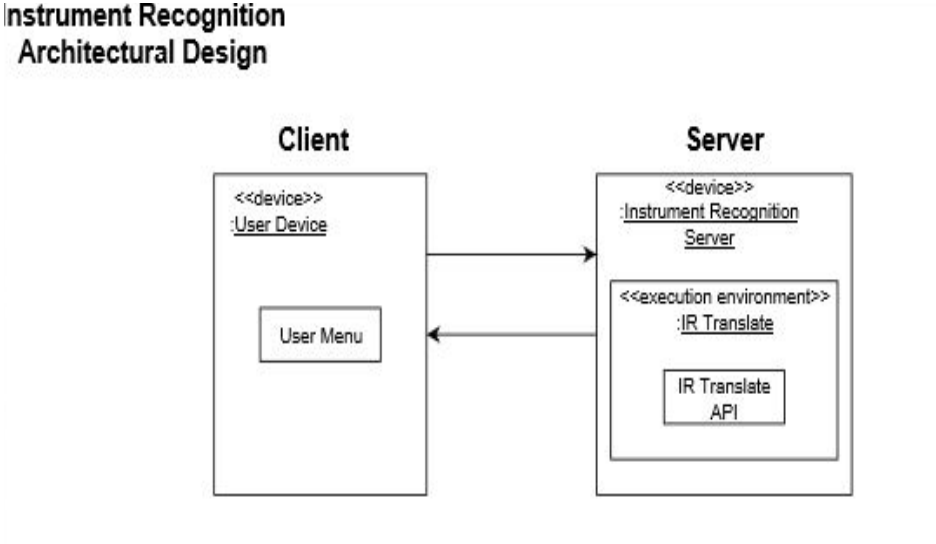
```
Expected=0, Got=0
Expected=0, Got=0
Expected=0, Got=0
Expected=0, Got=0
Expected=0, Got=0
Expected=1, Got=1
Expected=1, Got=1
Expected=1, Got=1
Expected=1, Got=1
Expected=1, Got=1
```




```
Scores: [96.35036496350365, 97.08029197080292, 97.44525547445255, 98.17518248175182
Mean Accuracy: 97.299%
```

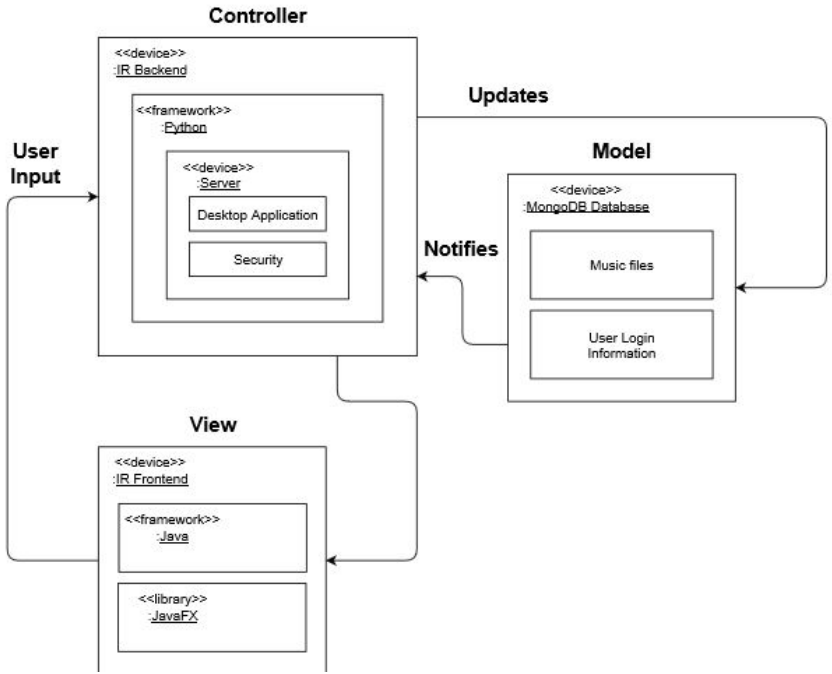
Architecture and Design Document & Trade Off Analysis

- **Client**
  - User's side of the software that contains a menu which allows them to interact with the software
- **Server**
  - Handles user login authentication, machine learning model, and uses algorithms to calculate the output.



Architecture and Design Document & Trade Off Analysis

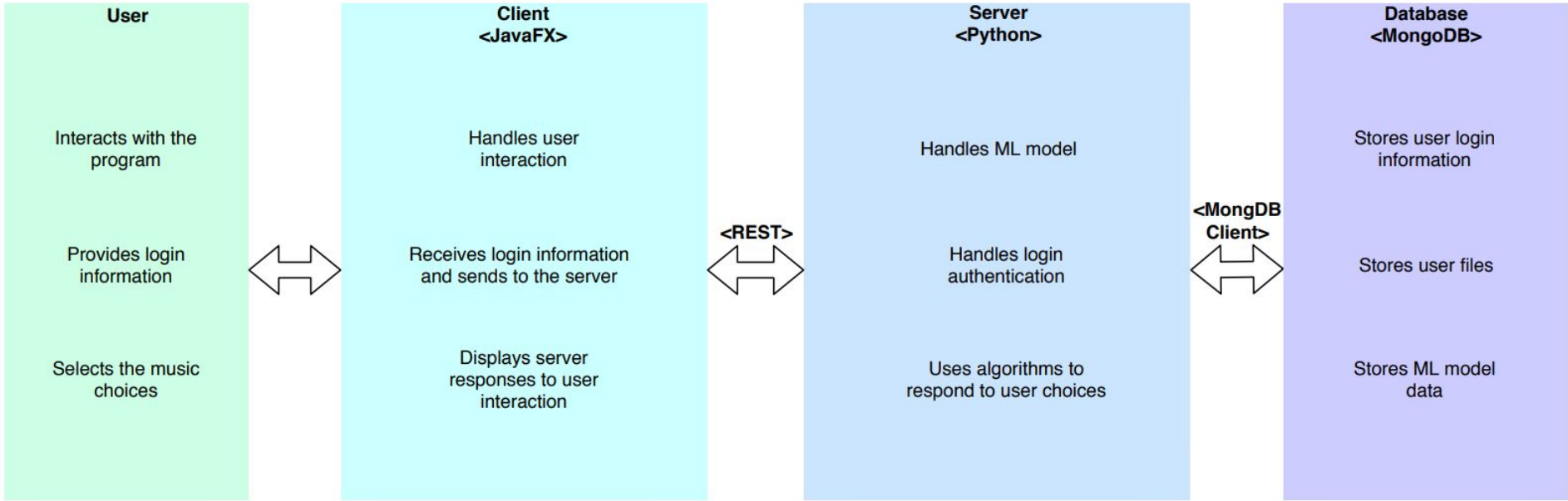
- Controller
  - Manage updates, serve as main process for Mode and View.
- View
  - Show the user the final outcome
- Model
  - Contain login information and music files.



Architecture and Design Document & Trade Off Analysis

Criterion	Weight	Python	Java
Object Oriented	25%	0	1
Data Management	50%	1	0
Ease of Use	10%	1	0
Libraries	15%	1	0

Architecture at high level

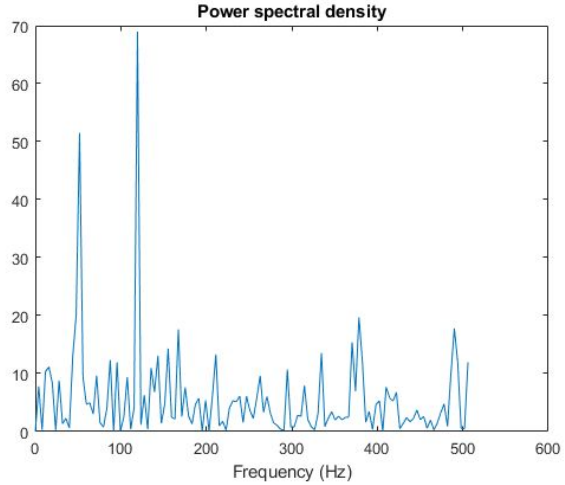


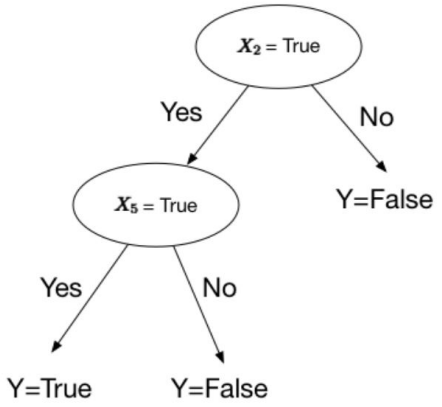
Alternatives

- Java
- Chose python because of the built in libraries and the ease of data manipulation

Sprint Goals

- Research the steps necessary to develop library
- Model small scale learning using a library
- Research different kinds of machine learning models to find what we need
- Get requirements for the kind of data we use





User Stories and Sprint Priorities

Factors for determining priority:

- Displayable progress
- Proof of concept
- Class deadlines

Story Progress:

- Front end log in display
- Working model -> Applicable model
- Draft feature list with potential options

Story Points:

- 68 projected, 63 completed

Retrospective:

- Less ambition for accurate projection
- Allocate more time for documentation

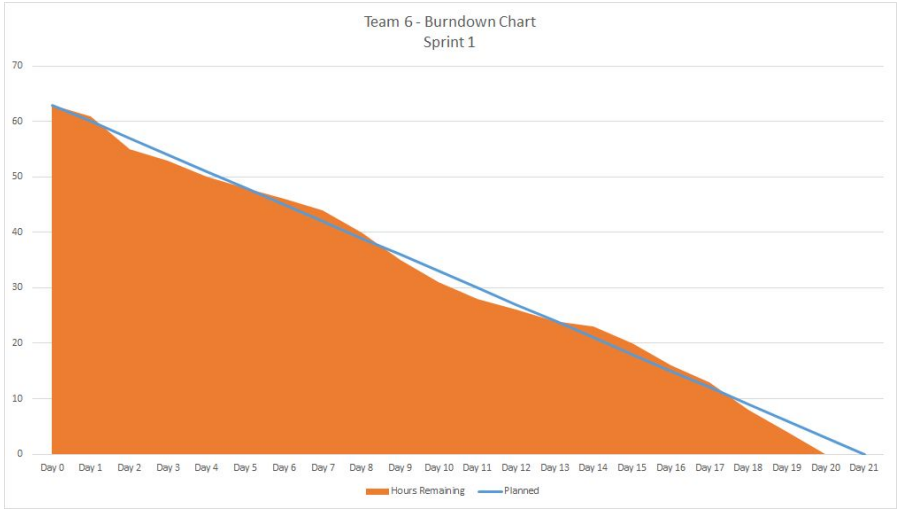
Tasks

Because our project is currently research-focused, we don’t have many user personas or tasks assigned.

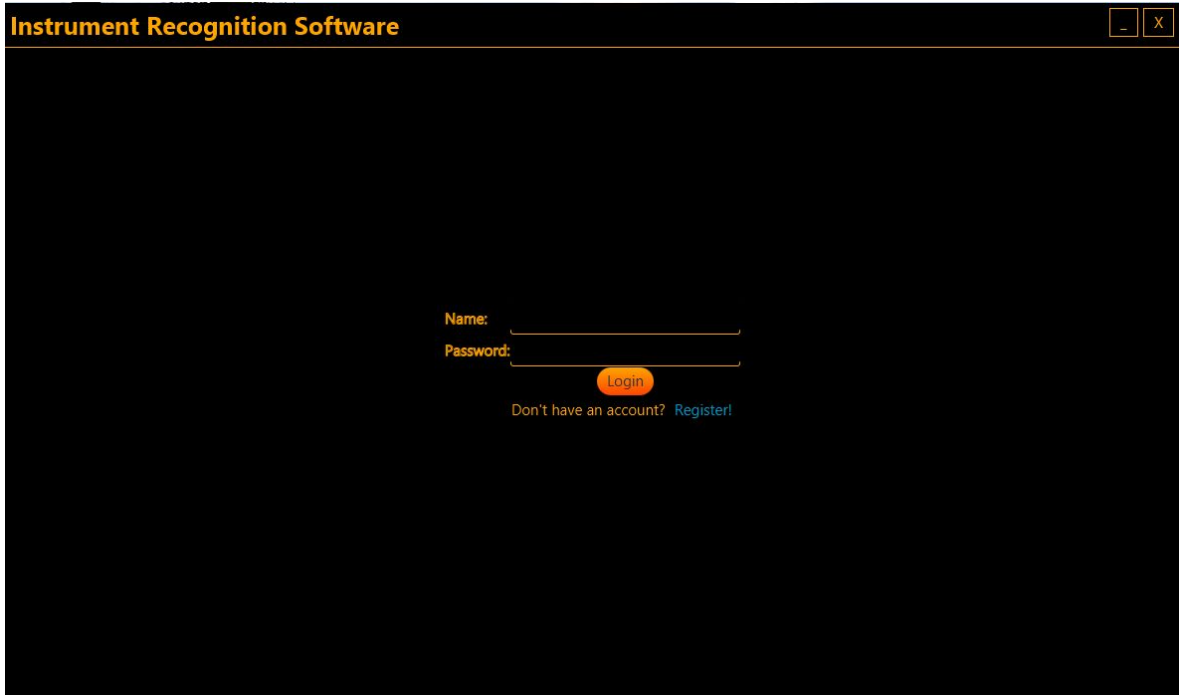
- As a user, I want to be able to log in and use this service
- As a user, I want to search for songs by the instrument type I want to hear
- As a user, I want to know what instruments are played in a song that I listened to

Burndown Chart

- Our work progress towards this sprint's goals went relatively smooth, excluding the hiccups that are the result of our other classes taking up our time
- We hope to better optimize our time and efficiency in our work velocity



Frontend


Instrument Recognition Software
_
X
Name:
Password:
Login
Don't have an account? Register!

Code Review

- Basic model of a decision tree
  - No project-specific data used
  - Final model will be in a similar category but different model
- Front end log in is ready for evaluation
- Most features will not be presentable because they are all dependant on a finalized machine learning model