

Business Requirements Document

1. Preface

a. State why we are interested in a Sound Recognition app:

i. To identify unique actors in audio in order to accurately label sources.

These labels can be then employed to sort, recommend, or identify clips of interest.

b. What technological innovation and depth we plan to bring to solving the problem:

i. How is our product different from those on the market?

There are almost no such other applications as ours that can isolate specific sounds from an instrument or person (quick Google search reveals this).

2. Appendix

a. User Persona (Company)

Sony

- A Japanese company headquartered in Tokyo, founded in 1946
- Owners of the largest music entertainment business in the world
- One of the leading companies in the consumer and professional markets
- Markets include music entertainment business, video game console business, video game publishing business

Instrument Recognition Software

3. Mini Business Plan

a. Strategy

Lean Canvas				
Problem	Solution	Value proposition	Unfair advantage	Customer segments
<ul style="list-style-type: none">• In today's modern world companies and individuals handle such large quantities of data that it is incredibly impractical to classify it manually. Without clear and accurate classification of mass data it is near impossible to locate desired files or sort them by a criteria. The scope of this project has been limited to musical data. How can we identify and pinpoint desired data in a set that has not been labeled manually?	<ul style="list-style-type: none">• We will develop one Instrument Recognition Software that can accept musical data, produce labels corresponding to the instrument(s) that produced it, and search by these labels. To achieve this goal, we plan to implement search engine, sound recognition, audio conversion, and machine learning concepts.	<ul style="list-style-type: none">• Our solution identifies unique actors in audio and accurately labels sources to sort, recommend, or identify clips of interest.	<ul style="list-style-type: none">• There currently are not any technologies that fill our intended niche	<ul style="list-style-type: none">• Primary consumers would be large music corporations such as Apple or Spotify whom our problem statement would apply to. These companies may also desire to provide their consumers with our additional search criteria to gain a personal advantage over their competitors. Additionally, companies like YouTube(Google) may be interested in a way to get meaningful information out of large quantities of audio without manual analysis.
	Key metrics		Channels	
	<ul style="list-style-type: none">• Our success can be measured by: the number of unique labels, the accuracy of audio labeling, and the accuracy of a search on a labeled database		<ul style="list-style-type: none">• We will release a computer application accessible to the public while also reaching out to the listed companies with intentions of being adopted into their current services.	
Cost structure		Revenue streams		
<ul style="list-style-type: none">• No fixed variable cost predicted.		<ul style="list-style-type: none">• Use the feature as an additional paid option within a product. Pay per periodical access.		

Instrument Recognition Software

b. SWOT

SWOT	
Strengths	Weaknesses
<p>Your advantages:</p> <ul style="list-style-type: none">• It has sound recognition• It can process and analyze various audio file types• It can analyze the frequency spectrum• It can isolate the sounds in an audio clip and determine each individual instrument being played• It allows users to search based on instruments and vocal sounds• Is easy to use by consumers because it requires no preprocessing of audio data• Automates mundane/wasteful task• Employs a multitude of machine learning models• Fulfills an otherwise unmet need throughout the industry increasing market value	<p>Areas for improvement:</p> <ul style="list-style-type: none">• Not a necessity for general music consumers• Difficult to differentiate different sounds - the application can become confused and misinterpret background noise as an instrument playing• Difficult to find resources to find inspiration from since the idea of sound recognition is still mostly in research and development• Not all audio clips follow the same rhythm or pacing - it may also confuse the frequencies within a faster-paced audio clip of one instrument with another instrument• Team is not specialized in auditory science - start from scratch in learning auditory science• Will require immense processing power for large file batches• Has a significantly large run time for initial classification• May not be able to differentiate between instruments with extreme similarities• Requires giant sets of training data• Cannot guarantee a product that fulfills desired operations until the final phases of development• Cannot produce a sufficiently representative model of the developing product for user evaluation
Opportunities	Threats
<p>Situations to apply your advantages:</p> <ul style="list-style-type: none">• AI training for a wider range of sound• Allowing other compression formats such as MP3 or FLAC• AI training for differentiating overlapping frequencies• Expanding to sounds from musical instruments• Being able to expand towards many industries• Various Python libraries allow for ease of design and implementation• Can specialize in specific ethnic music• Can be adapted for all kinds of instruments• Has potential to be abstracted to general audio classification such as animal identification through audio	<p>Where you are at risk:</p> <ul style="list-style-type: none">• Training AI to recognize specific sounds will require significant amount of computing power• Having a wide range of sounds mean that it would require a lot of time to be able to recognize all of the sounds in the universe we live in• Since this technology is new and is still in research and development it is hard to find programs already in the market to seek inspiration from• Having such an ambitious idea will require enormous levels of effort that may not be fulfilled• Lack of prior knowledge of the field may leave potential issues unaccounted for• It may be necessary for the development time of this software to exceed the course duration causing us to miss deadlines• A competitor may develop a product that fills our niche over our development period eliminating the competitive advantage• Due to the solution not yet existing, the solution may not be possible with our given level of technology• Our processing time may be so large that it renders our product impractical• Time spent toward the project may be limited by the availability of the students involved

i. Swot Analysis

1. Are there better solutions already?

There are similar solutions but none of the market solutions overlap with our idea.

2. Could there be better solutions?

Our main goal is to improve previous solutions to the same problem in a way that has never been successfully done before.

3. How is (or is not) your solution a unique solution to the problem?

Our solution is a unique solution because there are no other applications that process sound in such an amount of detail.

c. Market Segments

Segment Profile			
Geographic	Demographic	Behavioral	Psychographic
This market segment is world wide.	This market segment is comparable to Sony and Spotify. It will consists of industries interested in sound recognition.	Users of these services would be more satisfied or inclined to use one provider over another as it provides companies the technology no one else has.	This market segment targets companies that are in the competition for new technologies to gain customers.
Size		Growth potential	
The market segment size contains about 217 million users based on Spotify user count but we are unsure about Sony's so the numbers might be higher.		The segment can grow exponentially once the technology is implemented.	
Competitive activity		Risk	
Companies such as Sony and Spotify are competing to get this technology.		Success in this market segment will be challenging since we are a small group with little resources, however, the competition is low since the technology is still in research and development.	
Approach			
To succesfully serve this segment we will provide a strong solution to the basic problem, starting small to create a universal model for growth.			

d. Primary Market Research

i. Interview potential company representatives

Potential interview questions:

What is your name and role in the company?

What is the main focus of your company in the market?

Can you provide an estimate on the number of users?

How engaged do users ten to be with new features after release?

How much of the user engagement is in the music industry?

e. Secondary Market Search

i. Addressable market size:

As a tool sold to an enterprise, our application would appeal to the same market demographics as the company we sell our innovation to. For example, if adopted by Spotify our product would impact their 217 million monthly users.

Statistic taken from this article:

<https://venturebeat.com/2019/04/29/spotify-passes-100-million-premium-users-as-revenue-climbs-33-yoy/>

ii. Validation of user need:

Many companies and research groups have published various papers and reports on deciphering unique actors in audio through isolation, known instruments, and their classification. This continual investment in research by well-renowned institutions shows a strong financial and intellectual interest in the subject. We expect these parties, and by extension their resources, to be interested in our product as it contributes to and expands on the topics they have invested into.

iii. Relevant research in the field:

The ability to analyze audio begins with the transformation of pure analog sound to a digital representation that can be dissected by a computer.

Techniques to do so are discussed here:

<https://arxiv.org/pdf/1804.03160.pdf>.

Once a data from audio is created, it can then be evaluated by programs looking for similar patterns in sound. When inspecting frequency, these patterns can be used to identify unique instruments as shown in this paper

<https://arxiv.org/pdf/1705.04971.pdf>.

After instruments can be identified individually they must then be recognizable in an ensemble. The process for identifying unique actors in an audio file has been studied here

<https://docs.google.com/document/d/1leVnL66ahwdShc2iij1WGtcNPE69WeGAcKykj4RV4uY/edit>.

Our project will require extensive background research on sound conversion and machine learning which we will base on the research papers above. We plan to discuss our potential product with a Sony representative to further narrow the scope of topics we must learn to deliver a desirable product. Over the course of our primary research interview we aim to identify and isolate desired features for our product that will guide our research and development.

f. Commercialization Strategy

- i. Patent the product / technology
- ii. Provide for sale through a well-known platform / online service
- iii. Partner with researchers

g. Monetization Strategy

- i. Revenue from initial fees to purchase the product
- ii. Revenue from monthly fee which charges the consumers once a month
- iii. Revenue from advertisements

4. Total Solution

a. Describe our solution

Help users identify instruments in sound snippets through comparing the sound spectrums using machine learning.

- b. Explain clearly how the solution meets the user, business or market needs

Users can identify instruments used in music segment and find other songs that use said instruments. The market does not have a model that would precisely identify an instrument based on the sound spectrum comparison.

- c. Provide solution alternatives and show trade off analysis

An alternative solution would be to use visual recognition instead of sound recognition that determines the instruments used in a song based on what player played, problem with this is that it would require an mp4 file instead of an mp3 file since you would need visual images to determine the instruments used in the songs.

- d. How did we decide to choose the particular approach we did?

We decided to choose the particular approach because we thought that it would be the most efficient way to determine what instruments will be used in the songs. Since the sound spectrum is measurable, this allows us to determine sound frequencies and know which instruments it came from based on the frequency.

5. Optional

- a. We may elect to solve a complex problem needing an extensive solution, all of which cannot be solved within CECS 491

Our solution is so extensive that we will have to scale down for the purposes of finishing within the time frame of CECS 491.

- b. We can select a part of the solution identified from market research to be implemented in CECS 491

We will create a solution that will cover only one or two parts of the complex problem.

- c. Which parts of the total problem is our solution responding to? Which parts remain unsolved? What is our rationale for our decision?

Our solution is responding to detecting a single instrument within a song.

The unsolved parts of the problem include implementing that solution for more than one instrument, processing multiple songs at a time, and providing future song recommendations.

6. Analysis:

- (a) What is the state of the art research currently.

There is no research on the specific solution we are trying to implement but there are similar solutions.

- (b) What is a reasonable deliverable for your team.

A female/male voice recognition program by the end of the semester.