## ****Troy Sabia****

****C964****

****Task A****

## ****Project Proposal: Automated Metal Genre Classification for Streaming Platform****

### 1. Summary of the Problem

In the current digital music ecosystem, metal fans often struggle to find platforms that deliver a consistent, genre-specific listening experience. General-purpose streaming services use either human curation or genre tags that are prone to error and bias. As your platform plans to serve a metal-focused audience, the need for an automated, scalable solution to accurately filter out non-metal content becomes critical.

### 2. Customer Benefit and Decision Support

The proposed system will automatically classify incoming audio files as either "metal" or "non-metal" using a trained machine learning model. This allows your platform to maintain brand identity by ensuring only metal tracks are streamed, eliminate the cost and variability of manual tagging or curation, and scale your catalog efficiently without compromising genre integrity.

This tool directly supports content acquisition decisions and platform quality assurance.

### 3. Outline of the Data Product

The data product is a lightweight, efficient classification tool consisting of:

* A preprocessing pipeline that converts music files into 15-second mono audio clips and extracts Mel Frequency Cepstral Coefficients (MFCCs).
* A convolutional neural network (CNN) trained to distinguish metal from non-metal tracks based on those MFCCs.
* A set of evaluation tools to validate performance and ensure future expandability, including subgenre classification.

### 4. Description of the Data

To train the model, we use publicly available datasets such as the GTZAN genre collection as well as additional MP3s collected from the web. Only audio features (MFCCs) are extracted and used for training; no raw music is redistributed or republished. All audio inputs are transformed into non-reversible, non-identifiable feature representations, ensuring compliance with fair use under educational and research contexts.

### 5. Objectives and Hypotheses

**Primary Objective:** Build a proof-of-concept tool that can classify a music clip as "metal" or "non-metal" with at least 90% accuracy.

**Hypothesis:** Given a properly trained neural network and a well-balanced dataset, it is possible to achieve high classification accuracy using only short audio segments and extracted features.

### 6. Project Methodology

The development process will follow these stages:

1. **Data Collection & Preprocessing**: Standardize audio to 22.05 kHz mono format, extract 15-second clips, and convert to MFCCs.
2. **Model Design**: Implement a CNN architecture tailored to time-frequency data.
3. **Training**: Use supervised learning with class-balanced training data and learning rate scheduling.
4. **Evaluation**: Measure accuracy and per-class precision/recall using held-out validation and test sets.
5. **Packaging**: Save the trained model and create a script for batch classification of new music files.

### 7. Funding Requirements

The total estimated cost for development and training is **$1,500**, which covers:

* 10 hours of development and testing time.
* GPU server rental for model training and fine-tuning.  
  No additional hardware or proprietary software is required.

### 8. Impact on Stakeholders

**For the startup**: Enables automated content vetting and consistent user experience.  
**For users**: Guarantees genre purity and avoids alienating metal fans with off-brand content.  
**For curators**: Reduces manual labor and improves scalability.

### 9. Ethical and Legal Considerations

#### All audio used is either from publicly available sources or widely accepted academic datasets.

* Only non-identifiable, non-reversible feature data is used for training.
* No music is distributed or republished.
* Commercial deployment will require legal review of music rights and licenses to ensure full compliance.

### 10. Relevant Expertise

I am a computer science undergraduate with formal coursework in machine learning and experience building audio plugins and DSP tools using the JUCE framework and Reaper’s JSFX. My background includes hands-on development in both creative and technical audio environments, making me uniquely qualified to bridge music and AI effectively.