

1) What is the problem that you are attempting to solve?

**I would like to categorize music/songs by genre (classical, pop, country, etc.) using the features of a sample sound file (.wav)**

2) How is your solution valuable?

**It could be used to create groups of songs similar to Pandora music stations that group songs by genre and other characteristics, used by music lovers to create personalized playlists.**

3) What is your data source, and how will you access it?

**The source for the training data would be:**

**<https://www.kaggle.com/andradaolteanu/gtzan-dataset-music-genre-classification>**

**This is a Kaggle dataset that contains 1000 sound files (.wav) that are converted into image files (.png) using spectral analysis, categorized by genre. It also has a data file that contains other characteristics of the sound files. I will find other sound files as a test set to attempt to determine what genre they belong to.**

4) What techniques from the program do you anticipate using?

**I will be using deep learning techniques such as neural networking, TensorFlow and Keras to analyze the image files created from the sound files. I will also use unsupervised learning techniques such as K-Means to categorize the data by genre. The librosa library will be used to convert the sound files to spectral image files**

5) What is the biggest challenge that you anticipate facing?

**I have found from reading several articles that categorizing music by genre is a difficult task. I will need to make my data modeling and analysis very precise to determine the differences in music features between each genre.**