Name: \_Oliver James Tan, Chieu Le Heng\_\_\_ ( 5 ) Date: \_\_\_\_\_4 Feb 2020\_\_\_

**Project Proposal**



|  |  |
| --- | --- |
| **Project Title:** | Musical analysis, tagging and classification by way of AI |
| **Team Members:** | Oliver James Tan, Chieu Le Heng |
| **Objective:** | Classifying music by ~~genre~~ composer using AI which analyses stylistic choices in the piece (e.g. instrumentation, lyrics, singing style, metre, etc.). This information can be used to sort songs by ~~genre~~ composer as well as measure their similarity to other songs. |
| **Target Users:** | Musicians and laymen just looking for music they want to listen to  People looking to more easily classify music as well as find similar songs. |
| **Project Plan and Timeline** | |  |  | | --- | --- | | T1W6 - T1W8 | Try out and replicate [2], [3], [4] (See references), at least to an extent which is functional. | | T1W9 - T2W2 | Innovate and try adding novel improvements  If possible, integrate parts of the above into a single functional application. | | T2W3 | Refine, and make a report/presentation | |
| **Resources and Tools** | Libraries PyTorch/TensorFlow  [Librosa](https://github.com/librosa/librosa) (Python music analysis library) Tools *In general, audio files are large and the following feature extraction methods may be used. After that normalise and classify using a standard classification (k-nearest neighbor, logistic regression &c.) or some image classification (e.g., CNNs) if spectrograms are involved*  Spectrogram (of frequency)  Spectral centroid (centroid of frequencies)  Spectral rolloff (curve of frequency?)  MFCC (Mel Frequency Cepstral Coefficient): set of 10-20 features  Chroma Frequency (frequencies sorted into the 12 bins, 1 for each semitone, across octaves) Datasets [GTZAN (classic 2002 dataset)](http://opihi.cs.uvic.ca/sound/genres.tar.gz)  [Million Song Dataset](http://millionsongdataset.com/)  Others in [2]  <https://www.kaggle.com/jembishop1/classical-music-piano-rolls/discussion/94171>? |
| **Target Outcome and Benefits** | 1. tagging music database 2. integration in music app/search engine, as a service similar to Shazam/SoundHound |
| **References** | [1] Music Genre Classification with Python <https://towardsdatascience.com/music-genre-classification-with-python-c714d032f0d8>  (From a github search)  [2] List of AI in Music Resources <https://github.com/ybayle/awesome-deep-learning-music>  [3] <https://en.wikipedia.org/wiki/Music_Genome_Project>  [4] Using CNNs on GTZAN dataset <https://github.com/Hguimaraes/gtzan.keras>  This list of references is incomplete and more will be found over the course of our research. |

## Own notes

Maybe map music to latent space with Magenta’s encoding, and use standard classification algos on that

Then characterise