

Readme

1. For q2, please find the following files:

- > q2_df_filtered.csv
- > q2_df_full.csv
- > q2_df_group.csv
- > q2_deloitte.ipynb
- > q2_deloitte.html

2. For Part 1 (q1), the work has been done in 3 files:

- > part1_single_wrangling.ipynb
- > part1_feat_eng.ipynb
- > part1_modeling.ipynb

The pipeline broadly consists of:

1. Features generation...
- 2, outlier removal,
3. random under-sampling,
4. model training,
5. model evaluation
6. model prediction using PS2 data and identification of piece/files that are NOT classified using the classifier

For feature generation, there were the basic steps

1. Convert midi to wav
2. Break wavs into 15s, 30s, 60s interval wavs. This to be done for the target PS1 files as well
3. Generate MCSSes and spectral metrics for each file
4. Store features and PS in separate css

TOOLS USED FOR AUDIO PROCESSING

=====

1. FluidSynth, <https://www.fluidsynth.org/>, last accessed: [09/16/2023]
2. Midi2Audio, <https://github.com/bzamecnik/midi2audio>, last accessed:[09/16/2023]
3. PyDub, <https://github.com/jiaaro/pydub>, last accessed: [09/16/2023]
4. Librosa, librosa/librosa: 0.10.1, last accessed: [09/16/2023]

RESULTS

=====

For 30 sec classifier:

num of files from PS2 classified as not belonging to the 4 composers: 35

filenames (not from 4 composers) :

0.23120017256495873_adj.mid

0.3334293069825759_adj.mid
0.1960551158929671_adj.mid
0.30993012400290976_adj.mid
0.9353908329958764_adj.mid
0.647959423719129_adj.mid
0.48367685897240176_adj.mid
0.3264565808984162_adj.mid
0.06402123326764841_adj.mid
0.21198476749665085_adj.mid
0.3620067189216978_adj.mid
0.09167358800381353_adj.mid
0.7491289879531658_adj.mid
0.3559970176888735_adj.mid
0.549470161204349_adj.mid
0.10250888365879718_adj.mid
0.36321860283443286_adj.mid
0.047574444458241216_adj.mid
0.21901852969811753_adj.mid
0.905499891236923_adj.mid
0.9511403301279795_adj.mid
0.07186746659481313_adj.mid
0.22047111832936942_adj.mid
0.539512676743813_adj.mid
0.033313986422223163_adj.mid
0.981087291054314_adj.mid
0.10222964826466285_adj.mid
0.337517805339117_adj.mid
0.1755252422917658_adj.mid
0.5807329043589801_adj.mid
0.26551079719260606_adj.mid

0.33695573887576447_adj.mid
0.002716920481628_adj.mid
0.10630249969742178_adj.mid
0.8014751784512073_adj.mid

for 60 second classifier

num of files from PS2 identified as not belonging to the 4
composers: 35

filenames (not from composers)

0.23120017256495873_adj.mid
0.3334293069825759_adj.mid
0.1960551158929671_adj.mid
0.30993012400290976_adj.mid
0.9353908329958764_adj.mid
0.647959423719129_adj.mid
0.48367685897240176_adj.mid
0.3264565808984162_adj.mid
0.06402123326764841_adj.mid
0.21198476749665085_adj.mid
0.3620067189216978_adj.mid
0.09167358800381353_adj.mid
0.7491289879531658_adj.mid
0.3559970176888735_adj.mid
0.549470161204349_adj.mid
0.10250888365879718_adj.mid
0.36321860283443286_adj.mid
0.047574444458241216_adj.mid
0.21901852969811753_adj.mid
0.905499891236923_adj.mid

0.9511403301279795_adj.mid
0.07186746659481313_adj.mid
0.22047111832936942_adj.mid
0.539512676743813_adj.mid
0.033313986422223163_adj.mid
0.981087291054314_adj.mid
0.10222964826466285_adj.mid
0.337517805339117_adj.mid
0.1755252422917658_adj.mid
0.5807329043589801_adj.mid
0.26551079719260606_adj.mid
0.33695573887576447_adj.mid
0.002716920481628_adj.mid
0.10630249969742178_adj.mid
0.8014751784512073_adj.mid

for 15 second classifier

num of files from PS2 identified as not belonging to the 4
composers: 35

Filenames (not from composers)

0.23120017256495873_adj.mid
0.3334293069825759_adj.mid
0.1960551158929671_adj.mid
0.30993012400290976_adj.mid
0.9353908329958764_adj.mid
0.647959423719129_adj.mid
0.48367685897240176_adj.mid
0.3264565808984162_adj.mid
0.06402123326764841_adj.mid

0.21198476749665085_adj.mid
0.3620067189216978_adj.mid
0.09167358800381353_adj.mid
0.7491289879531658_adj.mid
0.3559970176888735_adj.mid
0.549470161204349_adj.mid
0.10250888365879718_adj.mid
0.36321860283443286_adj.mid
0.047574444458241216_adj.mid
0.21901852969811753_adj.mid
0.905499891236923_adj.mid
0.9511403301279795_adj.mid
0.07186746659481313_adj.mid
0.22047111832936942_adj.mid
0.539512676743813_adj.mid
0.033313986422223163_adj.mid
0.981087291054314_adj.mid
0.10222964826466285_adj.mid
0.337517805339117_adj.mid
0.1755252422917658_adj.mid
0.5807329043589801_adj.mid
0.26551079719260606_adj.mid
0.33695573887576447_adj.mid
0.002716920481628_adj.mid
0.10630249969742178_adj.mid
0.8014751784512073_adj.mid