Objective

- Build effective music genre classification models using a variety of machine learning techniques
- Accurately classify genre of new music tracks with associated features

Dataset

Our Free Music Archive (FMA) dataset includes 106,574 tracks of music, split into 16 different genres, with 518 associated features extracted with LibROSA and Echonest.

Preprocessing

- Conversion of audio raw data to 518 audio features using Python package LibROSA
- Splitting of all 25000 tracks into training set, validation set and test set, each with size of 19922, 2505, 2573
- Random shuffling of all training data

Models

1. Softmax
   \[ p(y = i | x; \theta) = \phi_i = \frac{\exp(\theta_i^T x)}{\sum_k \exp(\theta_k^T x)} \]

2. SVM-RBF
   \[ \min_{\gamma, w, b} \frac{1}{2} ||w||^2 \]
   \[ \text{s.t. } y_i (w^T x_i + b) \geq \gamma_i, \ i = 1, ..., m \]
   \[ L_{\text{binary}} (z, y) = \max (0, 1 - yz) \]
   \[ x(x, z) = \exp \left( -\frac{1}{2} ||x - z||^2 \right) \]

3. Neural Network

   Neural Networks with sigmoid (layer 2 and 3) and softmax (layer 4) as activation functions. Features selection applied for overfitting reduction.

4. KNN-PCA/Model Selection

Results

<table>
<thead>
<tr>
<th>Models</th>
<th>Train accuracy</th>
<th>Test accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Softmax</td>
<td>52.87%</td>
<td>51.03%</td>
</tr>
<tr>
<td>Logistic Regression</td>
<td>67.45%</td>
<td>62.61%</td>
</tr>
<tr>
<td>Neural Network</td>
<td>74.93%</td>
<td>63.19%</td>
</tr>
<tr>
<td>SVM Linear</td>
<td>67.38%</td>
<td>61.48%</td>
</tr>
<tr>
<td>SVM RBF</td>
<td>85.4%</td>
<td>66.07%</td>
</tr>
<tr>
<td>KNN-Model Selection</td>
<td>99.98%</td>
<td>57.87%</td>
</tr>
</tbody>
</table>

- SVM with RBF kernel gains test accuracy of 66.07%, increasing 30% from baseline, Softmax model;
- Improvement in test accuracy by L1 regularization and best subset feature selection

Future work

- Model ensembling: combining classifiers by voting or averaging to improve performance
- Feature refining: add other musically relevant features for better classification results
- Real application: new music tracks can turn into features the same way as we mentioned, and applied our machine learning models to predict its genre.

References