

Music Genre Classification with LSTMs

Feature extraction

Spectral features 🔗

<code>chroma_stft</code> <code>(([y, sr, S, norm, n_fft, ...])</code>	Compute a chromagram from a waveform or power spec
<code>chroma_cqt</code> <code>(([y, sr, C, hop_length, fmin, ...])</code>	Constant-Q chromagram
<code>chroma_cens</code> <code>(([y, sr, C, hop_length, fmin, ...])</code>	Computes the chroma variant "Chroma Energy Normaliz
<code>melspectrogram</code> <code>(([y, sr, S, n_fft, ...])</code>	Compute a mel-scaled spectrogram.
<code>mfcc</code> <code>(([y, sr, S, n_mfcc, dct_type, norm])</code>	Mel-frequency cepstral coefficients (MFCCs)
<code>rmse</code> <code>(([y, S, frame_length, hop_length, ...])</code>	Compute root-mean-square (RMS) energy for each fram
<code>spectral_centroid</code> <code>(([y, sr, S, n_fft, ...])</code>	Compute the spectral centroid.
<code>spectral_bandwidth</code> <code>(([y, sr, S, n_fft, ...])</code>	Compute p'th-order spectral bandwidth:
<code>spectral_contrast</code> <code>(([y, sr, S, n_fft, ...])</code>	Compute spectral contrast [R6fcc01153df-1]
<code>spectral_flatness</code> <code>(([y, S, n_fft, hop_length, ...])</code>	Compute spectral flatness
<code>spectral_rolloff</code> <code>(([y, sr, S, n_fft, ...])</code>	Compute roll-off frequency
<code>poly_features</code> <code>(([y, sr, S, n_fft, hop_length, ...])</code>	Get coefficients of fitting an nth-order polynomial to the
<code>tonnetz</code> <code>(([y, sr, chroma])</code>	Computes the tonal centroid features (tonnetz), followin
<code>zero_crossing_rate</code> <code>([y[, frame_length, ...])</code>	Compute the zero-crossing rate of an audio time series.

经大量研究表明, MFCCs (Mel Frequency Cepstral Coefficents) 是一种在自动语音和说话人识别中广泛使用的特征。在语音识别领域, MFCCs 在人工特征方面被证明十分有效, 所以在此保留 `mfcc` 这一特征。

此外, `chroma_stft`, `chroma_cqt`, `chroma_cens` 相似性很高, 可视为一种特征, 同理, `spectral_centroid`, `spectral_bandwidth`, `spectral_contrast`, `spectral_flatness`, `spectral_rolloff` 也可视为一种特征。

因此可得如下特征组合实验结果。

1.

```
for i, file in enumerate(list_of_audiofiles):
    y, sr = librosa.load(file)
    mfcc = librosa.feature.mfcc(y=y, sr=sr, hop_length=self.hop_length, n_mfcc=13)
    spectral_center = librosa.feature.spectral_centroid(y=y, sr=sr, hop_length=self.hop_length)
    chroma = librosa.feature.chroma_stft(y=y, sr=sr, hop_length=self.hop_length)
    rmse = librosa.feature.rmse(y=y, hop_length=self.hop_length)

    splits = re.split(' [.]', file)
    genre = re.split(' [/]', splits[1])[3]
    target.append(genre)

    data[i, :, 0:13] = mfcc.T[0:timeseries_length, :]
    data[i, :, 13:14] = spectral_center.T[0:timeseries_length, :]
    data[i, :, 14:26] = chroma.T[0:timeseries_length, :]
    data[i, :, 26:33] = rmse.T[0:timeseries_length, :]
```

Epoch 400/400

```
35/420 [=>.....] - ETA: 4s - loss: 0.3981 - acc: 0.8286
70/420 [====>.....] - ETA: 4s - loss: 0.3539 - acc: 0.8714
105/420 [=====>.....] - ETA: 3s - loss: 0.3192 - acc: 0.8857
140/420 [======>.....] - ETA: 3s - loss: 0.3459 - acc: 0.8643
175/420 [======>.....] - ETA: 2s - loss: 0.3430 - acc: 0.8571
210/420 [======>.....] - ETA: 2s - loss: 0.3372 - acc: 0.8571
245/420 [======>.....] - ETA: 2s - loss: 0.3379 - acc: 0.8612
280/420 [======>.....] - ETA: 1s - loss: 0.3226 - acc: 0.8679
315/420 [======>.....] - ETA: 1s - loss: 0.3272 - acc: 0.8635
350/420 [======>.....] - ETA: 0s - loss: 0.3336 - acc: 0.8686
385/420 [======>...] - ETA: 0s - loss: 0.3415 - acc: 0.8675
420/420 [=====] - 5s 12ms/step - loss: 0.3838 - acc: 0.8595
```

Validating ...

```
35/120 [=====>.....] - ETA: 1s
70/120 [======>.....] - ETA: 0s
105/120 [======>....] - ETA: 0s
120/120 [=====] - 1s 7ms/step
Dev loss: 1.373569389184316
Dev accuracy: 0.5833333544433117
```

Testing ...

```
35/60 [======>.....] - ETA: 0s
60/60 [=====] - 0s 5ms/step
Test loss: 1.2082031120856602
Test accuracy: 0.6166666721304258
```

2.

```
for i, file in enumerate(list_of_audiofiles):
    y, sr = librosa.load(file)

    mfcc = librosa.feature.mfcc(y=y, sr=sr, hop_length=self.hop_length, n_mfcc=13)
    spectral_center = librosa.feature.spectral_bandwidth(y=y, sr=sr, hop_length=self.hop_length)
    chroma = librosa.feature.chroma_stft(y=y, sr=sr, hop_length=self.hop_length)
    rmse = librosa.feature.rmse(y=y, hop_length=self.hop_length)

    splits = re.split('[.]', file)
    genre = re.split('[/]', splits[1])[3]
    target.append(genre)

    data[i, :, 0:13] = mfcc.T[0:timeseries_length, :]
    data[i, :, 13:14] = spectral_center.T[0:timeseries_length, :]
    data[i, :, 14:26] = chroma.T[0:timeseries_length, :]
    data[i, :, 26:33] = rmse.T[0:timeseries_length, :]
```

Epoch 400/400

```
35/420 [=>.....] - ETA: 3s - loss: 0.3491 - acc: 0.8857
70/420 [====>.....] - ETA: 3s - loss: 0.3915 - acc: 0.8286
105/420 [=====>.....] - ETA: 3s - loss: 0.4367 - acc: 0.8190
140/420 [=====>.....] - ETA: 2s - loss: 0.4475 - acc: 0.8214
175/420 [=====>.....] - ETA: 2s - loss: 0.4243 - acc: 0.8343
210/420 [=====>.....] - ETA: 2s - loss: 0.4109 - acc: 0.8429
245/420 [=====>.....] - ETA: 1s - loss: 0.4001 - acc: 0.8449
280/420 [=====>.....] - ETA: 1s - loss: 0.4070 - acc: 0.8393
315/420 [=====>.....] - ETA: 1s - loss: 0.4365 - acc: 0.8349
350/420 [=====>.....] - ETA: 0s - loss: 0.4444 - acc: 0.8286
385/420 [=====>.....] - ETA: 0s - loss: 0.4347 - acc: 0.8338
420/420 [=====>.....] - 4s 10ms/step - loss: 0.4330 - acc: 0.8333
```

Validating ...

```
35/120 [=====>.....] - ETA: 1s
70/120 [=====>.....] - ETA: 0s
105/120 [=====>.....] - ETA: 0s
120/120 [=====>.....] - 1s 7ms/step
Dev loss: 0.9716353416442871
Dev accuracy: 0.7083333507180214
```

Testing ...

```
35/60 [=====>.....] - ETA: 0s
60/60 [=====>.....] - 0s 2ms/step
Test loss: 1.0784357984860737
Test accuracy: 0.6666666865348816
```

3.

```
for i, file in enumerate(list_of_audiofiles):
    y, sr = librosa.load(file)
    mfcc = librosa.feature.mfcc(y=y, sr=sr, hop_length=self.hop_length, n_mfcc=13)
    spectral_center = librosa.feature.spectral_flatness(y=y, hop_length=self.hop_length)
    chroma = librosa.feature.chroma_stft(y=y, sr=sr, hop_length=self.hop_length)
    rmse = librosa.feature.rmse(y=y, hop_length=self.hop_length)

    splits = re.split(' [.]', file)
    genre = re.split(' [/]', splits[1])[3]
    target.append(genre)

    data[i, :, 0:13] = mfcc.T[0:timeseries_length, :]
    data[i, :, 13:14] = spectral_center.T[0:timeseries_length, :]
    data[i, :, 14:26] = chroma.T[0:timeseries_length, :]
    data[i, :, 26:33] = rmse.T[0:timeseries_length, :]
```

Epoch 400/400

```
35/420 [=>.....] - ETA: 4s - loss: 0.0208 - acc: 1.0000
70/420 [====>.....] - ETA: 4s - loss: 0.0180 - acc: 1.0000
105/420 [=====>.....] - ETA: 3s - loss: 0.0215 - acc: 1.0000
140/420 [=====>.....] - ETA: 3s - loss: 0.0181 - acc: 1.0000
175/420 [=====>.....] - ETA: 3s - loss: 0.0173 - acc: 1.0000
210/420 [=====>.....] - ETA: 2s - loss: 0.0177 - acc: 1.0000
245/420 [=====>.....] - ETA: 2s - loss: 0.0222 - acc: 0.9959
280/420 [=====>.....] - ETA: 1s - loss: 0.0394 - acc: 0.9929
315/420 [=====>.....] - ETA: 1s - loss: 0.0366 - acc: 0.9937
350/420 [=====>.....] - ETA: 0s - loss: 0.0336 - acc: 0.9943
385/420 [=====>.....] - ETA: 0s - loss: 0.0354 - acc: 0.9922
420/420 [=====>.....] - 5s 12ms/step - loss: 0.0331 - acc: 0.9929
```

Validating ...

```
35/120 [=====>.....] - ETA: 1s
70/120 [=====>.....] - ETA: 0s
105/120 [=====>.....] - ETA: 0s
120/120 [=====>.....] - 1s 7ms/step
```

Dev loss: 2.27315000196298

Dev accuracy: 0.6166666829958558

Testing ...

```
35/60 [=====>.....] - ETA: 0s
60/60 [=====>.....] - 0s 5ms/step
```

Test loss: 2.655836800734202

Test accuracy: 0.5333333512147268

4.

```
for i, file in enumerate(list_of_audiofiles):
    y, sr = librosa.load(file)

    mfcc = librosa.feature.mfcc(y=y, sr=sr, hop_length=self.hop_length, n_mfcc=13)
    spectral_center = librosa.feature.spectral_rolloff(y=y, sr=sr, hop_length=self.hop_length)
    chroma = librosa.feature.chroma_stft(y=y, sr=sr, hop_length=self.hop_length)
    rmse = librosa.feature.rmse(y=y, hop_length=self.hop_length)

    splits = re.split('[.]', file)
    genre = re.split('[/]', splits[1])[3]
    target.append(genre)

    data[i, :, 0:13] = mfcc.T[0:timeseries_length, :]
    data[i, :, 13:14] = spectral_center.T[0:timeseries_length, :]
    data[i, :, 14:26] = chroma.T[0:timeseries_length, :]
    data[i, :, 26:33] = rmse.T[0:timeseries_length, :]
```

Epoch 400/400

```
35/420 [=>.....] - ETA: 3s - loss: 0.5658 - acc: 0.8000
70/420 [====>.....] - ETA: 3s - loss: 0.4420 - acc: 0.8286
105/420 [=====>.....] - ETA: 3s - loss: 0.4326 - acc: 0.8381
140/420 [=====>.....] - ETA: 2s - loss: 0.4936 - acc: 0.7929
175/420 [=====>.....] - ETA: 2s - loss: 0.5304 - acc: 0.7771
210/420 [=====>.....] - ETA: 2s - loss: 0.5188 - acc: 0.7810
245/420 [=====>.....] - ETA: 2s - loss: 0.5043 - acc: 0.7878
280/420 [=====>.....] - ETA: 1s - loss: 0.5303 - acc: 0.7714
315/420 [=====>.....] - ETA: 1s - loss: 0.5046 - acc: 0.7841
350/420 [=====>.....] - ETA: 0s - loss: 0.5066 - acc: 0.7914
385/420 [=====>.....] - ETA: 0s - loss: 0.5076 - acc: 0.7948
420/420 [=====>.....] - 5s 12ms/step - loss: 0.5159 - acc: 0.7952
```

Validating ...

```
35/120 [=====>.....] - ETA: 2s
70/120 [=====>.....] - ETA: 0s
105/120 [=====>.....] - ETA: 0s
120/120 [=====>.....] - 1s 11ms/step
Dev loss: 0.9538268509010474
Dev accuracy: 0.6583333450059096
```

Testing ...

```
35/60 [=====>.....] - ETA: 0s
60/60 [=====>.....] - 0s 2ms/step
Test loss: 0.9062086641788483
Test accuracy: 0.6333333402872086
```

5.

```
for i, file in enumerate(list_of_audiofiles):
    y, sr = librosa.load(file)
    mfcc = librosa.feature.mfcc(y=y, sr=sr, hop_length=self.hop_length, n_mfcc=13)
    spectral_center = librosa.feature.spectral_centroid(y=y, hop_length=self.hop_length)
    chroma = librosa.feature.chroma_stft(y=y, sr=sr, hop_length=self.hop_length)
    zero_crossing_rate = librosa.feature.zero_crossing_rate(y=y)

    splits = re.split(' [.]', file)
    genre = re.split(' [/]', splits[1])[3]
    target.append(genre)

    data[i, :, 0:13] = mfcc.T[0:timeseries_length, :]
    data[i, :, 13:14] = spectral_center.T[0:timeseries_length, :]
    data[i, :, 14:26] = chroma.T[0:timeseries_length, :]
    data[i, :, 26:33] = zero_crossing_rate.T[0:timeseries_length, :]
```

Epoch 400/400

```
35/420 [=>.....] - ETA: 4s - loss: 0.5204 - acc: 0.7714
70/420 [====>.....] - ETA: 4s - loss: 0.4260 - acc: 0.8143
105/420 [=====>.....] - ETA: 3s - loss: 0.4514 - acc: 0.8095
140/420 [=====>.....] - ETA: 3s - loss: 0.4494 - acc: 0.8214
175/420 [=====>.....] - ETA: 2s - loss: 0.4155 - acc: 0.8457
210/420 [=====>.....] - ETA: 2s - loss: 0.3908 - acc: 0.8524
245/420 [=====>.....] - ETA: 2s - loss: 0.3719 - acc: 0.8571
280/420 [=====>.....] - ETA: 1s - loss: 0.3807 - acc: 0.8571
315/420 [=====>.....] - ETA: 1s - loss: 0.3891 - acc: 0.8476
350/420 [=====>.....] - ETA: 0s - loss: 0.4125 - acc: 0.8429
385/420 [=====>.....] - ETA: 0s - loss: 0.4103 - acc: 0.8468
420/420 [=====>.....] - 5s 12ms/step - loss: 0.4061 - acc: 0.8476
```

Validating ...

```
35/120 [=====>.....] - ETA: 1s
70/120 [=====>.....] - ETA: 0s
105/120 [=====>.....] - ETA: 0s
120/120 [=====>.....] - 1s 8ms/step
Dev loss: 1.0691008294622104
Dev accuracy: 0.6416666842997074
```

Testing ...

```
35/60 [=====>.....] - ETA: 0s
60/60 [=====>.....] - 0s 5ms/step
Test loss: 1.193656325340271
Test accuracy: 0.5500000094374021
```

6.

```
for i, file in enumerate(list_of_audiofiles):
    y, sr = librosa.load(file)

    mfcc = librosa.feature.mfcc(y=y, sr=sr, hop_length=self.hop_length, n_mfcc=13)
    spectral_center = librosa.feature.spectral_bandwidth(y=y, sr=sr, hop_length=self.hop_length)
    chroma = librosa.feature.chroma_stft(y=y, sr=sr, hop_length=self.hop_length)
    zero_crossing_rate = librosa.feature.zero_crossing_rate(y=y)

    splits = re.split('[.]', file)
    genre = re.split('[ /]', splits[1])[3]
    target.append(genre)

    data[i, :, 0:13] = mfcc.T[0:timeseries_length, :]
    data[i, :, 13:14] = spectral_center.T[0:timeseries_length, :]
    data[i, :, 14:26] = chroma.T[0:timeseries_length, :]
    data[i, :, 26:33] = zero_crossing_rate.T[0:timeseries_length, :]
```

Epoch 400/400

```
35/420 [=>.....] - ETA: 8s - loss: 0.4538 - acc: 0.8571
70/420 [====>.....] - ETA: 7s - loss: 0.4509 - acc: 0.8143
105/420 [=====>.....] - ETA: 6s - loss: 0.4328 - acc: 0.8286
140/420 [=====>.....] - ETA: 5s - loss: 0.4207 - acc: 0.8357
175/420 [======>.....] - ETA: 5s - loss: 0.4036 - acc: 0.8400
210/420 [======>.....] - ETA: 4s - loss: 0.4281 - acc: 0.8381
245/420 [======>.....] - ETA: 3s - loss: 0.4271 - acc: 0.8367
280/420 [======>.....] - ETA: 2s - loss: 0.4169 - acc: 0.8464
315/420 [======>.....] - ETA: 2s - loss: 0.4123 - acc: 0.8508
350/420 [======>.....] - ETA: 1s - loss: 0.4108 - acc: 0.8486
385/420 [======>.....] - ETA: 0s - loss: 0.4165 - acc: 0.8442
420/420 [=====] - 8s 20ms/step - loss: 0.4412 - acc: 0.8381
```

Validating ...

```
35/120 [=====>.....] - ETA: 5s
70/120 [=====>.....] - ETA: 1s
105/120 [======>.....] - ETA: 0s
120/120 [=====] - 3s 23ms/step
Dev loss: 1.2569773321350415
Dev accuracy: 0.6166666808227698
```

Testing ...

```
35/60 [======>.....] - ETA: 0s
60/60 [=====] - 0s 4ms/step
Test loss: 1.3598987857500713
Test accuracy: 0.5500000268220901
```


7.

```
for i, file in enumerate(list_of_audiofiles):
    y, sr = librosa.load(file)

    mfcc = librosa.feature.mfcc(y=y, sr=sr, hop_length=self.hop_length, n_mfcc=13)
    spectral_center = librosa.feature.spectral_flatness(y=y, hop_length=self.hop_length)
    chroma = librosa.feature.chroma_stft(y=y, sr=sr, hop_length=self.hop_length)
    zero_crossing_rate = librosa.feature.zero_crossing_rate(y=y, hop_length=self.hop_length)

    splits = re.split('[.]', file)
    genre = re.split('[ /]', splits[1])[3]
    target.append(genre)

    data[i, :, 0:13] = mfcc.T[0:timeseries_length, :]
    data[i, :, 13:14] = spectral_center.T[0:timeseries_length, :]
    data[i, :, 14:26] = chroma.T[0:timeseries_length, :]
    data[i, :, 26:33] = zero_crossing_rate.T[0:timeseries_length, :]
```

Epoch 400/400

```
35/420 [=>.....] - ETA: 3s - loss: 0.1554 - acc: 0.9429
70/420 [====>.....] - ETA: 3s - loss: 0.0852 - acc: 0.9714
105/420 [=====>.....] - ETA: 2s - loss: 0.0930 - acc: 0.9619
140/420 [=====>.....] - ETA: 2s - loss: 0.0707 - acc: 0.9714
175/420 [=====>.....] - ETA: 2s - loss: 0.0603 - acc: 0.9771
210/420 [=====>.....] - ETA: 1s - loss: 0.0510 - acc: 0.9810
245/420 [=====>.....] - ETA: 1s - loss: 0.0473 - acc: 0.9837
280/420 [=====>.....] - ETA: 1s - loss: 0.0613 - acc: 0.9821
315/420 [=====>.....] - ETA: 0s - loss: 0.0728 - acc: 0.9778
350/420 [=====>.....] - ETA: 0s - loss: 0.0688 - acc: 0.9800
385/420 [=====>.....] - ETA: 0s - loss: 0.0727 - acc: 0.9792
420/420 [=====>.....] - 4s 9ms/step - loss: 0.0675 - acc: 0.9810
```

Validating ...

```
35/120 [=====>.....] - ETA: 1s
70/120 [=====>.....] - ETA: 0s
105/120 [=====>.....] - ETA: 0s
120/120 [=====>.....] - 1s 7ms/step
Dev loss: 2.0576821342110634
Dev accuracy: 0.633333351883698
```

Testing ...

```
35/60 [=====>.....] - ETA: 0s
60/60 [=====>.....] - 0s 2ms/step
Test loss: 2.2607963581879935
Test accuracy: 0.533333328862985
```


8.

```
for i, file in enumerate(list_of_audiofiles):
    y, sr = librosa.load(file)
    mfcc = librosa.feature.mfcc(y=y, sr=sr, hop_length=self.hop_length, n_mfcc=13)
    spectral_center = librosa.feature.spectral_rolloff(y=y, hop_length=self.hop_length)
    chroma = librosa.feature.chroma_stft(y=y, sr=sr, hop_length=self.hop_length)
    zero_crossing_rate = librosa.feature.zero_crossing_rate(y=y)

    splits = re.split(' [.]', file)
    genre = re.split(' [/]', splits[1])[3]
    target.append(genre)

    data[i, :, 0:13] = mfcc.T[0:timeseries_length, :]
    data[i, :, 13:14] = spectral_center.T[0:timeseries_length, :]
    data[i, :, 14:26] = chroma.T[0:timeseries_length, :]
    data[i, :, 26:33] = zero_crossing_rate.T[0:timeseries_length, :]
```

Epoch 400/400

```
35/420 [=>.....] - ETA: 4s - loss: 0.4505 - acc: 0.8000
70/420 [====>.....] - ETA: 4s - loss: 0.4440 - acc: 0.8143
105/420 [=====>.....] - ETA: 3s - loss: 0.4322 - acc: 0.8286
140/420 [=====>.....] - ETA: 3s - loss: 0.4767 - acc: 0.8143
175/420 [=====>.....] - ETA: 2s - loss: 0.4647 - acc: 0.8171
210/420 [=====>.....] - ETA: 2s - loss: 0.4810 - acc: 0.8048
245/420 [=====>.....] - ETA: 2s - loss: 0.4850 - acc: 0.8122
280/420 [=====>.....] - ETA: 1s - loss: 0.5203 - acc: 0.8071
315/420 [=====>.....] - ETA: 1s - loss: 0.5180 - acc: 0.8032
350/420 [=====>.....] - ETA: 0s - loss: 0.5303 - acc: 0.7971
385/420 [=====>....] - ETA: 0s - loss: 0.5399 - acc: 0.7870
420/420 [=====>====] - 5s 12ms/step - loss: 0.5269 - acc: 0.7976
```

Validating ...

```
35/120 [=====>.....] - ETA: 1s
70/120 [=====>.....] - ETA: 0s
105/120 [=====>.....] - ETA: 0s
120/120 [=====>====] - 1s 7ms/step
Dev loss: 0.9134505813320478
Dev accuracy: 0.6583333536982536
```

Testing ...

```
35/60 [=====>.....] - ETA: 0s
60/60 [=====>====] - 0s 5ms/step
Test loss: 1.199605683485667
Test accuracy: 0.5500000044703484
```

9.

```
for i, file in enumerate(list_of_audiofiles):
    y, sr = librosa.load(file)
    mfcc = librosa.feature.mfcc(y=y, sr=sr, hop_length=self.hop_length, n_mfcc=13)
    spectral_center = librosa.feature.spectral_centroid(y=y, hop_length=self.hop_length)
    chroma = librosa.feature.chroma_cqt(y=y, sr=sr, hop_length=self.hop_length)
    rmse = librosa.feature.rmse(y=y, hop_length=self.hop_length)

    splits = re.split('[.]', file)
    genre = re.split('[ /]', splits[1])[3]
    target.append(genre)

    data[i, :, 0:13] = mfcc.T[0:timeseries_length, :]
    data[i, :, 13:14] = spectral_center.T[0:timeseries_length, :]
    data[i, :, 14:26] = chroma.T[0:timeseries_length, :]
    data[i, :, 26:33] = rmse.T[0:timeseries_length, :]
```

Epoch 400/400

```
35/420 [=>.....] - ETA: 4s - loss: 0.1938 - acc: 0.9143
70/420 [====>.....] - ETA: 4s - loss: 0.2412 - acc: 0.9000
105/420 [=====>.....] - ETA: 3s - loss: 0.2878 - acc: 0.8952
140/420 [=====>.....] - ETA: 3s - loss: 0.2778 - acc: 0.9071
175/420 [=====>.....] - ETA: 3s - loss: 0.2677 - acc: 0.9086
210/420 [=====>.....] - ETA: 2s - loss: 0.2925 - acc: 0.9048
245/420 [=====>.....] - ETA: 2s - loss: 0.2869 - acc: 0.8980
280/420 [=====>.....] - ETA: 1s - loss: 0.2831 - acc: 0.9000
315/420 [=====>.....] - ETA: 1s - loss: 0.2996 - acc: 0.8921
350/420 [=====>.....] - ETA: 0s - loss: 0.3010 - acc: 0.8886
385/420 [=====>...] - ETA: 0s - loss: 0.3332 - acc: 0.8753
420/420 [=====] - 5s 12ms/step - loss: 0.3233 - acc: 0.8810
```

Validating ...

```
35/120 [=====>.....] - ETA: 1s
70/120 [=====>.....] - ETA: 0s
105/120 [=====>....] - ETA: 0s
120/120 [=====] - 1s 7ms/step
Dev loss: 1.1678892870744069
Dev accuracy: 0.6583333536982536
```

Testing ...

```
35/60 [=====>.....] - ETA: 0s
60/60 [=====] - 0s 5ms/step
Test loss: 1.2402155896027882
Test accuracy: 0.6500000109275182
```

10.

```
for i, file in enumerate(list_of_audiofiles):
    y, sr = librosa.load(file)
    mfcc = librosa.feature.mfcc(y=y, sr=sr, hop_length=self.hop_length, n_mfcc=13)
    spectral_center = librosa.feature.spectral_bandwidth(y=y, sr=sr, hop_length=self.hop_length)
    chroma = librosa.feature.chroma_cqt(y=y, sr=sr, hop_length=self.hop_length)
    rmse = librosa.feature.rmse(y=y)

    splits = re.split(' [.]', file)
    genre = re.split(' [/]', splits[1])[3]
    target.append(genre)

    data[i, :, 0:13] = mfcc.T[0:timeseries_length, :]
    data[i, :, 13:14] = spectral_center.T[0:timeseries_length, :]
    data[i, :, 14:26] = chroma.T[0:timeseries_length, :]
    data[i, :, 26:33] = rmse.T[0:timeseries_length, :]
```

Epoch 400/400

```
35/420 [=>.....] - ETA: 4s - loss: 0.5335 - acc: 0.7143
70/420 [====>.....] - ETA: 4s - loss: 0.5802 - acc: 0.7286
105/420 [=====>.....] - ETA: 3s - loss: 0.4968 - acc: 0.8000
140/420 [=====>.....] - ETA: 3s - loss: 0.5133 - acc: 0.7929
175/420 [=====>.....] - ETA: 3s - loss: 0.4896 - acc: 0.8000
210/420 [=====>.....] - ETA: 2s - loss: 0.5001 - acc: 0.7952
245/420 [=====>.....] - ETA: 2s - loss: 0.4949 - acc: 0.7918
280/420 [=====>.....] - ETA: 1s - loss: 0.5077 - acc: 0.7893
315/420 [=====>.....] - ETA: 1s - loss: 0.5133 - acc: 0.7841
350/420 [=====>.....] - ETA: 0s - loss: 0.5151 - acc: 0.7829
385/420 [=====>.....] - ETA: 0s - loss: 0.5163 - acc: 0.7870
420/420 [=====>.....] - 5s 12ms/step - loss: 0.4985 - acc: 0.7976
```

Validating ...

```
35/120 [=====>.....] - ETA: 1s
70/120 [=====>.....] - ETA: 0s
105/120 [=====>.....] - ETA: 0s
120/120 [=====>.....] - 1s 7ms/step
```

Dev loss: 1.0489776755372684

Dev accuracy: 0.6083333510905504

Testing ...

```
35/60 [=====>.....] - ETA: 0s
60/60 [=====>.....] - 0s 5ms/step
```

Test loss: 1.5920234620571136

Test accuracy: 0.5333333512147268

11.

```
for i, file in enumerate(list_of_audiofiles):
    y, sr = librosa.load(file)

    mfcc = librosa.feature.mfcc(y=y, sr=sr, hop_length=self.hop_length, n_mfcc=13)
    spectral_center = librosa.feature.spectral_flatness(y=y, hop_length=self.hop_length)
    chroma = librosa.feature.chroma_cqt(y=y, sr=sr, hop_length=self.hop_length)
    rmse = librosa.feature.rmse(y=y, hop_length=self.hop_length)

    splits = re.split('[.]', file)
    genre = re.split('[ /]', splits[1])[3]
    target.append(genre)

    data[i, :, 0:13] = mfcc.T[0:timeseries_length, :]
    data[i, :, 13:14] = spectral_center.T[0:timeseries_length, :]
    data[i, :, 14:26] = chroma.T[0:timeseries_length, :]
    data[i, :, 26:33] = rmse.T[0:timeseries_length, :]
```

Epoch 400/400

```
35/420 [=>.....] - ETA: 2s - loss: 0.0080 - acc: 1.0000
70/420 [====>.....] - ETA: 2s - loss: 0.0571 - acc: 0.9714
105/420 [=====>.....] - ETA: 2s - loss: 0.0568 - acc: 0.9714
140/420 [=====>.....] - ETA: 2s - loss: 0.0504 - acc: 0.9786
175/420 [=====>.....] - ETA: 2s - loss: 0.0418 - acc: 0.9829
210/420 [=====>.....] - ETA: 1s - loss: 0.0633 - acc: 0.9762
245/420 [=====>.....] - ETA: 1s - loss: 0.0570 - acc: 0.9796
280/420 [=====>.....] - ETA: 1s - loss: 0.0505 - acc: 0.9821
315/420 [=====>.....] - ETA: 0s - loss: 0.0474 - acc: 0.9841
350/420 [=====>.....] - ETA: 0s - loss: 0.0601 - acc: 0.9800
385/420 [=====>.....] - ETA: 0s - loss: 0.0552 - acc: 0.9818
420/420 [=====>.....] - 3s 8ms/step - loss: 0.0519 - acc: 0.9833
```

Validating ...

```
35/120 [=====>.....] - ETA: 0s
105/120 [=====>.....] - ETA: 0s
120/120 [=====>.....] - 0s 4ms/step
Dev loss: 1.5269340202212334
Dev accuracy: 0.6833333522081375
```

Testing ...

```
35/60 [=====>.....] - ETA: 0s
60/60 [=====>.....] - 0s 2ms/step
Test loss: 2.4769909381866455
Test accuracy: 0.500000124176344
```

12.

```
for i, file in enumerate(list_of_audiofiles):
    y, sr = librosa.load(file)
    mfcc = librosa.feature.mfcc(y=y, sr=sr, hop_length=self.hop_length, n_mfcc=13)
    spectral_center = librosa.feature.spectral_rolloff(y=y, hop_length=self.hop_length)
    chroma = librosa.feature.chroma_cqt(y=y, sr=sr, hop_length=self.hop_length)
    rmse = librosa.feature.rmse(y=y, hop_length=self.hop_length)

    splits = re.split(' [.]', file)
    genre = re.split(' [/]', splits[1])[3]
    target.append(genre)

    data[i, :, 0:13] = mfcc.T[0:timeseries_length, :]
    data[i, :, 13:14] = spectral_center.T[0:timeseries_length, :]
    data[i, :, 14:26] = chroma.T[0:timeseries_length, :]
    data[i, :, 26:33] = rmse.T[0:timeseries_length, :]
```

Epoch 400/400

```
35/420 [=>.....] - ETA: 4s - loss: 0.3650 - acc: 0.8571
70/420 [====>.....] - ETA: 3s - loss: 0.3188 - acc: 0.8714
105/420 [=====>.....] - ETA: 3s - loss: 0.4842 - acc: 0.8286
140/420 [=====>.....] - ETA: 3s - loss: 0.4743 - acc: 0.8143
175/420 [=====>.....] - ETA: 2s - loss: 0.4710 - acc: 0.8229
210/420 [=====>.....] - ETA: 2s - loss: 0.4542 - acc: 0.8333
245/420 [=====>.....] - ETA: 2s - loss: 0.4430 - acc: 0.8367
280/420 [=====>.....] - ETA: 1s - loss: 0.4404 - acc: 0.8393
315/420 [=====>.....] - ETA: 1s - loss: 0.4376 - acc: 0.8349
350/420 [=====>.....] - ETA: 0s - loss: 0.4282 - acc: 0.8429
385/420 [=====>.....] - ETA: 0s - loss: 0.4110 - acc: 0.8494
420/420 [=====>.....] - 5s 12ms/step - loss: 0.4000 - acc: 0.8571
```

Validating ...

```
35/120 [=====>.....] - ETA: 1s
70/120 [=====>.....] - ETA: 0s
105/120 [=====>.....] - ETA: 0s
120/120 [=====>.....] - 1s 7ms/step
Dev loss: 1.1132373834649723
Dev accuracy: 0.6166666857898235
```

Testing ...

```
35/60 [=====>.....] - ETA: 0s
60/60 [=====>.....] - 0s 5ms/step
Test loss: 0.9893808513879776
Test accuracy: 0.6500000208616257
```

13.

```
for i, file in enumerate(list_of_audiofiles):
    y, sr = librosa.load(file)
    mfcc = librosa.feature.mfcc(y=y, sr=sr, hop_length=self.hop_length, n_mfcc=13)
    spectral_center = librosa.feature.spectral_centroid(y=y, sr=sr, hop_length=self.hop_length)
    chroma = librosa.feature.chroma_cens(y=y, sr=sr, hop_length=self.hop_length)
    rmse = librosa.feature.rmse(y=y, hop_length=self.hop_length)

    splits = re.split(' [.]', file)
    genre = re.split(' [/]', splits[1])[3]
    target.append(genre)

    data[i, :, 0:13] = mfcc.T[0:timeseries_length, :]
    data[i, :, 13:14] = spectral_center.T[0:timeseries_length, :]
    data[i, :, 14:26] = chroma.T[0:timeseries_length, :]
    data[i, :, 26:33] = rmse.T[0:timeseries_length, :]
```

Epoch 400/400

```
35/420 [=>.....] - ETA: 1s - loss: 0.2383 - acc: 0.9714
70/420 [=====>.....] - ETA: 1s - loss: 0.3148 - acc: 0.9000
105/420 [=====>.....] - ETA: 1s - loss: 0.4082 - acc: 0.8571
140/420 [=====>.....] - ETA: 1s - loss: 0.3624 - acc: 0.8786
175/420 [=====>.....] - ETA: 0s - loss: 0.3770 - acc: 0.8743
210/420 [=====>.....] - ETA: 0s - loss: 0.3632 - acc: 0.8810
245/420 [=====>.....] - ETA: 0s - loss: 0.3586 - acc: 0.8816
280/420 [=====>.....] - ETA: 0s - loss: 0.3301 - acc: 0.8929
315/420 [=====>.....] - ETA: 0s - loss: 0.3383 - acc: 0.8889
350/420 [=====>.....] - ETA: 0s - loss: 0.3478 - acc: 0.8800
385/420 [=====>.....] - ETA: 0s - loss: 0.3613 - acc: 0.8701
420/420 [=====>.....] - 2s 4ms/step - loss: 0.3626 - acc: 0.8643
```

Validating ...

```
35/120 [=====>.....] - ETA: 0s
105/120 [=====>.....] - ETA: 0s
120/120 [=====>.....] - 0s 3ms/step
Dev loss: 1.1321315790216129
Dev accuracy: 0.6666666828095913
```

Testing ...

```
35/60 [=====>.....] - ETA: 0s
60/60 [=====>.....] - 0s 909us/step
Test loss: 1.2191942036151886
Test accuracy: 0.6833333422740301
```

14.

```
for i, file in enumerate(list_of_audiofiles):
    y, sr = librosa.load(file)
    mfcc = librosa.feature.mfcc(y=y, sr=sr, hop_length=self.hop_length, n_mfcc=13)
    spectral_center = librosa.feature.spectral_centroid(y=y, sr=sr, hop_length=self.hop_length)
    chroma = librosa.feature.chroma_cens(y=y, sr=sr, hop_length=self.hop_length)
    spectral_contrast = librosa.feature.spectral_contrast(y=y, sr=sr, hop_length=self.hop_length)

    splits = re.split('[.]', file)
    genre = re.split('[ /]', splits[1])[3]
    target.append(genre)

    data[i, :, 0:13] = mfcc.T[0:timeseries_length, :]
    data[i, :, 13:14] = spectral_center.T[0:timeseries_length, :]
    data[i, :, 14:26] = chroma.T[0:timeseries_length, :]
    data[i, :, 26:33] = spectral_contrast.T[0:timeseries_length, :]
```

Epoch 400/400

```
35/420 [=>.....] - ETA: 1s - loss: 0.3663 - acc: 0.8857
70/420 [====>.....] - ETA: 1s - loss: 0.2992 - acc: 0.9000
105/420 [=====>.....] - ETA: 1s - loss: 0.3019 - acc: 0.9048
140/420 [=====>.....] - ETA: 1s - loss: 0.2965 - acc: 0.8929
175/420 [=====>.....] - ETA: 1s - loss: 0.3143 - acc: 0.8800
210/420 [=====>.....] - ETA: 0s - loss: 0.3222 - acc: 0.8810
245/420 [=====>.....] - ETA: 0s - loss: 0.3409 - acc: 0.8735
280/420 [=====>.....] - ETA: 0s - loss: 0.3224 - acc: 0.8857
315/420 [=====>.....] - ETA: 0s - loss: 0.3129 - acc: 0.8889
350/420 [=====>.....] - ETA: 0s - loss: 0.3051 - acc: 0.8914
385/420 [=====>.....] - ETA: 0s - loss: 0.3115 - acc: 0.8857
420/420 [=====>.....] - 2s 4ms/step - loss: 0.3237 - acc: 0.8810
```

Validating ...

```
35/120 [=====>.....] - ETA: 0s
105/120 [=====>.....] - ETA: 0s
120/120 [=====>.....] - 0s 3ms/step
Dev loss: 1.0959213078022003
Dev accuracy: 0.6416666842997074
```

Testing ...

```
35/60 [=====>.....] - ETA: 0s
60/60 [=====>.....] - 0s 884us/step
Test loss: 1.1959176063537598
Test accuracy: 0.6500000208616257
```


15.

```
for i, file in enumerate(list_of_audiofiles):
    y, sr = librosa.load(file)
    mfcc = librosa.feature.mfcc(y=y, sr=sr, hop_length=self.hop_length, n_mfcc=13)
    spectral_center = librosa.feature.spectral_centroid(y=y, sr=sr, hop_length=self.hop_length)
    chroma = librosa.feature.chroma_cens(y=y, sr=sr, hop_length=self.hop_length)
    spectral_bandwidth = librosa.feature.spectral_bandwidth(y=y, sr=sr, hop_length=self.hop_length)

    splits = re.split(' [.]', file)
    genre = re.split(' [/]', splits[1])[3]
    target.append(genre)

    data[i, :, 0:13] = mfcc.T[0:timeseries_length, :]
    data[i, :, 13:14] = spectral_center.T[0:timeseries_length, :]
    data[i, :, 14:26] = chroma.T[0:timeseries_length, :]
    data[i, :, 26:33] = spectral_bandwidth.T[0:timeseries_length, :]
```

Epoch 400/400

```
35/420 [=>.....] - ETA: 1s - loss: 1.1163 - acc: 0.5143
70/420 [====>.....] - ETA: 1s - loss: 0.9843 - acc: 0.6286
105/420 [=====>.....] - ETA: 1s - loss: 1.0822 - acc: 0.6000
140/420 [=====>.....] - ETA: 1s - loss: 1.0109 - acc: 0.6357
175/420 [=====>.....] - ETA: 0s - loss: 0.9601 - acc: 0.6514
210/420 [=====>.....] - ETA: 0s - loss: 0.9193 - acc: 0.6667
245/420 [=====>.....] - ETA: 0s - loss: 0.9080 - acc: 0.6694
280/420 [=====>.....] - ETA: 0s - loss: 0.8786 - acc: 0.6750
315/420 [=====>.....] - ETA: 0s - loss: 0.8894 - acc: 0.6667
350/420 [=====>.....] - ETA: 0s - loss: 0.8771 - acc: 0.6714
385/420 [=====>.....] - ETA: 0s - loss: 0.8778 - acc: 0.6649
420/420 [=====>.....] - 2s 4ms/step - loss: 0.8833 - acc: 0.6667
```

Validating ...

```
35/120 [=====>.....] - ETA: 0s
105/120 [=====>.....] - ETA: 0s
120/120 [=====>.....] - 0s 3ms/step
Dev loss: 1.020694226026535
Dev accuracy: 0.6166666895151138
```

Testing ...

```
35/60 [=====>.....] - ETA: 0s
60/60 [=====>.....] - 0s 90µs/step
Test loss: 0.9509770373503367
Test accuracy: 0.5833333432674408
```

16.

```
for i, file in enumerate(list_of_audiofiles):
    y, sr = librosa.load(file)
    mfcc = librosa.feature.mfcc(y=y, sr=sr, hop_length=self.hop_length, n_mfcc=13)
    spectral_center = librosa.feature.spectral_centroid(y=y, sr=sr, hop_length=self.hop_length)
    chroma = librosa.feature.chroma_cens(y=y, sr=sr, hop_length=self.hop_length)
    spectral_flatness = librosa.feature.spectral_flatness(y=y, hop_length=self.hop_length)

    splits = re.split(' [.]', file)
    genre = re.split(' [/]', splits[1])[3]
    target.append(genre)

    data[i, :, 0:13] = mfcc.T[0:timeseries_length, :]
    data[i, :, 13:14] = spectral_center.T[0:timeseries_length, :]
    data[i, :, 14:26] = chroma.T[0:timeseries_length, :]
    data[i, :, 26:33] = spectral_flatness.T[0:timeseries_length, :]
```

Epoch 400/400

```
35/420 [=>.....] - ETA: 1s - loss: 0.2677 - acc: 0.8857
70/420 [====>.....] - ETA: 1s - loss: 0.3010 - acc: 0.8714
105/420 [=====>.....] - ETA: 1s - loss: 0.2863 - acc: 0.8952
140/420 [=====>.....] - ETA: 1s - loss: 0.2556 - acc: 0.9071
175/420 [=====>.....] - ETA: 0s - loss: 0.2629 - acc: 0.9029
210/420 [=====>.....] - ETA: 0s - loss: 0.2766 - acc: 0.8952
245/420 [=====>.....] - ETA: 0s - loss: 0.2941 - acc: 0.8980
280/420 [=====>.....] - ETA: 0s - loss: 0.2730 - acc: 0.9071
315/420 [=====>.....] - ETA: 0s - loss: 0.2676 - acc: 0.9079
350/420 [=====>.....] - ETA: 0s - loss: 0.2711 - acc: 0.9114
385/420 [=====>.....] - ETA: 0s - loss: 0.2693 - acc: 0.9117
420/420 [=====>.....] - 2s 4ms/step - loss: 0.2690 - acc: 0.9119
```

Validating ...

```
35/120 [=====>.....] - ETA: 0s
105/120 [=====>.....] - ETA: 0s
120/120 [=====>.....] - 0s 3ms/step
Dev loss: 0.8947887097795805
Dev accuracy: 0.7583333477377892
```

Testing ...

```
35/60 [=====>.....] - ETA: 0s
60/60 [=====>.....] - 0s 907us/step
Test loss: 1.029723157485326
Test accuracy: 0.6333333551883698
```

17.

```
for i, file in enumerate(list_of_audiofiles):
    y, sr = librosa.load(file)
    mfcc = librosa.feature.mfcc(y=y, sr=sr, hop_length=self.hop_length, n_mfcc=13)
    spectral_center = librosa.feature.spectral_centroid(y=y, sr=sr, hop_length=self.hop_length)
    chroma = librosa.feature.chroma_cens(y=y, sr=sr, hop_length=self.hop_length)
    spectral_rolloff = librosa.feature.spectral_rolloff(y=y, sr=sr, hop_length=self.hop_length)

    splits = re.split(' [.]', file)
    genre = re.split(' [/]', splits[1])[3]
    target.append(genre)

    data[i, :, 0:13] = mfcc.T[0:timeseries_length, :]
    data[i, :, 13:14] = spectral_center.T[0:timeseries_length, :]
    data[i, :, 14:26] = chroma.T[0:timeseries_length, :]
    data[i, :, 26:33] = spectral_rolloff.T[0:timeseries_length, :]
```

Epoch 400/400

```
35/420 [=>.....] - ETA: 2s - loss: 0.8835 - acc: 0.6857
70/420 [====>.....] - ETA: 1s - loss: 1.0689 - acc: 0.5571
105/420 [=====>.....] - ETA: 1s - loss: 1.0933 - acc: 0.5810
140/420 [=====>.....] - ETA: 1s - loss: 1.0484 - acc: 0.5929
175/420 [=====>.....] - ETA: 1s - loss: 1.0241 - acc: 0.6229
210/420 [=====>.....] - ETA: 1s - loss: 1.0300 - acc: 0.6190
245/420 [=====>.....] - ETA: 0s - loss: 1.0319 - acc: 0.6245
280/420 [=====>.....] - ETA: 0s - loss: 1.0403 - acc: 0.6214
315/420 [=====>.....] - ETA: 0s - loss: 1.0488 - acc: 0.6222
350/420 [=====>.....] - ETA: 0s - loss: 1.0733 - acc: 0.6086
385/420 [=====>.....] - ETA: 0s - loss: 1.0566 - acc: 0.6104
420/420 [=====>.....] - 2s 5ms/step - loss: 1.0416 - acc: 0.6143
```

Validating ...

```
35/120 [=====>.....] - ETA: 0s
105/120 [=====>.....] - ETA: 0s
120/120 [=====>.....] - 0s 3ms/step
Dev loss: 1.000188333292802
Dev accuracy: 0.6000000151495138
```

Testing ...

```
35/60 [=====>.....] - ETA: 0s
60/60 [=====>.....] - 0s 1ms/step
Test loss: 1.244888613621394
Test accuracy: 0.4666666661699613
```

18.

```
for i, file in enumerate(list_of_audiofiles):
    y, sr = librosa.load(file)
    mfcc = librosa.feature.mfcc(y=y, sr=sr, hop_length=self.hop_length, n_mfcc=13)
    spectral_center = librosa.feature.spectral_centroid(y=y, sr=sr, hop_length=self.hop_length)
    chroma = librosa.feature.chroma_cqt(y=y, sr=sr, hop_length=self.hop_length)
    spectral_bandwidth = librosa.feature.spectral_bandwidth(y=y, sr=sr, hop_length=self.hop_length)

    splits = re.split(' [.]', file)
    genre = re.split(' [/]', splits[1])[3]
    target.append(genre)

    data[i, :, 0:13] = mfcc.T[0:timeseries_length, :]
    data[i, :, 13:14] = spectral_center.T[0:timeseries_length, :]
    data[i, :, 14:26] = chroma.T[0:timeseries_length, :]
    data[i, :, 26:33] = spectral_bandwidth.T[0:timeseries_length, :]
```

Epoch 400/400

```
35/420 [=>.....] - ETA: 1s - loss: 1.0677 - acc: 0.4857
70/420 [====>.....] - ETA: 1s - loss: 1.1393 - acc: 0.5286
105/420 [=====>.....] - ETA: 1s - loss: 1.0420 - acc: 0.5429
140/420 [=====>.....] - ETA: 1s - loss: 0.9928 - acc: 0.5643
175/420 [=====>.....] - ETA: 0s - loss: 0.9586 - acc: 0.5886
210/420 [=====>.....] - ETA: 0s - loss: 0.9879 - acc: 0.5810
245/420 [=====>.....] - ETA: 0s - loss: 0.9900 - acc: 0.5878
280/420 [=====>.....] - ETA: 0s - loss: 0.9577 - acc: 0.6107
315/420 [=====>.....] - ETA: 0s - loss: 0.9553 - acc: 0.6063
350/420 [=====>.....] - ETA: 0s - loss: 0.9623 - acc: 0.6000
385/420 [=====>...] - ETA: 0s - loss: 0.9617 - acc: 0.5974
420/420 [=====] - 2s 4ms/step - loss: 0.9507 - acc: 0.6024
```

Validating ...

```
35/120 [=====>.....] - ETA: 0s
105/120 [=====>.....] - ETA: 0s
120/120 [=====] - 0s 3ms/step
Dev loss: 0.9821401908993721
Dev accuracy: 0.6250000136593977
```

Testing ...

```
35/60 [=====>.....] - ETA: 0s
60/60 [=====] - 0s 904us/step
Test loss: 0.9612920631965002
Test accuracy: 0.6833333472410837
```

19.

```
for i, file in enumerate(list_of_audiofiles):
    y, sr = librosa.load(file)
    mfcc = librosa.feature.mfcc(y=y, sr=sr, hop_length=self.hop_length, n_mfcc=13)
    spectral_center = librosa.feature.spectral_centroid(y=y, sr=sr, hop_length=self.hop_length)
    chroma = librosa.feature.chroma_cqt(y=y, sr=sr, hop_length=self.hop_length)
    spectral_contrast = librosa.feature.spectral_contrast(y=y, sr=sr, hop_length=self.hop_length)

    splits = re.split(' [.]', file)
    genre = re.split(' [/]', splits[1])[3]
    target.append(genre)

    data[i, :, 0:13] = mfcc.T[0:timeseries_length, :]
    data[i, :, 13:14] = spectral_center.T[0:timeseries_length, :]
    data[i, :, 14:26] = chroma.T[0:timeseries_length, :]
    data[i, :, 26:33] = spectral_contrast.T[0:timeseries_length, :]
```

Epoch 400/400

```
35/420 [=>.....] - ETA: 1s - loss: 0.3826 - acc: 0.8286
70/420 [====>.....] - ETA: 1s - loss: 0.3748 - acc: 0.8429
105/420 [=====>.....] - ETA: 1s - loss: 0.4167 - acc: 0.8571
140/420 [======>.....] - ETA: 1s - loss: 0.4069 - acc: 0.8714
175/420 [======>.....] - ETA: 0s - loss: 0.4087 - acc: 0.8743
210/420 [======>.....] - ETA: 0s - loss: 0.4276 - acc: 0.8619
245/420 [======>.....] - ETA: 0s - loss: 0.4582 - acc: 0.8490
280/420 [======>.....] - ETA: 0s - loss: 0.4495 - acc: 0.8500
315/420 [======>.....] - ETA: 0s - loss: 0.4462 - acc: 0.8444
350/420 [======>.....] - ETA: 0s - loss: 0.4395 - acc: 0.8400
385/420 [======>...] - ETA: 0s - loss: 0.4242 - acc: 0.8494
420/420 [=====] - 2s 4ms/step - loss: 0.4264 - acc: 0.8476
```

Validating ...

```
35/120 [=====>.....] - ETA: 0s
105/120 [======>.....] - ETA: 0s
120/120 [=====] - 0s 3ms/step
Dev loss: 0.9936995257933935
Dev accuracy: 0.6750000156462193
```

Testing ...

```
35/60 [======>.....] - ETA: 0s
60/60 [=====] - 0s 1ms/step
Test loss: 1.0625159045060475
Test accuracy: 0.6666666865348816
```

20.

```
for i, file in enumerate(list_of_audiofiles):
    y, sr = librosa.load(file)

    mfcc = librosa.feature.mfcc(y=y, sr=sr, hop_length=self.hop_length, n_mfcc=13)
    spectral_center = librosa.feature.spectral_bandwidth(y=y, hop_length=self.hop_length)
    chroma = librosa.feature.chroma_cens(y=y, sr=sr, hop_length=self.hop_length)
    rmse = librosa.feature.rmse(y=y, hop_length=self.hop_length)

    splits = re.split('[. ]', file)
    genre = re.split('[ /]', splits[1])[3]
    target.append(genre)

    data[i, :, 0:13] = mfcc.T[0:timeseries_length, :]
    data[i, :, 13:14] = spectral_center.T[0:timeseries_length, :]
    data[i, :, 14:26] = chroma.T[0:timeseries_length, :]
    data[i, :, 26:33] = rmse.T[0:timeseries_length, :]
```

Epoch 400/400

```
35/420 [=>.....] - ETA: 3s - loss: 0.5413 - acc: 0.8571
70/420 [====>.....] - ETA: 2s - loss: 0.5927 - acc: 0.8000
105/420 [=====>.....] - ETA: 2s - loss: 0.5818 - acc: 0.7714
140/420 [=====>.....] - ETA: 2s - loss: 0.5376 - acc: 0.7857
175/420 [=====>.....] - ETA: 2s - loss: 0.5079 - acc: 0.8057
210/420 [=====>.....] - ETA: 1s - loss: 0.4997 - acc: 0.8238
245/420 [=====>.....] - ETA: 1s - loss: 0.4711 - acc: 0.8327
280/420 [=====>.....] - ETA: 1s - loss: 0.4670 - acc: 0.8357
315/420 [=====>.....] - ETA: 0s - loss: 0.4763 - acc: 0.8317
350/420 [=====>.....] - ETA: 0s - loss: 0.4891 - acc: 0.8257
385/420 [=====>.....] - ETA: 0s - loss: 0.4933 - acc: 0.8234
420/420 [=====>.....] - 3s 8ms/step - loss: 0.5021 - acc: 0.8143
```

Validating ...

```
35/120 [=====>.....] - ETA: 1s
105/120 [=====>.....] - ETA: 0s
120/120 [=====>.....] - 1s 5ms/step
Dev loss: 1.2569006780783336
Dev accuracy: 0.6333333464960257
```

Testing ...

```
35/60 [=====>.....] - ETA: 0s
60/60 [=====>.....] - 0s 2ms/step
Test loss: 1.367306540409724
Test accuracy: 0.5166666780908903
```

21.

```
for i, file in enumerate(list_of_audiofiles):
    y, sr = librosa.load(file)

    mfcc = librosa.feature.mfcc(y=y, sr=sr, hop_length=self.hop_length, n_mfcc=13)
    spectral_center = librosa.feature.spectral_flatness(y=y, hop_length=self.hop_length)
    chroma = librosa.feature.chroma_cens(y=y, sr=sr, hop_length=self.hop_length)
    rmse = librosa.feature.rmse(y=y, hop_length=self.hop_length)

    splits = re.split('[ ]', file)
    genre = re.split('[ /]', splits[1])[3]
    target.append(genre)

    data[i, :, 0:13] = mfcc.T[0:timeseries_length, :]
    data[i, :, 13:14] = spectral_center.T[0:timeseries_length, :]
    data[i, :, 14:26] = chroma.T[0:timeseries_length, :]
    data[i, :, 26:33] = rmse.T[0:timeseries_length, :]
```

Epoch 400/400

```
35/420 [=>.....] - ETA: 3s - loss: 0.0185 - acc: 1.0000
70/420 [====>.....] - ETA: 2s - loss: 0.0108 - acc: 1.0000
105/420 [=====>.....] - ETA: 2s - loss: 0.0238 - acc: 0.9905
140/420 [=====>.....] - ETA: 2s - loss: 0.0211 - acc: 0.9929
175/420 [=====>.....] - ETA: 2s - loss: 0.0179 - acc: 0.9943
210/420 [=====>.....] - ETA: 1s - loss: 0.0206 - acc: 0.9952
245/420 [=====>.....] - ETA: 1s - loss: 0.0248 - acc: 0.9959
280/420 [=====>.....] - ETA: 1s - loss: 0.0229 - acc: 0.9964
315/420 [=====>.....] - ETA: 0s - loss: 0.0217 - acc: 0.9968
350/420 [=====>.....] - ETA: 0s - loss: 0.0215 - acc: 0.9971
385/420 [=====>.....] - ETA: 0s - loss: 0.0206 - acc: 0.9974
420/420 [=====>.....] - 3s 8ms/step - loss: 0.0214 - acc: 0.9976
```

Validating ...

```
35/120 [=====>.....] - ETA: 0s
105/120 [=====>.....] - ETA: 0s
120/120 [=====>.....] - 0s 4ms/step
Dev loss: 2.0063950767119727
Dev accuracy: 0.6250000223517418
```

Testing ...

```
35/60 [=====>.....] - ETA: 0s
60/60 [=====>.....] - 0s 1ms/step
Test loss: 2.245579570531845
Test accuracy: 0.48333333681027096
```


22.

```
for i, file in enumerate(list_of_audiofiles):
    y, sr = librosa.load(file)
    mfcc = librosa.feature.mfcc(y=y, sr=sr, hop_length=self.hop_length, n_mfcc=13)
    spectral_center = librosa.feature.spectral_contrast(y=y, hop_length=self.hop_length)
    chroma = librosa.feature.chroma_stft(y=y, sr=sr, hop_length=self.hop_length)
    rmse = librosa.feature.rmse(y=y, hop_length=self.hop_length)

    splits = re.split('[.]', file)
    genre = re.split('[/]', splits[1])[3]
    target.append(genre)

    data[i, :, 0:13] = mfcc.T[0:timeseries_length, :]
    data[i, :, 13:20] = spectral_center.T[0:timeseries_length, :]
    data[i, :, 20:32] = chroma.T[0:timeseries_length, :]
    data[i, :, 32:39] = rmse.T[0:timeseries_length, :]
```

Epoch 400/400

```
35/420 [=>.....] - ETA: 4s - loss: 0.0504 - acc: 0.9714
70/420 [====>.....] - ETA: 4s - loss: 0.0995 - acc: 0.9714
105/420 [=====>.....] - ETA: 3s - loss: 0.1221 - acc: 0.9524
140/420 [=====>.....] - ETA: 3s - loss: 0.0974 - acc: 0.9643
175/420 [=====>.....] - ETA: 3s - loss: 0.0836 - acc: 0.9714
210/420 [=====>.....] - ETA: 2s - loss: 0.0778 - acc: 0.9762
245/420 [=====>.....] - ETA: 2s - loss: 0.0775 - acc: 0.9714
280/420 [=====>.....] - ETA: 1s - loss: 0.0761 - acc: 0.9714
315/420 [=====>.....] - ETA: 1s - loss: 0.0838 - acc: 0.9651
350/420 [=====>.....] - ETA: 0s - loss: 0.0780 - acc: 0.9686
385/420 [=====>.....] - ETA: 0s - loss: 0.0770 - acc: 0.9662
420/420 [=====>.....] - 5s 12ms/step - loss: 0.0786 - acc: 0.9643
```

Validating ...

```
35/120 [=====>.....] - ETA: 1s
70/120 [=====>.....] - ETA: 0s
105/120 [=====>.....] - ETA: 0s
120/120 [=====>.....] - 1s 7ms/step
Dev loss: 1.9086458186308544
Dev accuracy: 0.6416666842997074
```

Testing ...

```
35/60 [=====>.....] - ETA: 0s
60/60 [=====>.....] - 0s 5ms/step
Test loss: 2.6331248680750527
Test accuracy: 0.5333333487311999
```

23.

```
for i, file in enumerate(list_of_audiofiles):
    y, sr = librosa.load(file)

    mfcc = librosa.feature.mfcc(y=y, sr=sr, hop_length=self.hop_length, n_mfcc=13)
    spectral_center = librosa.feature.spectral_contrast(y=y, hop_length=self.hop_length)
    chroma = librosa.feature.chroma_stft(y=y, sr=sr, hop_length=self.hop_length)
    zero_crossing_rate = librosa.feature.zero_crossing_rate(y=y, hop_length=self.hop_length)

    splits = re.split('[.]', file)
    genre = re.split('[/]', splits[1])[3]
    target.append(genre)

    data[i, :, 0:13] = mfcc.T[0:timeseries_length, :]
    data[i, :, 13:20] = spectral_center.T[0:timeseries_length, :]
    data[i, :, 20:32] = chroma.T[0:timeseries_length, :]
    data[i, :, 32:39] = zero_crossing_rate.T[0:timeseries_length, :]
```

Epoch 400/400

```
35/420 [=>.....] - ETA: 3s - loss: 0.0654 - acc: 0.9714
70/420 [====>.....] - ETA: 2s - loss: 0.0770 - acc: 0.9571
105/420 [=====>.....] - ETA: 2s - loss: 0.0543 - acc: 0.9714
140/420 [=====>.....] - ETA: 2s - loss: 0.0425 - acc: 0.9786
175/420 [=====>.....] - ETA: 1s - loss: 0.0347 - acc: 0.9829
210/420 [=====>.....] - ETA: 1s - loss: 0.0316 - acc: 0.9857
245/420 [=====>.....] - ETA: 1s - loss: 0.0424 - acc: 0.9837
280/420 [=====>.....] - ETA: 1s - loss: 0.0389 - acc: 0.9857
315/420 [=====>.....] - ETA: 0s - loss: 0.0355 - acc: 0.9873
350/420 [=====>.....] - ETA: 0s - loss: 0.0344 - acc: 0.9886
385/420 [=====>.....] - ETA: 0s - loss: 0.0430 - acc: 0.9870
420/420 [=====>.....] - 3s 8ms/step - loss: 0.0452 - acc: 0.9833
```

Validating ...

```
35/120 [=====>.....] - ETA: 0s
70/120 [=====>.....] - ETA: 0s
105/120 [=====>.....] - ETA: 0s
120/120 [=====>.....] - 1s 5ms/step
Dev loss: 2.1123732874790826
Dev accuracy: 0.59166669100523
```

Testing ...

```
35/60 [=====>.....] - ETA: 0s
60/60 [=====>.....] - 0s 2ms/step
Test loss: 2.146716525157293
Test accuracy: 0.5833333432674408
```

24.

```
for i, file in enumerate(list_of_audiofiles):
    y, sr = librosa.load(file)

    mfcc = librosa.feature.mfcc(y=y, sr=sr, hop_length=self.hop_length, n_mfcc=13)
    spectral_center = librosa.feature.spectral_contrast(y=y, hop_length=self.hop_length)
    chroma = librosa.feature.chroma_cqt(y=y, sr=sr, hop_length=self.hop_length)
    rmse = librosa.feature.rmse(y=y, hop_length=self.hop_length)

    splits = re.split('[.]', file)
    genre = re.split('[ /]', splits[1])[3]
    target.append(genre)

    data[i, :, 0:13] = mfcc.T[0:timeseries_length, :]
    data[i, :, 13:20] = spectral_center.T[0:timeseries_length, :]
    data[i, :, 20:32] = chroma.T[0:timeseries_length, :]
    data[i, :, 32:39] = rmse.T[0:timeseries_length, :]
```

Epoch 400/400

```
35/420 [=>.....] - ETA: 2s - loss: 0.0085 - acc: 1.0000
70/420 [====>.....] - ETA: 2s - loss: 0.0500 - acc: 0.9857
105/420 [=====>.....] - ETA: 2s - loss: 0.0586 - acc: 0.9810
140/420 [=====>.....] - ETA: 1s - loss: 0.0645 - acc: 0.9786
175/420 [=====>.....] - ETA: 1s - loss: 0.0722 - acc: 0.9714
210/420 [=====>.....] - ETA: 1s - loss: 0.0888 - acc: 0.9714
245/420 [=====>.....] - ETA: 1s - loss: 0.0807 - acc: 0.9755
280/420 [=====>.....] - ETA: 0s - loss: 0.0899 - acc: 0.9714
315/420 [=====>.....] - ETA: 0s - loss: 0.0869 - acc: 0.9714
350/420 [=====>.....] - ETA: 0s - loss: 0.0831 - acc: 0.9714
385/420 [=====>.....] - ETA: 0s - loss: 0.0774 - acc: 0.9740
420/420 [=====>.....] - 3s 7ms/step - loss: 0.0716 - acc: 0.9762
```

Validating ...

```
35/120 [=====>.....] - ETA: 0s
105/120 [=====>.....] - ETA: 0s
120/120 [=====>.....] - 1s 4ms/step
Dev loss: 2.672145207722982
Dev accuracy: 0.5416666852931181
```

Testing ...

```
35/60 [=====>.....] - ETA: 0s
60/60 [=====>.....] - 0s 2ms/step
Test loss: 2.8862125078837075
Test accuracy: 0.5000000024835268
```

25.

```
for i, file in enumerate(list_of_audiofiles):
    y, sr = librosa.load(file)
    mfcc = librosa.feature.mfcc(y=y, sr=sr, hop_length=self.hop_length, n_mfcc=13)
    spectral_center = librosa.feature.spectral_contrast(y=y, hop_length=self.hop_length)
    chroma = librosa.feature.chroma_cens(y=y, sr=sr, hop_length=self.hop_length)
    rmse = librosa.feature.rmse(y=y, hop_length=self.hop_length)

    splits = re.split(' [.]', file)
    genre = re.split(' [/]', splits[1])[3]
    target.append(genre)

    data[i, :, 0:13] = mfcc.T[0:timeseries_length, :]
    data[i, :, 13:20] = spectral_center.T[0:timeseries_length, :]
    data[i, :, 20:32] = chroma.T[0:timeseries_length, :]
    data[i, :, 32:39] = rmse.T[0:timeseries_length, :]
```

Epoch 400/400

```
35/420 [=>.....] - ETA: 4s - loss: 0.0085 - acc: 1.0000
70/420 [====>.....] - ETA: 4s - loss: 0.0518 - acc: 0.9857
105/420 [=====>.....] - ETA: 3s - loss: 0.0557 - acc: 0.9810
140/420 [=====>.....] - ETA: 3s - loss: 0.0604 - acc: 0.9786
175/420 [=====>.....] - ETA: 3s - loss: 0.0509 - acc: 0.9829
210/420 [=====>.....] - ETA: 2s - loss: 0.0549 - acc: 0.9810
245/420 [=====>.....] - ETA: 2s - loss: 0.0512 - acc: 0.9796
280/420 [=====>.....] - ETA: 1s - loss: 0.0522 - acc: 0.9786
315/420 [=====>.....] - ETA: 1s - loss: 0.0473 - acc: 0.9810
350/420 [=====>.....] - ETA: 0s - loss: 0.0455 - acc: 0.9800
385/420 [=====>.....] - ETA: 0s - loss: 0.0444 - acc: 0.9818
420/420 [=====>.....] - 5s 12ms/step - loss: 0.0445 - acc: 0.9810
```

Validating ...

```
35/120 [=====>.....] - ETA: 1s
70/120 [=====>.....] - ETA: 0s
105/120 [=====>.....] - ETA: 0s
120/120 [=====>.....] - 1s 7ms/step
```

Dev loss: 2.2819462964932122

Dev accuracy: 0.5583333531394601

Testing ...

```
35/60 [=====>.....] - ETA: 0s
60/60 [=====>.....] - 0s 5ms/step
```

Test loss: 2.19755748907725

Test accuracy: 0.5166666830579439

26.

```
for i, file in enumerate(list_of_audiofiles):
    y, sr = librosa.load(file)

    mfcc = librosa.feature.mfcc(y=y, sr=sr, hop_length=self.hop_length, n_mfcc=13)
    spectral_center = librosa.feature.spectral_centroid(y=y, hop_length=self.hop_length)
    chroma = librosa.feature.chroma_stft(y=y, sr=sr, hop_length=self.hop_length)
    rmse = librosa.feature.rmse(y=y, hop_length=self.hop_length)
    zero_crossing_rate = librosa.feature.zero_crossing_rate(y=y, hop_length=self.hop_length)
    spectral_bandwidth = librosa.feature.spectral_bandwidth(y=y, hop_length=self.hop_length)

    splits = re.split('[.]', file)
    genre = re.split('[ /]', splits[1])[3]
    target.append(genre)

    data[i, :, 0:13] = mfcc.T[0:timeseries_length, :]
    data[i, :, 13:20] = spectral_center.T[0:timeseries_length, :]
    data[i, :, 20:32] = chroma.T[0:timeseries_length, :]
    data[i, :, 32:39] = rmse.T[0:timeseries_length, :]
    data[i, :, 39:46] = zero_crossing_rate.T[0:timeseries_length, :]
    data[i, :, 46:53] = spectral_bandwidth.T[0:timeseries_length, :]
```

Epoch 400/400

```
35/420 [=>.....] - ETA: 2s - loss: 0.8533 - acc: 0.6571
70/420 [====>.....] - ETA: 2s - loss: 0.8750 - acc: 0.6286
105/420 [=====>.....] - ETA: 1s - loss: 0.8393 - acc: 0.6571
140/420 [=====>.....] - ETA: 1s - loss: 0.7916 - acc: 0.7071
175/420 [=====>.....] - ETA: 1s - loss: 0.8837 - acc: 0.6629
210/420 [=====>.....] - ETA: 1s - loss: 0.9138 - acc: 0.6571
245/420 [=====>.....] - ETA: 1s - loss: 0.9585 - acc: 0.6286
280/420 [=====>.....] - ETA: 0s - loss: 0.9686 - acc: 0.6179
315/420 [=====>.....] - ETA: 0s - loss: 0.9631 - acc: 0.6254
350/420 [=====>.....] - ETA: 0s - loss: 0.9564 - acc: 0.6286
385/420 [=====>.....] - ETA: 0s - loss: 0.9598 - acc: 0.6182
420/420 [=====>.....] - 2s 6ms/step - loss: 0.9559 - acc: 0.6167
```

Validating ...

```
35/120 [=====>.....] - ETA: 0s
105/120 [=====>.....] - ETA: 0s
120/120 [=====>.....] - 0s 4ms/step
Dev loss: 1.0772419820229213
Dev accuracy: 0.5916666785875956
```

Testing ...

```
35/60 [=====>.....] - ETA: 0s
60/60 [=====>.....] - 0s 1ms/step
Test loss: 1.0228710174560547
Test accuracy: 0.5166666780908903
```