

openmic

November 25, 2019

```
[1]: import librosa as lb
import librosa.display
import pandas as pd
import scipy
import json
import numpy as np
import sklearn
from sklearn.metrics import classification_report
from sklearn.model_selection import train_test_split
import os
from pylab import plot, show, figure, imshow, xlim, ylim, title
import matplotlib.pyplot as plt
import keras
from keras.utils import np_utils
from keras import layers
from keras import models
```

Using TensorFlow backend.

```
[2]: #CONSTANTS

DATA_DIR = "openmic-2018/"
CATEGORY_COUNT = 8

[3]: df = pd.read_csv('openmic-2018/openmic-2018-aggregated-labels.csv')
del df['relevance']
del df['num_responses']

[4]: labels = df.values
labels

[4]: array([[ '000046_3840', 'clarinet'],
            [ '000046_3840', 'flute'],
            [ '000046_3840', 'trumpet'],
            ...,
            [ '155311_453120', 'saxophone'],
            [ '155311_453120', 'trumpet'],
            [ '155311_453120', 'trombone']], dtype=object)
```

```
[5]: y, sr = lb.load(DATA_DIR + 'audio/000/000135_483840.ogg')
S = lb.feature.melspectrogram(y=y, sr=sr)

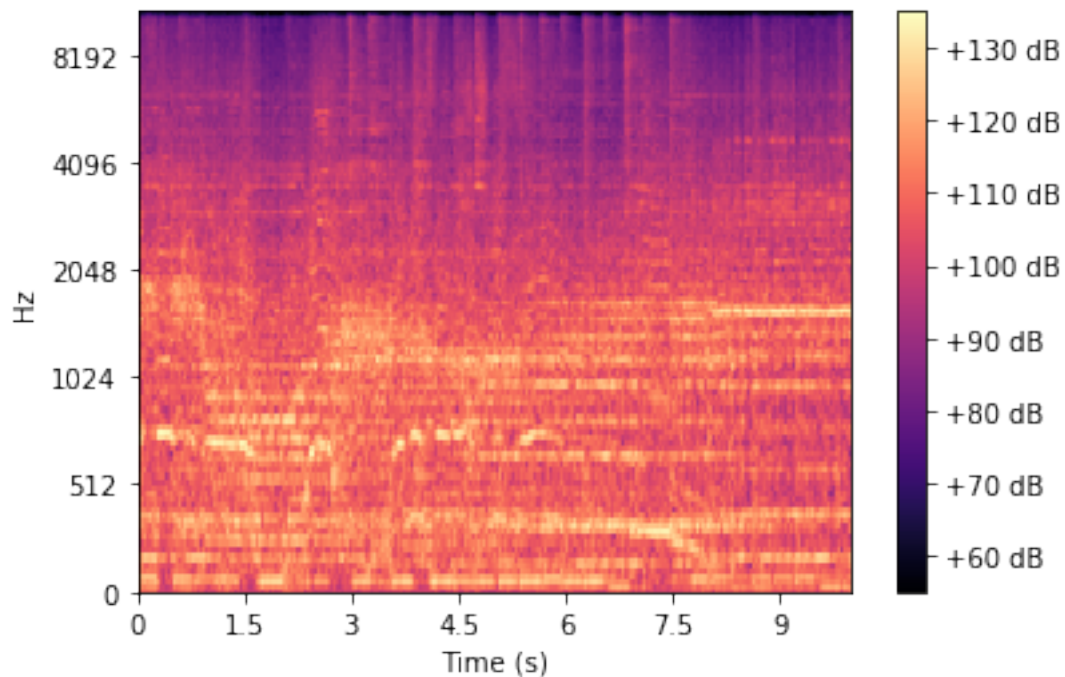
S_dB = lb.power_to_db(S, ref=0) # 10 * log10(S / ref)

print(y.shape)
print(sr)
print(S.shape)
print(S_dB.shape)
```

```
(220544,)
22050
(128, 431)
(128, 431)
```

```
[6]: librosa.display.specshow(S_dB, x_axis='s', y_axis='mel')
plt.colorbar(format='%+2.0f dB')
```

```
[6]: <matplotlib.colorbar.Colorbar at 0x27ec65b1390>
```



```
[7]: OPENMIC = np.load(os.path.join(DATA_DIR, 'openmic-2018.npz'), allow_pickle=True)
print(list(OPENMIC.keys()))
```

```
['X', 'Y_true', 'Y_mask', 'sample_key']
```

```
[8]: X, Y_true, Y_mask, sample_key = OPENMIC['X'], OPENMIC['Y_true'],  
      ↪OPENMIC['Y_mask'], OPENMIC['sample_key']  
      #print(X.shape)  
      #X = []  
      #print(len(sample_key))  
      #for key in sample_key:  
      #    key_dir = key[:3]  
      #    y, sr = lb.load(DATA_DIR + 'audio/' + key_dir + '/' + key + '.ogg')  
      #    X.append(lb.feature.melspectrogram(y=y, sr=sr))  
      #    print(len(X))
```

```
[9]: with open(os.path.join(DATA_DIR, 'class-map.json'), 'r') as f:  
      class_map = json.load(f)
```

```
[10]: split_train, split_test, X_train, X_test, Y_true_train, Y_true_test,  
      ↪Y_mask_train, Y_mask_test = train_test_split(sample_key, X, Y_true, Y_mask)  
      train_set = np.asarray(set(split_train))  
      test_set = np.asarray(set(split_test))  
      print('# Train: {}, # Test: {}'.format(len(split_train), len(split_test)))
```

```
# Train: 15000, # Test: 5000
```

```
[11]: print(X_train.shape)  
      print(X_test.shape)
```

```
(15000, 10, 128)
```

```
(5000, 10, 128)
```

```
[13]: # This dictionary will include the classifiers for each model  
      mymodels = dict()  
  
      # We'll iterate over all instrument classes, and fit a model for each one  
      # After training, we'll print a classification report for each instrument  
      for instrument in class_map:  
  
          # Map the instrument name to its column number  
          inst_num = class_map[instrument]  
  
          # Step 1: sub-sample the data  
  
          # First, we need to select down to the data for which we have annotations  
          # This is what the mask arrays are for  
          train_inst = Y_mask_train[:, inst_num]  
          test_inst = Y_mask_test[:, inst_num]  
  
          # Here, we're using the Y_mask_train array to slice out only the training,  
          ↪examples
```

```

# for which we have annotations for the given class
X_train_inst = X_train[train_inst]

# Step 3: simplify the data by averaging over time

# Let's arrange the data for a sklearn Random Forest model
# Instead of having time-varying features, we'll summarize each track by
→ its mean feature vector over time
X_train_inst_sklearn = np.mean(X_train_inst, axis=1)

# Again, we slice the labels to the annotated examples
# We threshold the label likelihoods at 0.5 to get binary labels
Y_true_train_inst = Y_true_train[train_inst, inst_num] >= 0.5

# Repeat the above slicing and dicing but for the test set
X_test_inst = X_test[test_inst]
X_test_inst_sklearn = np.mean(X_test_inst, axis=1)
Y_true_test_inst = Y_true_test[test_inst, inst_num] >= 0.5

X_train_inst = X_train_inst.astype('float32')
X_train_inst_sklearn = X_train_inst_sklearn.astype('float32')
X_train_inst_sklearn = lb.util.normalize(X_train_inst_sklearn)
# X_train_inst = S_dB
print(X_train_inst.shape)
shape = X_train_inst.shape
X_train_inst = X_train_inst.reshape(shape[0],1, shape[1], shape[2])
shape = X_test_inst.shape
X_test_inst = X_test_inst.reshape(shape[0],1, shape[1], shape[2])
#X_train_inst = X_train_inst.reshape(1,1,431,128)
print(X_train_inst.shape)
print(Y_true_train_inst[0])
# Step 3.
# Initialize a new classifier
import keras,os
from keras.models import Sequential
from keras.layers import Dense, Conv2D, MaxPool2D , Flatten
from keras.preprocessing.image import ImageDataGenerator
import numpy as np
model = models.Sequential()

# model.add(layers.Conv2D(filters=8,kernel_size=(3,3),activation='relu',
→ input_shape=(10,128,1,)))
model.
→ add(Conv2D(input_shape=(1,10,128),data_format="channels_first",filters=32,kernel_size=(3,3)
→ activation="relu",strides=(2,2)))

```

```

    model.add(Conv2D(filters=32, kernel_size=(3,3), padding="same",
→activation="relu", strides=(2,2)))
    model.add(MaxPool2D(pool_size=(3,3)))
    model.add(Conv2D(filters=128, kernel_size=(3,3), padding="same",
→activation="relu"))
    model.add(Conv2D(filters=128, kernel_size=(3,3), padding="same",
→activation="relu"))

    model.add(layers.Flatten())
    model.add(layers.Dense(units=512, activation='relu'))
    model.add(layers.Dense(units=256, activation='relu'))
    model.add(layers.Dense(units=100, activation='relu'))
    model.add(layers.Dense(units=1, activation='sigmoid'))

    model.compile(loss='binary_crossentropy',
                  optimizer='Adam',
                  metrics=['acc'])

# model.summary()
# Step 4.
model.fit(X_train_inst, Y_true_train_inst , epochs=10, batch_size=64)

# Step 5.
# Finally, we'll evaluate the model on both train and test
Y_pred_train = model.predict(X_train_inst)
Y_pred_test = model.predict(X_test_inst)
Y_pred_train_bool = Y_pred_train > 0.3 #THRESHOLD (should be lower than 0.
→5)
Y_pred_test_bool = Y_pred_test > 0.3 #THRESHOLD (should be lower than 0.5)
print(Y_pred_train[0])
print('-' * 52)
print(instrument)
print('\tTRAIN')
print(classification_report(Y_true_train_inst, Y_pred_train_bool))

print(Y_true_train_inst[3])
print(Y_pred_train[3])
print('\tTEST')
print(classification_report(Y_true_test_inst, Y_pred_test_bool))
sum = 0
# for i, prob in enumerate(Y_pred_train):
#     print (i)
#     print (prob)
#     sum += prob
# print(sum)
# Store the classifier in our dictionary
mymodels[instrument] = model

```

```

(1539, 10, 128)
(1539, 1, 10, 128)
False
Epoch 1/10
1539/1539 [=====] - 4s 2ms/step - loss: 0.8286 - acc:
0.7057
Epoch 2/10
1539/1539 [=====] - 2s 1ms/step - loss: 0.5063 - acc:
0.7661
Epoch 3/10
1539/1539 [=====] - 2s 1ms/step - loss: 0.4889 - acc:
0.7810
Epoch 4/10
1539/1539 [=====] - 2s 1ms/step - loss: 0.4506 - acc:
0.7888
Epoch 5/10
1539/1539 [=====] - 2s 1ms/step - loss: 0.3925 - acc:
0.8109
Epoch 6/10
1539/1539 [=====] - 2s 1ms/step - loss: 0.3575 - acc:
0.8395
Epoch 7/10
1539/1539 [=====] - 2s 1ms/step - loss: 0.3421 - acc:
0.8447
Epoch 8/10
1539/1539 [=====] - 2s 1ms/step - loss: 0.3203 - acc:
0.8545A: 1s - los
Epoch 9/10
1539/1539 [=====] - 2s 1ms/step - loss: 0.2954 - acc:
0.8824
Epoch 10/10
1539/1539 [=====] - 2s 1ms/step - loss: 0.2957 - acc:
0.8661
[0.02420208]
-----

```

accordion

TRAIN	precision	recall	f1-score	support
False	0.95	0.91	0.93	1169
True	0.74	0.84	0.79	370
accuracy			0.89	1539
macro avg	0.84	0.87	0.86	1539
weighted avg	0.90	0.89	0.89	1539

True

[0.6667123]

TEST	precision	recall	f1-score	support
False	0.89	0.85	0.87	413
True	0.55	0.63	0.59	119
accuracy			0.80	532
macro avg	0.72	0.74	0.73	532
weighted avg	0.81	0.80	0.81	532

(1679, 10, 128)

(1679, 1, 10, 128)

True

Epoch 1/10

1679/1679 [=====] - 4s 2ms/step - loss: 0.7042 - acc: 0.6587

Epoch 2/10

1679/1679 [=====] - 2s 1ms/step - loss: 0.4520 - acc: 0.7790

Epoch 3/10

1679/1679 [=====] - 2s 1ms/step - loss: 0.4088 - acc: 0.8064

Epoch 4/10

1679/1679 [=====] - 2s 1ms/step - loss: 0.3692 - acc: 0.8130

Epoch 5/10

1679/1679 [=====] - 3s 2ms/step - loss: 0.3128 - acc: 0.8612

Epoch 6/10

1679/1679 [=====] - 3s 2ms/step - loss: 0.2672 - acc: 0.8821

Epoch 7/10

1679/1679 [=====] - 2s 1ms/step - loss: 0.2523 - acc: 0.8970

Epoch 8/10

1679/1679 [=====] - 2s 1ms/step - loss: 0.1602 - acc: 0.9369

Epoch 9/10

1679/1679 [=====] - 2s 1ms/step - loss: 0.1417 - acc: 0.9410

Epoch 10/10

1679/1679 [=====] - 2s 1ms/step - loss: 0.0892 - acc: 0.9690

[0.99322283]

banjo

TRAIN

precision	recall	f1-score	support
-----------	--------	----------	---------

False	0.97	0.99	0.98	1120
True	0.98	0.95	0.96	559
accuracy			0.98	1679
macro avg	0.98	0.97	0.97	1679
weighted avg	0.98	0.98	0.98	1679

True
[0.9989266]

TEST	precision	recall	f1-score	support
False	0.81	0.88	0.84	366
True	0.69	0.57	0.62	173
accuracy			0.78	539
macro avg	0.75	0.72	0.73	539
weighted avg	0.77	0.78	0.77	539

(1416, 10, 128)

(1416, 1, 10, 128)

False

Epoch 1/10

1416/1416 [=====] - 4s 2ms/step - loss: 0.8830 - acc: 0.6780

Epoch 2/10

1416/1416 [=====] - 2s 1ms/step - loss: 0.5550 - acc: 0.7295

Epoch 3/10

1416/1416 [=====] - 2s 1ms/step - loss: 0.4950 - acc: 0.7613

Epoch 4/10

1416/1416 [=====] - 2s 1ms/step - loss: 0.4179 - acc: 0.8072

Epoch 5/10

1416/1416 [=====] - 2s 1ms/step - loss: 0.3977 - acc: 0.8185

Epoch 6/10

1416/1416 [=====] - 2s 1ms/step - loss: 0.3409 - acc: 0.8482

Epoch 7/10

1416/1416 [=====] - 2s 1ms/step - loss: 0.2942 - acc: 0.8764

Epoch 8/10

1416/1416 [=====] - 2s 1ms/step - loss: 0.2433 - acc: 0.9054

Epoch 9/10

1416/1416 [=====] - 2s 2ms/step - loss: 0.2353 - acc:
0.9054A: 0s - loss: 0.2331 - acc: 0.9

Epoch 10/10

1416/1416 [=====] - 2s 2ms/step - loss: 0.1678 - acc:
0.9301

[0.95407677]

bass

TRAIN				
	precision	recall	f1-score	support
False	1.00	0.70	0.82	1007
True	0.58	1.00	0.73	409
accuracy			0.79	1416
macro avg	0.79	0.85	0.78	1416
weighted avg	0.88	0.79	0.80	1416

False

[0.0468536]

TEST				
	precision	recall	f1-score	support
False	0.89	0.56	0.69	332
True	0.44	0.83	0.58	140
accuracy			0.64	472
macro avg	0.67	0.70	0.63	472
weighted avg	0.76	0.64	0.66	472

(1441, 10, 128)

(1441, 1, 10, 128)

False

Epoch 1/10

1441/1441 [=====] - 4s 3ms/step - loss: 0.9748 - acc:
0.5718

Epoch 2/10

1441/1441 [=====] - 2s 1ms/step - loss: 0.5209 - acc:
0.7245

Epoch 3/10

1441/1441 [=====] - 2s 1ms/step - loss: 0.4502 - acc:
0.7668

Epoch 4/10

1441/1441 [=====] - 2s 1ms/step - loss: 0.3901 - acc:
0.8071

Epoch 5/10

1441/1441 [=====] - 2s 1ms/step - loss: 0.3758 - acc:
0.8189

Epoch 6/10
 1441/1441 [=====] - 2s 1ms/step - loss: 0.3448 - acc: 0.8411
 Epoch 7/10
 1441/1441 [=====] - 2s 1ms/step - loss: 0.3110 - acc: 0.8605
 Epoch 8/10
 1441/1441 [=====] - 2s 1ms/step - loss: 0.3104 - acc: 0.8577A: 1s - loss: 0.3454 - ETA: 0s - loss: 0.3014 - acc: 0.8
 Epoch 9/10
 1441/1441 [=====] - 2s 1ms/step - loss: 0.3513 - acc: 0.8223
 Epoch 10/10
 1441/1441 [=====] - 2s 1ms/step - loss: 0.2368 - acc: 0.9008
 [0.00059983]

 cello

	precision	recall	f1-score	support
False	0.99	0.87	0.93	829
True	0.85	0.98	0.91	612
accuracy			0.92	1441
macro avg	0.92	0.93	0.92	1441
weighted avg	0.93	0.92	0.92	1441

True
 [0.92815673]

	precision	recall	f1-score	support
False	0.83	0.72	0.77	296
True	0.67	0.80	0.73	212
accuracy			0.75	508
macro avg	0.75	0.76	0.75	508
weighted avg	0.77	0.75	0.75	508

(1759, 10, 128)
 (1759, 1, 10, 128)
 False

Epoch 1/10
 1759/1759 [=====] - 4s 2ms/step - loss: 0.6442 - acc: 0.7504
 Epoch 2/10
 1759/1759 [=====] - 2s 1ms/step - loss: 0.4823 - acc:

```

0.7658
Epoch 3/10
1759/1759 [=====] - 3s 1ms/step - loss: 0.4430 - acc:
0.7834
Epoch 4/10
1759/1759 [=====] - 2s 1ms/step - loss: 0.4259 - acc:
0.7874
Epoch 5/10
1759/1759 [=====] - 2s 1ms/step - loss: 0.4028 - acc:
0.7897
Epoch 6/10
1759/1759 [=====] - 2s 1ms/step - loss: 0.3288 - acc:
0.8368A: 0s - loss: 0.3405 - acc: 0.830 - ETA: 0s - loss: 0.3370 - acc
Epoch 7/10
1759/1759 [=====] - 2s 1ms/step - loss: 0.3163 - acc:
0.8499
Epoch 8/10
1759/1759 [=====] - 2s 1ms/step - loss: 0.2021 - acc:
0.9187
Epoch 9/10
1759/1759 [=====] - 2s 1ms/step - loss: 0.1539 - acc:
0.9403A: 2s - los
Epoch 10/10
1759/1759 [=====] - 2s 1ms/step - loss: 0.1341 - acc:
0.9551A: 0s - loss: 0.1261 - acc: 0
[0.05214179]

```

clarinet

TRAIN					
	precision	recall	f1-score	support	
False	0.96	0.99	0.97	1354	
True	0.95	0.87	0.91	405	
accuracy			0.96	1759	
macro avg	0.96	0.93	0.94	1759	
weighted avg	0.96	0.96	0.96	1759	

False

[0.00090608]

TEST					
	precision	recall	f1-score	support	
False	0.82	0.88	0.85	498	
True	0.36	0.27	0.31	128	
accuracy			0.75	626	
macro avg	0.59	0.58	0.58	626	

```

weighted avg      0.73      0.75      0.74      626

(1343, 10, 128)
(1343, 1, 10, 128)
False
Epoch 1/10
1343/1343 [=====] - 4s 3ms/step - loss: 1.0250 - acc:
0.6687
Epoch 2/10
1343/1343 [=====] - 2s 1ms/step - loss: 0.3174 - acc:
0.8801
Epoch 3/10
1343/1343 [=====] - 2s 1ms/step - loss: 0.2486 - acc:
0.9144
Epoch 4/10
1343/1343 [=====] - 2s 1ms/step - loss: 0.2292 - acc:
0.9188
Epoch 5/10
1343/1343 [=====] - 2s 1ms/step - loss: 0.1975 - acc:
0.9322
Epoch 6/10
1343/1343 [=====] - 2s 1ms/step - loss: 0.2003 - acc:
0.9367
Epoch 7/10
1343/1343 [=====] - 2s 1ms/step - loss: 0.1854 - acc:
0.9367
Epoch 8/10
1343/1343 [=====] - 2s 1ms/step - loss: 0.1865 - acc:
0.9345
Epoch 9/10
1343/1343 [=====] - 2s 1ms/step - loss: 0.1439 - acc:
0.9516
Epoch 10/10
1343/1343 [=====] - 2s 1ms/step - loss: 0.1251 - acc:
0.9598
[5.9604645e-08]

```

cymbals

	TRAIN	precision	recall	f1-score	support
	False	1.00	0.93	0.97	473
	True	0.97	1.00	0.98	870
	accuracy			0.98	1343
	macro avg	0.98	0.97	0.97	1343
	weighted avg	0.98	0.98	0.98	1343

```

True
[0.98833585]
      TEST
      precision    recall  f1-score   support

 False         0.91      0.85      0.88        151
  True         0.91      0.95      0.93        241

 accuracy              0.91        392
 macro avg         0.91      0.90      0.90        392
weighted avg         0.91      0.91      0.91        392

(1321, 10, 128)
(1321, 1, 10, 128)
False
Epoch 1/10
1321/1321 [=====] - 4s 3ms/step - loss: 0.5865 - acc:
0.7434
Epoch 2/10
1321/1321 [=====] - 2s 2ms/step - loss: 0.2039 - acc:
0.9281
Epoch 3/10
1321/1321 [=====] - 2s 1ms/step - loss: 0.1801 - acc:
0.9296
Epoch 4/10
1321/1321 [=====] - 2s 1ms/step - loss: 0.1412 - acc:
0.9553
Epoch 5/10
1321/1321 [=====] - 1s 1ms/step - loss: 0.0959 - acc:
0.9682
Epoch 6/10
1321/1321 [=====] - 1s 1ms/step - loss: 0.0531 - acc:
0.9833
Epoch 7/10
1321/1321 [=====] - 1s 1ms/step - loss: 0.0291 - acc:
0.9917
Epoch 8/10
1321/1321 [=====] - 2s 1ms/step - loss: 0.1367 - acc:
0.9508
Epoch 9/10
1321/1321 [=====] - 2s 1ms/step - loss: 0.0681 - acc:
0.9750
Epoch 10/10
1321/1321 [=====] - 2s 1ms/step - loss: 0.0236 - acc:
0.9924
[0.]
-----
drums

```

TRAIN				
	precision	recall	f1-score	support
False	1.00	1.00	1.00	479
True	1.00	1.00	1.00	842
accuracy			1.00	1321
macro avg	1.00	1.00	1.00	1321
weighted avg	1.00	1.00	1.00	1321

True
[0.99965954]

TEST				
	precision	recall	f1-score	support
False	0.91	0.86	0.88	162
True	0.92	0.95	0.93	264
accuracy			0.91	426
macro avg	0.91	0.90	0.91	426
weighted avg	0.91	0.91	0.91	426

(1570, 10, 128)

(1570, 1, 10, 128)

False

Epoch 1/10

1570/1570 [=====] - 4s 2ms/step - loss: 0.9802 - acc: 0.6242

Epoch 2/10

1570/1570 [=====] - 2s 1ms/step - loss: 0.5765 - acc: 0.7057

Epoch 3/10

1570/1570 [=====] - 2s 1ms/step - loss: 0.5343 - acc: 0.7102

Epoch 4/10

1570/1570 [=====] - 2s 1ms/step - loss: 0.4799 - acc: 0.7580

Epoch 5/10

1570/1570 [=====] - 2s 1ms/step - loss: 0.4580 - acc: 0.7732

Epoch 6/10

1570/1570 [=====] - 2s 1ms/step - loss: 0.4390 - acc: 0.7885

Epoch 7/10

1570/1570 [=====] - 2s 1ms/step - loss: 0.4289 - acc: 0.8019

Epoch 8/10

1570/1570 [=====] - 2s 1ms/step - loss: 0.3413 - acc:

```

0.8420
Epoch 9/10
1570/1570 [=====] - 2s 1ms/step - loss: 0.2997 - acc:
0.8732
Epoch 10/10
1570/1570 [=====] - 2s 1ms/step - loss: 0.2171 - acc:
0.9000
[0.29924047]

```

flute

TRAIN					
	precision	recall	f1-score	support	
False	1.00	0.83	0.91	1076	
True	0.73	1.00	0.85	494	
accuracy			0.89	1570	
macro avg	0.87	0.92	0.88	1570	
weighted avg	0.92	0.89	0.89	1570	

False

[0.00135657]

TEST					
	precision	recall	f1-score	support	
False	0.84	0.59	0.69	361	
True	0.43	0.75	0.55	153	
accuracy			0.63	514	
macro avg	0.64	0.67	0.62	514	
weighted avg	0.72	0.63	0.65	514	

(1230, 10, 128)

(1230, 1, 10, 128)

False

Epoch 1/10

```

1230/1230 [=====] - 4s 3ms/step - loss: 0.8445 - acc:
0.7455

```

Epoch 2/10

```

1230/1230 [=====] - 1s 1ms/step - loss: 0.2476 - acc:
0.9122

```

Epoch 3/10

```

1230/1230 [=====] - 1s 1ms/step - loss: 0.1325 - acc:
0.9634

```

Epoch 4/10

```

1230/1230 [=====] - 1s 1ms/step - loss: 0.1119 - acc:
0.9659

```

Epoch 5/10

```

1230/1230 [=====] - 2s 2ms/step - loss: 0.1309 - acc:
0.9488
Epoch 6/10
1230/1230 [=====] - 3s 2ms/step - loss: 0.1290 - acc:
0.9602
Epoch 7/10
1230/1230 [=====] - 3s 2ms/step - loss: 0.0894 - acc:
0.9740
Epoch 8/10
1230/1230 [=====] - 2s 2ms/step - loss: 0.0664 - acc:
0.9846
Epoch 9/10
1230/1230 [=====] - 3s 2ms/step - loss: 0.0363 - acc:
0.9894
Epoch 10/10
1230/1230 [=====] - 2s 2ms/step - loss: 0.0260 - acc:
0.9943
[3.874302e-07]

```

guitar

	TRAIN				
		precision	recall	f1-score	support
	False	1.00	0.97	0.99	378
	True	0.99	1.00	0.99	852
	accuracy			0.99	1230
	macro avg	0.99	0.99	0.99	1230
	weighted avg	0.99	0.99	0.99	1230

True

[0.999998]

	TEST				
		precision	recall	f1-score	support
	False	0.93	0.91	0.92	134
	True	0.96	0.97	0.96	286
	accuracy			0.95	420
	macro avg	0.94	0.94	0.94	420
	weighted avg	0.95	0.95	0.95	420

(1345, 10, 128)

(1345, 1, 10, 128)

True

Epoch 1/10

```

1345/1345 [=====] - 5s 4ms/step - loss: 0.8501 - acc:
0.6104A: 11s - loss: 1.

```


Epoch 2/10
1345/1345 [=====] - 2s 1ms/step - loss: 0.4930 - acc:
0.7279
Epoch 3/10
1345/1345 [=====] - 3s 2ms/step - loss: 0.4327 - acc:
0.7948
Epoch 4/10
1345/1345 [=====] - 2s 2ms/step - loss: 0.3678 - acc:
0.8320
Epoch 5/10
1345/1345 [=====] - 3s 2ms/step - loss: 0.3244 - acc:
0.8506
Epoch 6/10
1345/1345 [=====] - 2s 1ms/step - loss: 0.3129 - acc:
0.8572A: 1s - loss: 0.
Epoch 7/10
1345/1345 [=====] - 2s 1ms/step - loss: 0.3694 - acc:
0.8409
Epoch 8/10
1345/1345 [=====] - 2s 2ms/step - loss: 0.2657 - acc:
0.8914
Epoch 9/10
1345/1345 [=====] - 2s 2ms/step - loss: 0.1968 - acc:
0.9197
Epoch 10/10
1345/1345 [=====] - 2s 1ms/step - loss: 0.1436 - acc:
0.9532
[0.93632215]

mallet_percussion

TRAIN					
	precision	recall	f1-score	support	
False	0.71	1.00	0.83	800	
True	0.99	0.41	0.58	545	
accuracy			0.76	1345	
macro avg	0.85	0.70	0.71	1345	
weighted avg	0.82	0.76	0.73	1345	

True

[0.98925805]

TEST					
	precision	recall	f1-score	support	
False	0.68	0.97	0.80	269	
True	0.88	0.35	0.50	188	

accuracy			0.71	457
macro avg	0.78	0.66	0.65	457
weighted avg	0.76	0.71	0.67	457

(1864, 10, 128)

(1864, 1, 10, 128)

False

Epoch 1/10

1864/1864 [=====] - 5s 3ms/step - loss: 0.8315 - acc: 0.6615

Epoch 2/10

1864/1864 [=====] - 3s 1ms/step - loss: 0.4982 - acc: 0.7430

Epoch 3/10

1864/1864 [=====] - 2s 1ms/step - loss: 0.4716 - acc: 0.7559

Epoch 4/10

1864/1864 [=====] - 3s 1ms/step - loss: 0.4560 - acc: 0.7677

Epoch 5/10

1864/1864 [=====] - 3s 2ms/step - loss: 0.4211 - acc: 0.7822

Epoch 6/10

1864/1864 [=====] - 3s 1ms/step - loss: 0.4010 - acc: 0.8053

Epoch 7/10

1864/1864 [=====] - 3s 1ms/step - loss: 0.3639 - acc: 0.8289A: 2s -

Epoch 8/10

1864/1864 [=====] - 3s 2ms/step - loss: 0.3250 - acc: 0.8466

Epoch 9/10

1864/1864 [=====] - 3s 1ms/step - loss: 0.3472 - acc: 0.8401

Epoch 10/10

1864/1864 [=====] - 3s 1ms/step - loss: 0.2752 - acc: 0.8739A: 1s - loss: 0.2931 -

[6.7949295e-06]

mandolin

TRAIN

	precision	recall	f1-score	support
False	1.00	0.77	0.87	1223
True	0.70	1.00	0.82	641
accuracy			0.85	1864
macro avg	0.85	0.89	0.85	1864

weighted avg	0.89	0.85	0.85	1864
--------------	------	------	------	------

True

[0.97260296]

TEST

	precision	recall	f1-score	support
False	0.89	0.58	0.70	396
True	0.51	0.86	0.64	204
accuracy			0.67	600
macro avg	0.70	0.72	0.67	600
weighted avg	0.76	0.67	0.68	600

(1449, 10, 128)

(1449, 1, 10, 128)

True

Epoch 1/10

1449/1449 [=====] - 4s 3ms/step - loss: 1.0355 - acc: 0.6632

Epoch 2/10

1449/1449 [=====] - 2s 1ms/step - loss: 0.4317 - acc: 0.8012

Epoch 3/10

1449/1449 [=====] - 2s 1ms/step - loss: 0.3757 - acc: 0.8171

Epoch 4/10

1449/1449 [=====] - 2s 1ms/step - loss: 0.3604 - acc: 0.8233

Epoch 5/10

1449/1449 [=====] - 2s 1ms/step - loss: 0.3185 - acc: 0.8420

Epoch 6/10

1449/1449 [=====] - 2s 1ms/step - loss: 0.3039 - acc: 0.8496

Epoch 7/10

1449/1449 [=====] - 2s 1ms/step - loss: 0.3003 - acc: 0.8606

Epoch 8/10

1449/1449 [=====] - 2s 1ms/step - loss: 0.2076 - acc: 0.9068

Epoch 9/10

1449/1449 [=====] - 2s 1ms/step - loss: 0.1692 - acc: 0.9303

Epoch 10/10

1449/1449 [=====] - 2s 1ms/step - loss: 0.1405 - acc: 0.9413

[0.9638747]

organ

TRAIN					
	precision	recall	f1-score	support	
False	1.00	0.94	0.97	983	
True	0.89	1.00	0.94	466	
accuracy			0.96	1449	
macro avg	0.95	0.97	0.96	1449	
weighted avg	0.96	0.96	0.96	1449	

False

[5.4091215e-05]

TEST					
	precision	recall	f1-score	support	
False	0.90	0.75	0.82	304	
True	0.60	0.81	0.69	137	
accuracy			0.77	441	
macro avg	0.75	0.78	0.75	441	
weighted avg	0.80	0.77	0.78	441	

(1291, 10, 128)

(1291, 1, 10, 128)

False

Epoch 1/10

1291/1291 [=====] - 5s 4ms/step - loss: 0.6743 - acc: 0.7266

Epoch 2/10

1291/1291 [=====] - 2s 1ms/step - loss: 0.1667 - acc: 0.9388

Epoch 3/10

1291/1291 [=====] - 2s 1ms/step - loss: 0.1288 - acc: 0.9558

Epoch 4/10

1291/1291 [=====] - 2s 1ms/step - loss: 0.1218 - acc: 0.9613

Epoch 5/10

1291/1291 [=====] - 2s 1ms/step - loss: 0.0967 - acc: 0.9675

Epoch 6/10

1291/1291 [=====] - 2s 1ms/step - loss: 0.1056 - acc: 0.9659

Epoch 7/10

1291/1291 [=====] - 2s 1ms/step - loss: 0.0887 - acc: 0.9752

Epoch 8/10
 1291/1291 [=====] - 2s 1ms/step - loss: 0.0523 - acc:
 0.9830
 Epoch 9/10
 1291/1291 [=====] - 2s 1ms/step - loss: 0.0382 - acc:
 0.9884
 Epoch 10/10
 1291/1291 [=====] - 2s 1ms/step - loss: 0.0523 - acc:
 0.9845
 [0.1406048]

 piano

TRAIN					
	precision	recall	f1-score	support	
False	1.00	0.89	0.94	414	
True	0.95	1.00	0.97	877	
accuracy			0.97	1291	
macro avg	0.98	0.95	0.96	1291	
weighted avg	0.97	0.97	0.96	1291	

False
 [0.00920945]

TEST					
	precision	recall	f1-score	support	
False	0.98	0.82	0.90	136	
True	0.92	0.99	0.96	293	
accuracy			0.94	429	
macro avg	0.95	0.91	0.93	429	
weighted avg	0.94	0.94	0.94	429	

(1733, 10, 128)
 (1733, 1, 10, 128)

True

Epoch 1/10
 1733/1733 [=====] - 8s 5ms/step - loss: 0.9876 - acc:
 0.5026A: 3s - loss: 1.1127 - ac
 Epoch 2/10
 1733/1733 [=====] - 2s 1ms/step - loss: 0.6010 - acc:
 0.6705
 Epoch 3/10
 1733/1733 [=====] - 3s 1ms/step - loss: 0.5303 - acc:
 0.7224
 Epoch 4/10
 1733/1733 [=====] - 3s 2ms/step - loss: 0.4989 - acc:

```

0.7340
Epoch 5/10
1733/1733 [=====] - 3s 2ms/step - loss: 0.4368 - acc:
0.7825
Epoch 6/10
1733/1733 [=====] - 2s 1ms/step - loss: 0.4158 - acc:
0.8096
Epoch 7/10
1733/1733 [=====] - 2s 1ms/step - loss: 0.3781 - acc:
0.8171
Epoch 8/10
1733/1733 [=====] - 2s 1ms/step - loss: 0.3367 - acc:
0.8488
Epoch 9/10
1733/1733 [=====] - 2s 1ms/step - loss: 0.4156 - acc:
0.8130
Epoch 10/10
1733/1733 [=====] - 2s 1ms/step - loss: 0.2951 - acc:
0.8731
[0.92669]

```

saxophone

TRAIN	precision	recall	f1-score	support
False	0.98	0.77	0.87	900
True	0.80	0.98	0.88	833
accuracy			0.87	1733
macro avg	0.89	0.88	0.87	1733
weighted avg	0.89	0.87	0.87	1733

True

[0.8732294]

TEST	precision	recall	f1-score	support
False	0.85	0.58	0.69	330
True	0.66	0.89	0.76	302
accuracy			0.73	632
macro avg	0.75	0.73	0.72	632
weighted avg	0.76	0.73	0.72	632

(1185, 10, 128)

(1185, 1, 10, 128)

False

Epoch 1/10

```

1185/1185 [=====] - 4s 4ms/step - loss: 0.6299 - acc:
0.7840
Epoch 2/10
1185/1185 [=====] - 2s 1ms/step - loss: 0.2077 - acc:
0.9367
Epoch 3/10
1185/1185 [=====] - 1s 1ms/step - loss: 0.1712 - acc:
0.9426
Epoch 4/10
1185/1185 [=====] - 1s 1ms/step - loss: 0.1402 - acc:
0.9603
Epoch 5/10
1185/1185 [=====] - 1s 1ms/step - loss: 0.1054 - acc:
0.9705
Epoch 6/10
1185/1185 [=====] - 2s 2ms/step - loss: 0.1158 - acc:
0.9654
Epoch 7/10
1185/1185 [=====] - 2s 2ms/step - loss: 0.0950 - acc:
0.9738
Epoch 8/10
1185/1185 [=====] - 3s 2ms/step - loss: 0.0810 - acc:
0.9806
Epoch 9/10
1185/1185 [=====] - 2s 1ms/step - loss: 0.0653 - acc:
0.9865
Epoch 10/10
1185/1185 [=====] - 1s 1ms/step - loss: 0.0575 - acc:
0.9831
[7.897615e-06]

```

synthesizer

TRAIN				
	precision	recall	f1-score	support
False	1.00	0.97	0.99	376
True	0.99	1.00	0.99	809
accuracy			0.99	1185
macro avg	0.99	0.99	0.99	1185
weighted avg	0.99	0.99	0.99	1185

True

[0.99991584]

TEST				
	precision	recall	f1-score	support
False	0.96	0.90	0.93	135

True	0.95	0.98	0.97	282
accuracy			0.95	417
macro avg	0.96	0.94	0.95	417
weighted avg	0.95	0.95	0.95	417

(2036, 10, 128)

(2036, 1, 10, 128)

True

Epoch 1/10

2036/2036 [=====] - 5s 3ms/step - loss: 0.7733 - acc: 0.6660

Epoch 2/10

2036/2036 [=====] - 2s 1ms/step - loss: 0.4876 - acc: 0.7677

Epoch 3/10

2036/2036 [=====] - 3s 1ms/step - loss: 0.4180 - acc: 0.8158

Epoch 4/10

2036/2036 [=====] - 2s 1ms/step - loss: 0.3678 - acc: 0.8418

Epoch 5/10

2036/2036 [=====] - 2s 1ms/step - loss: 0.3313 - acc: 0.8561

Epoch 6/10

2036/2036 [=====] - 2s 1ms/step - loss: 0.2626 - acc: 0.8924

Epoch 7/10

2036/2036 [=====] - 2s 1ms/step - loss: 0.2182 - acc: 0.9190

Epoch 8/10

2036/2036 [=====] - 3s 1ms/step - loss: 0.1714 - acc: 0.9322

Epoch 9/10

2036/2036 [=====] - 3s 1ms/step - loss: 0.1119 - acc: 0.9578

Epoch 10/10

2036/2036 [=====] - 3s 1ms/step - loss: 0.0983 - acc: 0.9661A: 2s - 1o

[0.989437]

trombone

TRAIN

	precision	recall	f1-score	support
False	0.99	0.99	0.99	1409
True	0.99	0.98	0.98	627

accuracy			0.99	2036
macro avg	0.99	0.98	0.99	2036
weighted avg	0.99	0.99	0.99	2036

False

[0.00032526]

TEST

	precision	recall	f1-score	support
False	0.83	0.86	0.85	488
True	0.69	0.63	0.66	236

accuracy			0.79	724
macro avg	0.76	0.75	0.75	724
weighted avg	0.78	0.79	0.78	724

(2183, 10, 128)

(2183, 1, 10, 128)

True

Epoch 1/10

2183/2183 [=====] - 6s 3ms/step - loss: 0.8680 - acc: 0.5767

Epoch 2/10

2183/2183 [=====] - 3s 1ms/step - loss: 0.5370 - acc: 0.7169A: 1s - loss: 0.562

Epoch 3/10

2183/2183 [=====] - 3s 1ms/step - loss: 0.4889 - acc: 0.7572

Epoch 4/10

2183/2183 [=====] - 2s 1ms/step - loss: 0.4261 - acc: 0.7975

Epoch 5/10

2183/2183 [=====] - 2s 1ms/step - loss: 0.4126 - acc: 0.8090

Epoch 6/10

2183/2183 [=====] - 2s 1ms/step - loss: 0.3808 - acc: 0.8296

Epoch 7/10

2183/2183 [=====] - 3s 1ms/step - loss: 0.3219 - acc: 0.8562

Epoch 8/10

2183/2183 [=====] - 3s 1ms/step - loss: 0.2890 - acc: 0.8754

Epoch 9/10

2183/2183 [=====] - 3s 1ms/step - loss: 0.2240 - acc: 0.9020A: 1s - loss: 0.

Epoch 10/10

2183/2183 [=====] - 2s 1ms/step - loss: 0.2024 - acc:

0.9180
[0.86600804]

trumpet

TRAIN	precision	recall	f1-score	support
False	1.00	0.82	0.90	1335
True	0.78	1.00	0.88	848
accuracy			0.89	2183
macro avg	0.89	0.91	0.89	2183
weighted avg	0.91	0.89	0.89	2183

True
[0.8923961]

TEST	precision	recall	f1-score	support
False	0.87	0.61	0.72	435
True	0.60	0.87	0.71	298
accuracy			0.71	733
macro avg	0.74	0.74	0.71	733
weighted avg	0.76	0.71	0.71	733

(1812, 10, 128)
(1812, 1, 10, 128)

True

Epoch 1/10

1812/1812 [=====] - 5s 3ms/step - loss: 0.6498 - acc:
0.7053

Epoch 2/10

1812/1812 [=====] - 2s 1ms/step - loss: 0.4708 - acc:
0.7472

Epoch 3/10

1812/1812 [=====] - 2s 1ms/step - loss: 0.4387 - acc:
0.7710

Epoch 4/10

1812/1812 [=====] - 3s 2ms/step - loss: 0.3833 - acc:
0.8102

Epoch 5/10

1812/1812 [=====] - 3s 1ms/step - loss: 0.3722 - acc:
0.8201

Epoch 6/10

1812/1812 [=====] - 2s 1ms/step - loss: 0.3077 - acc:
0.8571

Epoch 7/10

```

1812/1812 [=====] - 2s 1ms/step - loss: 0.2764 - acc:
0.8775
Epoch 8/10
1812/1812 [=====] - 3s 1ms/step - loss: 0.2015 - acc:
0.9178
Epoch 9/10
1812/1812 [=====] - 3s 1ms/step - loss: 0.1961 - acc:
0.9150A: 1s - loss: 0.1089 - acc: - ETA: 0s - loss: 0.1466 - a
Epoch 10/10
1812/1812 [=====] - 2s 1ms/step - loss: 0.1642 - acc:
0.9338
[0.8599187]

```

ukulele

TRAIN					
	precision	recall	f1-score	support	
False	0.99	0.98	0.98	1271	
True	0.94	0.98	0.96	541	
accuracy			0.98	1812	
macro avg	0.97	0.98	0.97	1812	
weighted avg	0.98	0.98	0.98	1812	

True

[0.974316]

TEST					
	precision	recall	f1-score	support	
False	0.80	0.81	0.80	416	
True	0.59	0.56	0.58	197	
accuracy			0.73	613	
macro avg	0.69	0.69	0.69	613	
weighted avg	0.73	0.73	0.73	613	

(1503, 10, 128)

(1503, 1, 10, 128)

True

Epoch 1/10

```

1503/1503 [=====] - 5s 3ms/step - loss: 1.0332 - acc:
0.6028

```

Epoch 2/10

```

1503/1503 [=====] - 2s 1ms/step - loss: 0.4492 - acc:
0.8011

```

Epoch 3/10

```

1503/1503 [=====] - 2s 1ms/step - loss: 0.3653 - acc:
0.8363

```

Epoch 4/10
1503/1503 [=====] - 2s 1ms/step - loss: 0.3303 - acc: 0.8417
Epoch 5/10
1503/1503 [=====] - 2s 1ms/step - loss: 0.3248 - acc: 0.8536
Epoch 6/10
1503/1503 [=====] - 2s 1ms/step - loss: 0.3090 - acc: 0.8709
Epoch 7/10
1503/1503 [=====] - 2s 1ms/step - loss: 0.2930 - acc: 0.8729
Epoch 8/10
1503/1503 [=====] - 2s 1ms/step - loss: 0.2492 - acc: 0.9002
Epoch 9/10
1503/1503 [=====] - 3s 2ms/step - loss: 0.2047 - acc: 0.9142A: 0s - loss: 0.2119 - acc: 0.90
Epoch 10/10
1503/1503 [=====] - 2s 2ms/step - loss: 0.1844 - acc: 0.9295
[0.943087]

violin

TRAIN				
	precision	recall	f1-score	support
False	0.99	0.90	0.94	629
True	0.93	1.00	0.96	874
accuracy			0.96	1503
macro avg	0.96	0.95	0.95	1503
weighted avg	0.96	0.96	0.96	1503

True

[0.95510423]

TEST				
	precision	recall	f1-score	support
False	0.81	0.71	0.76	231
True	0.79	0.88	0.83	299
accuracy			0.80	530
macro avg	0.80	0.79	0.79	530
weighted avg	0.80	0.80	0.80	530

(1188, 10, 128)

(1188, 1, 10, 128)

```

False
Epoch 1/10
1188/1188 [=====] - 8s 7ms/step - loss: 0.9186 - acc:
0.6490
Epoch 2/10
1188/1188 [=====] - 2s 1ms/step - loss: 0.3191 - acc:
0.8847
Epoch 3/10
1188/1188 [=====] - 2s 1ms/step - loss: 0.2256 - acc:
0.9234
Epoch 4/10
1188/1188 [=====] - 1s 1ms/step - loss: 0.1700 - acc:
0.9411
Epoch 5/10
1188/1188 [=====] - 1s 1ms/step - loss: 0.1404 - acc:
0.9529
Epoch 6/10
1188/1188 [=====] - 1s 1ms/step - loss: 0.1197 - acc:
0.9562
Epoch 7/10
1188/1188 [=====] - 1s 1ms/step - loss: 0.0945 - acc:
0.9621
Epoch 8/10
1188/1188 [=====] - 1s 1ms/step - loss: 0.1684 - acc:
0.9402
Epoch 9/10
1188/1188 [=====] - 1s 1ms/step - loss: 0.0842 - acc:
0.9722
Epoch 10/10
1188/1188 [=====] - 1s 1ms/step - loss: 0.0437 - acc:
0.9857
[1.1444092e-05]

```

voice

TRAIN				
	precision	recall	f1-score	support
False	1.00	0.98	0.99	429
True	0.99	1.00	0.99	759
accuracy			0.99	1188
macro avg	0.99	0.99	0.99	1188
weighted avg	0.99	0.99	0.99	1188

True

[0.99902856]

TEST

	precision	recall	f1-score	support
--	-----------	--------	----------	---------

False	0.90	0.84	0.87	147
True	0.90	0.94	0.92	229
accuracy			0.90	376
macro avg	0.90	0.89	0.90	376
weighted avg	0.90	0.90	0.90	376

```
[ ]: print(X_train_inst_sklearn)
      print(Y_pred_train)
```

```
[ ]:
```