openmic-MEL-DL

March 2, 2020

[1]:
```
import librosa as lb
import librosa.display
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
import keras
from keras.utils import np_utils
from keras import layers
from keras import models
from keras.models import Sequential
from keras.layers import Dense, Conv2D, MaxPool2D , Flatten, Dropout
from keras.preprocessing.image import ImageDataGenerator
from model_builder import build_example
from plotter import plot_history
import matplotlib.pyplot as plt
```

```
Using TensorFlow backend.
```

[2]:
```
# CONSTANTS

DATA_DIR = "openmic-2018/"
CATEGORY_COUNT = 8
LEARNING_RATE = 0.00001
THRESHOLD = 0.5
```

[ ]:
```
# MEL-SPECTOGRAM EXAMPLE

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)

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

[ ]:
```
MEL = []
sum = 0
for i in range(X.shape[0]):
    key = sample_key[i]
    key_pref = key[:3]
    y, sr = lb.load(DATA_DIR + 'audio/' + key_pref + '/' + key + '.ogg')
    S = lb.feature.melspectrogram(y=y, sr=sr)
    S_dB = lb.power_to_db(S, ref=0)
    MEL.append(S_dB[:,:430])
```

[ ]:
```
MEL_S = np.asarray(MEL)
print('Mel has shape: ' + str(MEL_S.shape))
```

[ ]:
```
# TODO SAVE WITHOUT X

np.savez('openmic-test-delete.npz', MEL = X, Y_true=Y_true, Y_mask=Y_mask,␣
 ↪sample_key=sample_key)
```

[ ]:
```
np.savez_compressed('openmic-mel-only.npz', MEL = MEL_S)
print('OpenMIC keys: ' + str(list(OPENMIC_2.keys())))
```

[3]:
```
OPENMIC_2 = np.load(os.path.join(DATA_DIR, 'openmic-mel.npz'),␣
 ↪allow_pickle=True)
X, Y_true, Y_mask, sample_key = OPENMIC_2['MEL'], OPENMIC_2['Y_true'],␣
 ↪OPENMIC_2['Y_mask'], OPENMIC_2['sample_key']
```

[ ]:
```
# LOAD DATA

OPENMIC = np.load(os.path.join(DATA_DIR, 'openmic-2018.npz'), allow_pickle=True)
print('OpenMIC keys: ' + str(list(OPENMIC.keys())))
X, Y_true, Y_mask, sample_key = OPENMIC['X'], OPENMIC['Y_true'],␣
 ↪OPENMIC['Y_mask'], OPENMIC['sample_key']
print('X has shape: ' + str(X.shape))
print('Y_true has shape: ' + str(Y_true.shape))
print('Y_mask has shape: ' + str(Y_mask.shape))
print('sample_key has shape: ' + str(sample_key.shape))
```

[4]:
```
# LOAD LABELS

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

```
print('OpenMIC instruments: ' + str(INSTRUMENTS))
```

```
OpenMIC instruments: {'accordion': 0, 'banjo': 1, 'bass': 2, 'cello': 3,
'clarinet': 4, 'cymbals': 5, 'drums': 6, 'flute': 7, 'guitar': 8,
'mallet_percussion': 9, 'mandolin': 10, 'organ': 11, 'piano': 12, 'saxophone':
13, 'synthesizer': 14, 'trombone': 15, 'trumpet': 16, 'ukulele': 17, 'violin':
18, 'voice': 19}
```

[5]:
```
# SPLIT DATA (TRAIN - TEST - VAL)

# CHANGE X TO MEL
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)
split_val, split_test, X_val, X_test, Y_true_val, Y_true_test, Y_mask_val,␣
 ↪Y_mask_test = train_test_split(split_test, X_test, Y_true_test, Y_mask_test,␣
 ↪test_size=0.5)
train_set = np.asarray(set(split_train))
test_set = np.asarray(set(split_test))
print('# Train: {}, # Val: {}, # Test: {}'.format(len(split_train),␣
 ↪len(split_test), len(split_val)))
```

```
# Train: 15000, # Val: 2500, # Test: 2500
```

[34]:
```
# DUPLICATE OF THE MODEL PREPROCESS

print(X_train.shape)
print(X_test.shape)

for instrument in INSTRUMENTS:

    # Map the instrument name to its column number
    inst_num = INSTRUMENTS[instrument]

    print(instrument)

    # TRAIN
    train_inst = Y_mask_train[:, inst_num]
    X_train_inst = X_train[train_inst]
    X_train_inst = X_train_inst.astype('float16')
    shape = X_train_inst.shape
    X_train_inst = X_train_inst.reshape(shape[0],1, shape[1], shape[2])
    Y_true_train_inst = Y_true_train[train_inst, inst_num] >= THRESHOLD
    i = 0
    for val in Y_true_train_inst:
        i += val
```

```
    print('TRAIN: ' + str(i) + ' true of ' + str(len(Y_true_train_inst)) + ' ('␣
↪+ str(round(i / len(Y_true_train_inst ) * 100,2)) + ' %)' )


    # TEST
    test_inst = Y_mask_test[:, inst_num]
    X_test_inst = X_test[test_inst]
    X_test_inst = X_test_inst.astype('float16')
    shape = X_test_inst.shape
    X_test_inst = X_test_inst.reshape(shape[0],1, shape[1], shape[2])
    Y_true_test_inst = Y_true_test[test_inst, inst_num] >= THRESHOLD

    i = 0
    for val in Y_true_test_inst:
        i += val

    print('TEST: ' + str(i) + ' true of ' + str(len(Y_true_test_inst)) + ' (' +␣
↪str(round(i / len(Y_true_test_inst ) * 100,2)) + ' %)' )


    # VALIDATION
    val_inst = Y_mask_val[:, inst_num]
    X_val_inst = X_val[val_inst]
    X_val_inst = X_val_inst.astype('float16')
    shape = X_val_inst.shape
    X_val_inst = X_val_inst.reshape(shape[0],1, shape[1], shape[2])
    Y_true_val_inst = Y_true_val[val_inst, inst_num] >= THRESHOLD


    i = 0
    for val in Y_true_val_inst:
        i += val
    print('VALIDATION: ' + str(i) + ' true of ' + str(len(Y_true_val_inst)) + '␣
↪(' + str(round(i / len(Y_true_val_inst ) * 100,2)) + ' %)' )
```

```
(15000, 128, 430)
(2500, 128, 430)
accordion
TRAIN: 367 true of 1540 (23.83 %)
TEST: 64 true of 279 (22.94 %)
VALIDATION: 58 true of 252 (23.02 %)
banjo
TRAIN: 532 true of 1620 (32.84 %)
TEST: 99 true of 307 (32.25 %)
VALIDATION: 101 true of 291 (34.71 %)
bass
TRAIN: 410 true of 1401 (29.26 %)
```

```
TEST: 65 true of 234 (27.78 %)
VALIDATION: 74 true of 253 (29.25 %)
cello
TRAIN: 643 true of 1490 (43.15 %)
TEST: 97 true of 243 (39.92 %)
VALIDATION: 84 true of 216 (38.89 %)
clarinet
TRAIN: 411 true of 1810 (22.71 %)
TEST: 59 true of 293 (20.14 %)
VALIDATION: 63 true of 282 (22.34 %)
cymbals
TRAIN: 816 true of 1280 (63.75 %)
TEST: 144 true of 225 (64.0 %)
VALIDATION: 151 true of 230 (65.65 %)
drums
TRAIN: 827 true of 1313 (62.99 %)
TEST: 144 true of 226 (63.72 %)
VALIDATION: 135 true of 208 (64.9 %)
flute
TRAIN: 484 true of 1562 (30.99 %)
TEST: 82 true of 277 (29.6 %)
VALIDATION: 81 true of 245 (33.06 %)
guitar
TRAIN: 859 true of 1232 (69.72 %)
TEST: 147 true of 215 (68.37 %)
VALIDATION: 132 true of 203 (65.02 %)
mallet_percussion
TRAIN: 561 true of 1393 (40.27 %)
TEST: 67 true of 191 (35.08 %)
VALIDATION: 105 true of 218 (48.17 %)
mandolin
TRAIN: 619 true of 1799 (34.41 %)
TEST: 107 true of 314 (34.08 %)
VALIDATION: 119 true of 351 (33.9 %)
organ
TRAIN: 454 true of 1427 (31.81 %)
TEST: 65 true of 224 (29.02 %)
VALIDATION: 84 true of 239 (35.15 %)
piano
TRAIN: 871 true of 1284 (67.83 %)
TEST: 164 true of 237 (69.2 %)
VALIDATION: 135 true of 199 (67.84 %)
saxophone
TRAIN: 841 true of 1769 (47.54 %)
TEST: 141 true of 286 (49.3 %)
VALIDATION: 153 true of 310 (49.35 %)
synthesizer
TRAIN: 798 true of 1178 (67.74 %)
```

```
TEST: 146 true of 225 (64.89 %)
VALIDATION: 147 true of 199 (73.87 %)
trombone
TRAIN: 653 true of 2058 (31.73 %)
TEST: 110 true of 362 (30.39 %)
VALIDATION: 100 true of 340 (29.41 %)
trumpet
TRAIN: 861 true of 2179 (39.51 %)
TEST: 145 true of 385 (37.66 %)
VALIDATION: 140 true of 352 (39.77 %)
ukulele
TRAIN: 542 true of 1805 (30.03 %)
TEST: 90 true of 298 (30.2 %)
VALIDATION: 106 true of 322 (32.92 %)
violin
TRAIN: 884 true of 1529 (57.82 %)
TEST: 138 true of 244 (56.56 %)
VALIDATION: 151 true of 260 (58.08 %)
voice
TRAIN: 752 true of 1174 (64.05 %)
TEST: 126 true of 197 (63.96 %)
VALIDATION: 110 true of 193 (56.99 %)
```

[15]:
```python
# VALAMI FANCY ADATKIÍRÁS
len(Y_true_val_inst)
```

[15]: 193

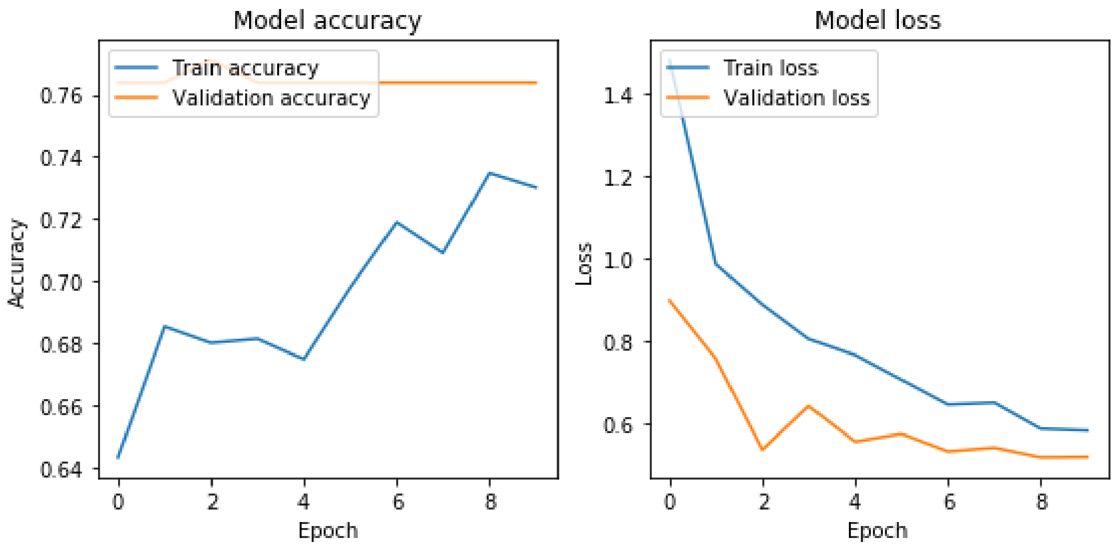
[9]:
```python
# This dictionary will include the classifiers for each model
mymodels = dict()
# We'll iterate over all istrument classes, and fit a model for each one
# After training, we'll print a classification report for each instrument
for instrument in INSTRUMENTS:

    # Map the instrument name to its column number
    inst_num = INSTRUMENTS[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
    # 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
    # Again, we slice the labels to the annotated examples
    # We thresold the label likelihoods at 0.5 to get binary labels

    # TRAIN
    train_inst = Y_mask_train[:, inst_num]
```

```
    X_train_inst = X_train[train_inst]
    X_train_inst = X_train_inst.astype('float16')
    shape = X_train_inst.shape
    X_train_inst = X_train_inst.reshape(shape[0],1, shape[1], shape[2])
    Y_true_train_inst = Y_true_train[train_inst, inst_num] >= THRESHOLD

    # TEST
    test_inst = Y_mask_test[:, inst_num]
    X_test_inst = X_test[test_inst]
    X_test_inst = X_test_inst.astype('float16')
    shape = X_test_inst.shape
    X_test_inst = X_test_inst.reshape(shape[0],1, shape[1], shape[2])
    Y_true_test_inst = Y_true_test[test_inst, inst_num] >= THRESHOLD

    # VALIDATION
    val_inst = Y_mask_val[:, inst_num]
    X_val_inst = X_val[val_inst]
    X_val_inst = X_val_inst.astype('float16')
    shape = X_val_inst.shape
    X_val_inst = X_val_inst.reshape(shape[0],1, shape[1], shape[2])
    Y_true_val_inst = Y_true_val[val_inst, inst_num] >= THRESHOLD


    # Step 3.
    # Initialize a new classifier
    model = models.Sequential()
    model.
↪add(Conv2D(input_shape=(1,128,430),data_format="channels_first",filters=32,kernel_size=(3,3
↪activation="relu"))
    model.add(Conv2D(filters=32,kernel_size=(3,3),padding="same",␣
↪activation="relu"))
    model.add(MaxPool2D(pool_size=(3,3),strides=(2,2)))
    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(MaxPool2D(pool_size=(2,2),strides=(2,2)))
    model.add(Dropout(0.2))
    model.add(Conv2D(filters=256, kernel_size=(3,3), padding="same",␣
↪activation="relu"))
    model.add(MaxPool2D(pool_size=(2,2),strides=(2,2)))
    model.add(layers.Flatten())
    model.add(layers.Dense(units=256, activation='relu'))
    model.add(layers.Dense(units=1, activation='sigmoid'))

    model.compile(loss='binary_crossentropy', optimizer=keras.optimizers.
↪Adam(lr= LEARNING_RATE), metrics = ['accuracy'])
```

```
    # model.summary()
    # Step 4.
    history = model.fit(X_train_inst,Y_true_train_inst , epochs=10,␣
 ↪batch_size=64, validation_data=(X_val_inst,Y_true_val_inst))

    plot_history()

    loss, acc = model.evaluate(X_test_inst, Y_true_test_inst)
    print('Test loss: {}'.format(loss))
    print('Test accuracy: {:.2%}'.format(acc))
    # 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 > THRESHOLD - 0.15 #THRESHOLD (should be␣
 ↪lower than 0.5)
    Y_pred_test_bool = Y_pred_test > THRESHOLD - 0.15 #THRESHOLD (should be␣
 ↪lower than 0.5)
    print('-' * 52)
    print(instrument)
    print('\tTRAIN')
    print(classification_report(Y_true_train_inst, Y_pred_train_bool))
    print('\tTEST')
    print(classification_report(Y_true_test_inst, Y_pred_test_bool))

    # Store the classifier in our dictionary
mymodels[instrument] = model
```

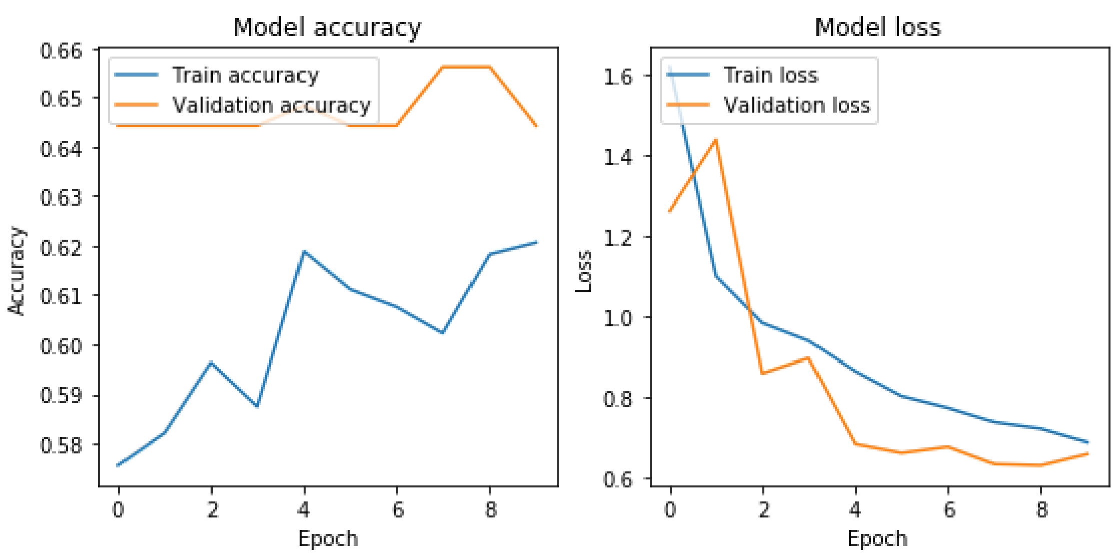
```
WARNING:tensorflow:From C:\Users\user\Anaconda3\lib\site-
packages\tensorflow\python\framework\op_def_library.py:263: colocate_with (from
tensorflow.python.framework.ops) is deprecated and will be removed in a future
version.
Instructions for updating:
Colocations handled automatically by placer.
WARNING:tensorflow:From C:\Users\user\Anaconda3\lib\site-
packages\keras\backend\tensorflow_backend.py:3733: calling dropout (from
tensorflow.python.ops.nn_ops) with keep_prob is deprecated and will be removed
in a future version.
Instructions for updating:
Please use `rate` instead of `keep_prob`. Rate should be set to `rate = 1 -
keep_prob`.
WARNING:tensorflow:From C:\Users\user\Anaconda3\lib\site-
packages\tensorflow\python\ops\math_ops.py:3066: to_int32 (from
tensorflow.python.ops.math_ops) is deprecated and will be removed in a future
version.
Instructions for updating:
```

```
Use tf.cast instead.
Train on 1519 samples, validate on 271 samples
Epoch 1/10
1519/1519 [==============================] - 243s 160ms/step - loss: 1.4823 -
acc: 0.6432 - val_loss: 0.8994 - val_acc: 0.7638
Epoch 2/10
1519/1519 [==============================] - 233s 154ms/step - loss: 0.9874 -
acc: 0.6853 - val_loss: 0.7584 - val_acc: 0.7638
Epoch 3/10
1519/1519 [==============================] - 232s 153ms/step - loss: 0.8897 -
acc: 0.6801 - val_loss: 0.5379 - val_acc: 0.7712
Epoch 4/10
1519/1519 [==============================] - 231s 152ms/step - loss: 0.8063 -
acc: 0.6814 - val_loss: 0.6440 - val_acc: 0.7638
Epoch 5/10
1519/1519 [==============================] - 239s 157ms/step - loss: 0.7674 -
acc: 0.6748 - val_loss: 0.5574 - val_acc: 0.7638
Epoch 6/10
1519/1519 [==============================] - 248s 163ms/step - loss: 0.7075 -
acc: 0.6978 - val_loss: 0.5762 - val_acc: 0.7638
Epoch 7/10
1519/1519 [==============================] - 253s 167ms/step - loss: 0.6475 -
acc: 0.7189 - val_loss: 0.5339 - val_acc: 0.7638
Epoch 8/10
1519/1519 [==============================] - 253s 167ms/step - loss: 0.6519 -
acc: 0.7090 - val_loss: 0.5428 - val_acc: 0.7638
Epoch 9/10
1519/1519 [==============================] - 252s 166ms/step - loss: 0.5890 -
acc: 0.7347 - val_loss: 0.5197 - val_acc: 0.7638
Epoch 10/10
1519/1519 [==============================] - 257s 169ms/step - loss: 0.5847 -
acc: 0.7301 - val_loss: 0.5208 - val_acc: 0.7638
```




```
281/281 [==============================] - 17s 61ms/step
Test loss: 0.5738391290780064
Test accuracy: 75.09%
----------------------------------------------------
accordion
        TRAIN
              precision    recall  f1-score   support

       False       0.77      0.98      0.86      1164
        True       0.47      0.07      0.12       355

    accuracy                           0.76      1519
   macro avg       0.62      0.52      0.49      1519
weighted avg       0.70      0.76      0.69      1519

        TEST
              precision    recall  f1-score   support

       False       0.75      0.97      0.85       211
        True       0.30      0.04      0.07        70

    accuracy                           0.74       281
   macro avg       0.53      0.50      0.46       281
weighted avg       0.64      0.74      0.65       281

Train on 1687 samples, validate on 253 samples
Epoch 1/10
1687/1687 [==============================] - 285s 169ms/step - loss: 1.6191 -
acc: 0.5756 - val_loss: 1.2620 - val_acc: 0.6443
```

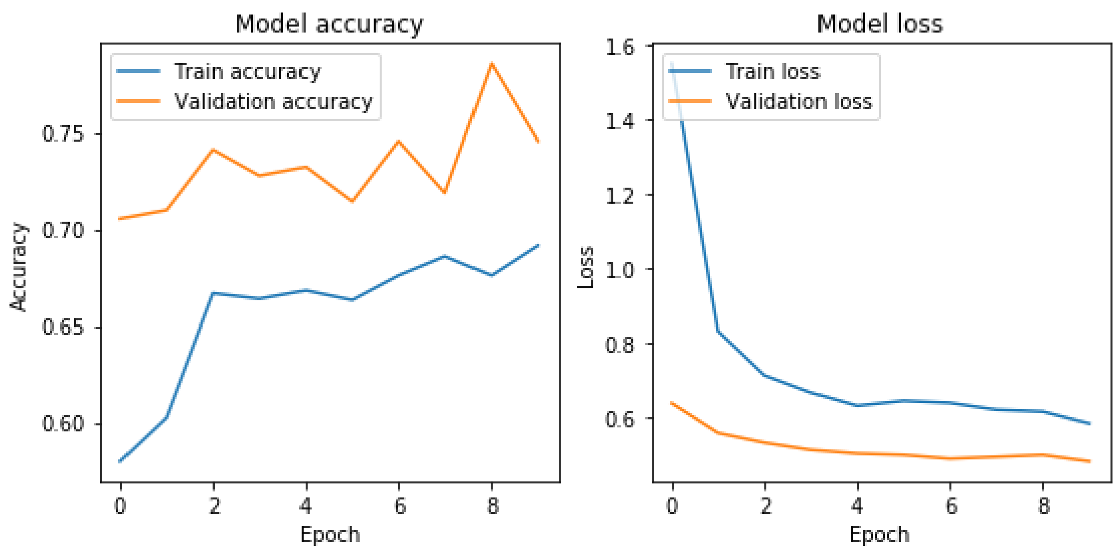
```
Epoch 2/10
1687/1687 [==============================] - 287s 170ms/step - loss: 1.1000 -
acc: 0.5821 - val_loss: 1.4386 - val_acc: 0.6443
Epoch 3/10
1687/1687 [==============================] - 290s 172ms/step - loss: 0.9832 -
acc: 0.5963 - val_loss: 0.8577 - val_acc: 0.6443
Epoch 4/10
1687/1687 [==============================] - 289s 171ms/step - loss: 0.9389 -
acc: 0.5874 - val_loss: 0.8959 - val_acc: 0.6443
Epoch 5/10
1687/1687 [==============================] - 289s 172ms/step - loss: 0.8633 -
acc: 0.6189 - val_loss: 0.6830 - val_acc: 0.6482
Epoch 6/10
1687/1687 [==============================] - 290s 172ms/step - loss: 0.8017 -
acc: 0.6111 - val_loss: 0.6609 - val_acc: 0.6443
Epoch 7/10
1687/1687 [==============================] - 286s 170ms/step - loss: 0.7723 -
acc: 0.6076 - val_loss: 0.6761 - val_acc: 0.6443
Epoch 8/10
1687/1687 [==============================] - 291s 172ms/step - loss: 0.7369 -
acc: 0.6023 - val_loss: 0.6332 - val_acc: 0.6561
Epoch 9/10
1687/1687 [==============================] - 293s 174ms/step - loss: 0.7206 -
acc: 0.6183 - val_loss: 0.6299 - val_acc: 0.6561
Epoch 10/10
1687/1687 [==============================] - 289s 171ms/step - loss: 0.6872 -
acc: 0.6206 - val_loss: 0.6582 - val_acc: 0.6443
```




```
278/278 [==============================] - 18s 66ms/step
```

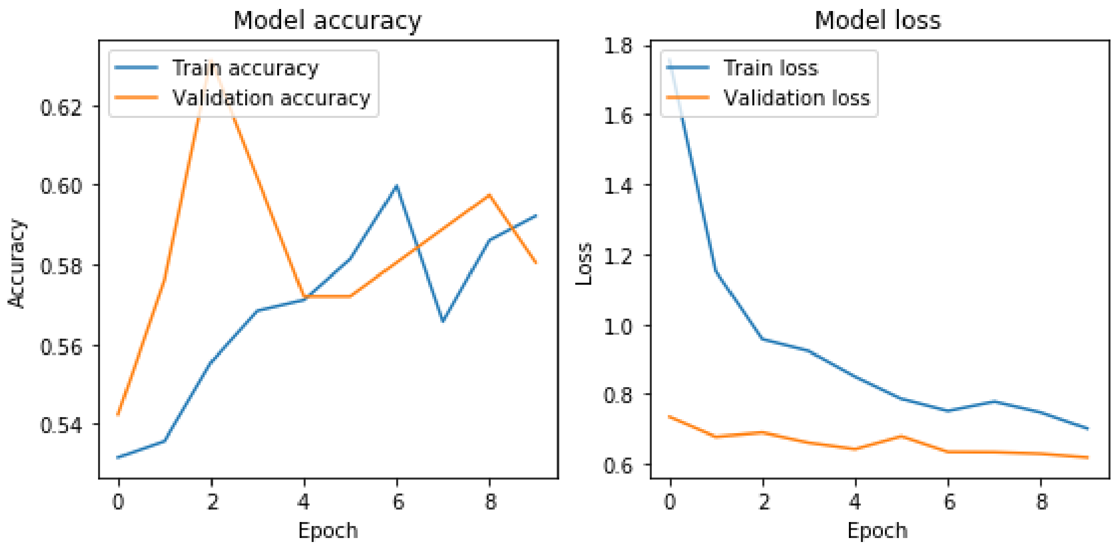
```
Test loss: 0.6368379172661321
Test accuracy: 66.55%
----------------------------------------------------
banjo
        TRAIN
              precision    recall  f1-score   support

       False       0.69      0.93      0.79      1138
        True       0.48      0.13      0.20       549

    accuracy                           0.67      1687
   macro avg       0.58      0.53      0.50      1687
weighted avg       0.62      0.67      0.60      1687

        TEST
              precision    recall  f1-score   support

       False       0.66      0.89      0.76       185
        True       0.32      0.11      0.16        93

    accuracy                           0.63       278
   macro avg       0.49      0.50      0.46       278
weighted avg       0.55      0.63      0.56       278

Train on 1416 samples, validate on 224 samples
Epoch 1/10
1416/1416 [==============================] - 244s 172ms/step - loss: 1.5489 -
acc: 0.5798 - val_loss: 0.6392 - val_acc: 0.7054
Epoch 2/10
1416/1416 [==============================] - 246s 174ms/step - loss: 0.8325 -
acc: 0.6024 - val_loss: 0.5599 - val_acc: 0.7098
Epoch 3/10
1416/1416 [==============================] - 245s 173ms/step - loss: 0.7140 -
acc: 0.6667 - val_loss: 0.5339 - val_acc: 0.7411
Epoch 4/10
1416/1416 [==============================] - 242s 171ms/step - loss: 0.6678 -
acc: 0.6638 - val_loss: 0.5146 - val_acc: 0.7277
Epoch 5/10
1416/1416 [==============================] - 245s 173ms/step - loss: 0.6324 -
acc: 0.6681 - val_loss: 0.5047 - val_acc: 0.7321
Epoch 6/10
1416/1416 [==============================] - 249s 176ms/step - loss: 0.6456 -
acc: 0.6631 - val_loss: 0.5009 - val_acc: 0.7143
Epoch 7/10
1416/1416 [==============================] - 245s 173ms/step - loss: 0.6402 -
acc: 0.6758 - val_loss: 0.4908 - val_acc: 0.7455
Epoch 8/10
1416/1416 [==============================] - 243s 172ms/step - loss: 0.6219 -
```

```
acc: 0.6857 - val_loss: 0.4957 - val_acc: 0.7188
Epoch 9/10
1416/1416 [==============================] - 243s 172ms/step - loss: 0.6169 -
acc: 0.6758 - val_loss: 0.5009 - val_acc: 0.7857
Epoch 10/10
1416/1416 [==============================] - 243s 172ms/step - loss: 0.5850 -
acc: 0.6914 - val_loss: 0.4838 - val_acc: 0.7455
```



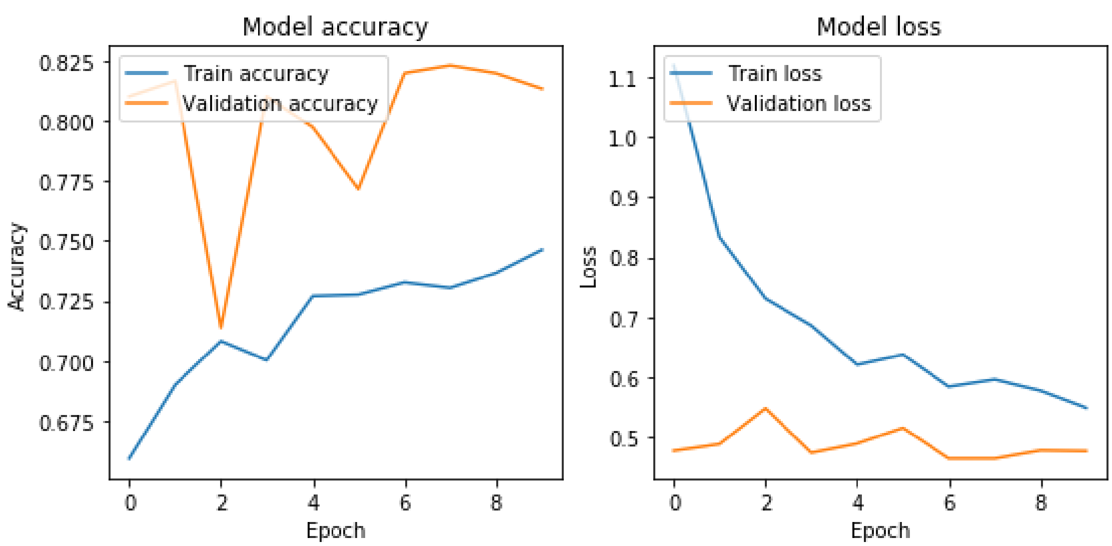

```
248/248 [==============================] - 15s 62ms/step
Test loss: 0.48448632417186616
Test accuracy: 76.61%
----------------------------------------------------
bass
        TRAIN
              precision    recall  f1-score   support

       False       0.83      0.71      0.77      1003
        True       0.48      0.65      0.55       413

    accuracy                           0.70      1416
   macro avg       0.66      0.68      0.66      1416
weighted avg       0.73      0.70      0.71      1416

        TEST
              precision    recall  f1-score   support

       False       0.89      0.79      0.83       178
        True       0.58      0.74      0.65        70
```

```
    accuracy                           0.77       248
   macro avg       0.73      0.76      0.74       248
weighted avg       0.80      0.77      0.78       248

Train on 1469 samples, validate on 236 samples
Epoch 1/10
1469/1469 [==============================] - 254s 173ms/step - loss: 1.7556 -
acc: 0.5317 - val_loss: 0.7347 - val_acc: 0.5424
Epoch 2/10
1469/1469 [==============================] - 251s 171ms/step - loss: 1.1518 -
acc: 0.5357 - val_loss: 0.6780 - val_acc: 0.5763
Epoch 3/10
1469/1469 [==============================] - 251s 171ms/step - loss: 0.9568 -
acc: 0.5555 - val_loss: 0.6902 - val_acc: 0.6314
Epoch 4/10
1469/1469 [==============================] - 235s 160ms/step - loss: 0.9233 -
acc: 0.5684 - val_loss: 0.6616 - val_acc: 0.6017
Epoch 5/10
1469/1469 [==============================] - 235s 160ms/step - loss: 0.8500 -
acc: 0.5711 - val_loss: 0.6432 - val_acc: 0.5720
Epoch 6/10
1469/1469 [==============================] - 234s 160ms/step - loss: 0.7865 -
acc: 0.5813 - val_loss: 0.6798 - val_acc: 0.5720
Epoch 7/10
1469/1469 [==============================] - 235s 160ms/step - loss: 0.7516 -
acc: 0.5997 - val_loss: 0.6352 - val_acc: 0.5805
Epoch 8/10
1469/1469 [==============================] - 234s 159ms/step - loss: 0.7784 -
acc: 0.5657 - val_loss: 0.6345 - val_acc: 0.5890
Epoch 9/10
1469/1469 [==============================] - 235s 160ms/step - loss: 0.7473 -
acc: 0.5861 - val_loss: 0.6299 - val_acc: 0.5975
Epoch 10/10
1469/1469 [==============================] - 236s 161ms/step - loss: 0.7013 -
acc: 0.5922 - val_loss: 0.6194 - val_acc: 0.5805
```




```
244/244 [==============================] - 15s 60ms/step
Test loss: 0.5947532761292379
Test accuracy: 65.16%
----------------------------------------------------
cello
        TRAIN
              precision    recall  f1-score   support

       False       0.77      0.53      0.63       843
        True       0.56      0.79      0.65       626

    accuracy                           0.64      1469
   macro avg       0.66      0.66      0.64      1469
weighted avg       0.68      0.64      0.64      1469

        TEST
              precision    recall  f1-score   support

       False       0.83      0.52      0.64       145
        True       0.55      0.84      0.66        99

    accuracy                           0.65       244
   macro avg       0.69      0.68      0.65       244
weighted avg       0.71      0.65      0.65       244

Train on 1758 samples, validate on 311 samples
Epoch 1/10
1758/1758 [==============================] - 281s 160ms/step - loss: 1.1186 -
acc: 0.6593 - val_loss: 0.4773 - val_acc: 0.8103
```

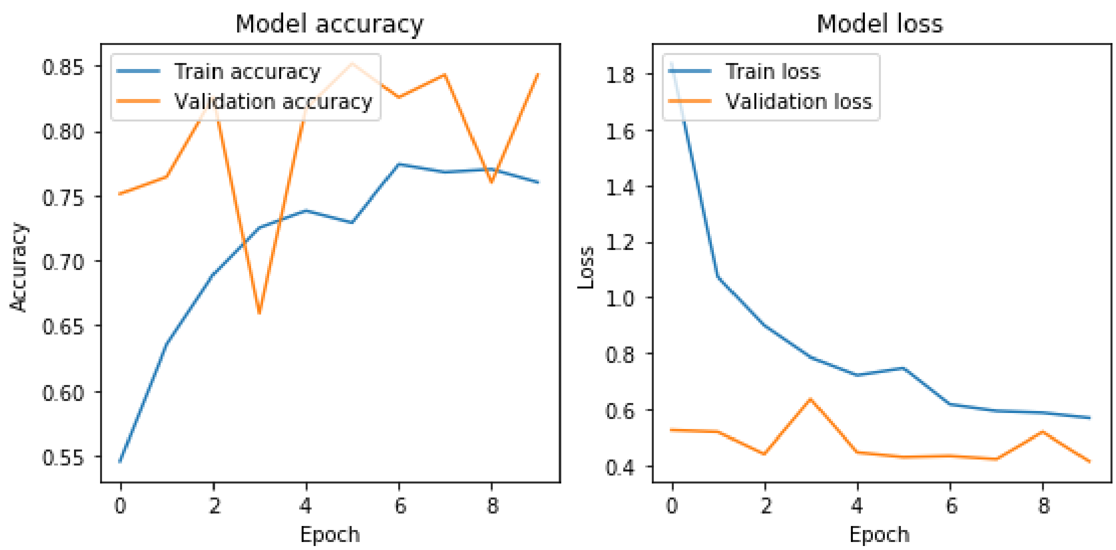
```
Epoch 2/10
1758/1758 [==============================] - 281s 160ms/step - loss: 0.8329 -
acc: 0.6900 - val_loss: 0.4884 - val_acc: 0.8167
Epoch 3/10
1758/1758 [==============================] - 284s 161ms/step - loss: 0.7305 -
acc: 0.7082 - val_loss: 0.5479 - val_acc: 0.7138
Epoch 4/10
1758/1758 [==============================] - 281s 160ms/step - loss: 0.6857 -
acc: 0.7002 - val_loss: 0.4740 - val_acc: 0.8103
Epoch 5/10
1758/1758 [==============================] - 280s 159ms/step - loss: 0.6204 -
acc: 0.7270 - val_loss: 0.4893 - val_acc: 0.7974
Epoch 6/10
1758/1758 [==============================] - 281s 160ms/step - loss: 0.6374 -
acc: 0.7275 - val_loss: 0.5152 - val_acc: 0.7717
Epoch 7/10
1758/1758 [==============================] - 282s 160ms/step - loss: 0.5840 -
acc: 0.7327 - val_loss: 0.4645 - val_acc: 0.8199
Epoch 8/10
1758/1758 [==============================] - 282s 160ms/step - loss: 0.5959 -
acc: 0.7304 - val_loss: 0.4645 - val_acc: 0.8232
Epoch 9/10
1758/1758 [==============================] - 280s 160ms/step - loss: 0.5773 -
acc: 0.7366 - val_loss: 0.4779 - val_acc: 0.8199
Epoch 10/10
1758/1758 [==============================] - 280s 159ms/step - loss: 0.5487 -
acc: 0.7463 - val_loss: 0.4769 - val_acc: 0.8135
```




```
316/316 [==============================] - 17s 55ms/step
Test loss: 0.5168347962294952
```

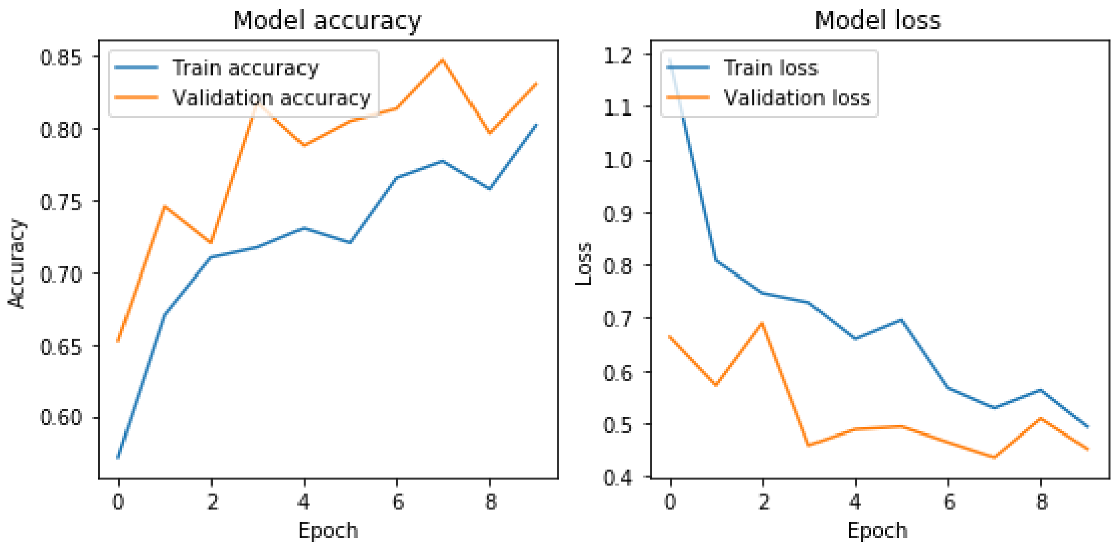
```
Test accuracy: 75.32%
----------------------------------------------------
clarinet
        TRAIN
              precision    recall  f1-score   support

       False       0.80      0.87      0.83      1357
        True       0.37      0.26      0.31       401

    accuracy                           0.73      1758
   macro avg       0.59      0.57      0.57      1758
weighted avg       0.70      0.73      0.71      1758

        TEST
              precision    recall  f1-score   support

       False       0.80      0.86      0.83       240
        True       0.43      0.34      0.38        76

    accuracy                           0.73       316
   macro avg       0.62      0.60      0.61       316
weighted avg       0.72      0.73      0.72       316

Train on 1292 samples, validate on 229 samples
Epoch 1/10
1292/1292 [==============================] - 207s 160ms/step - loss: 1.8346 -
acc: 0.5457 - val_loss: 0.5255 - val_acc: 0.7511
Epoch 2/10
1292/1292 [==============================] - 208s 161ms/step - loss: 1.0731 -
acc: 0.6354 - val_loss: 0.5193 - val_acc: 0.7642
Epoch 3/10
1292/1292 [==============================] - 208s 161ms/step - loss: 0.8994 -
acc: 0.6889 - val_loss: 0.4405 - val_acc: 0.8253
Epoch 4/10
1292/1292 [==============================] - 206s 159ms/step - loss: 0.7858 -
acc: 0.7252 - val_loss: 0.6367 - val_acc: 0.6594
Epoch 5/10
1292/1292 [==============================] - 205s 159ms/step - loss: 0.7221 -
acc: 0.7384 - val_loss: 0.4469 - val_acc: 0.8166
Epoch 6/10
1292/1292 [==============================] - 208s 161ms/step - loss: 0.7478 -
acc: 0.7291 - val_loss: 0.4300 - val_acc: 0.8515
Epoch 7/10
1292/1292 [==============================] - 206s 160ms/step - loss: 0.6178 -
acc: 0.7740 - val_loss: 0.4340 - val_acc: 0.8253
Epoch 8/10
1292/1292 [==============================] - 206s 160ms/step - loss: 0.5941 -
acc: 0.7678 - val_loss: 0.4227 - val_acc: 0.8428
```

```
Epoch 9/10
1292/1292 [==============================] - 204s 158ms/step - loss: 0.5877 -
acc: 0.7701 - val_loss: 0.5193 - val_acc: 0.7598
Epoch 10/10
1292/1292 [==============================] - 206s 160ms/step - loss: 0.5692 -
acc: 0.7601 - val_loss: 0.4148 - val_acc: 0.8428
```



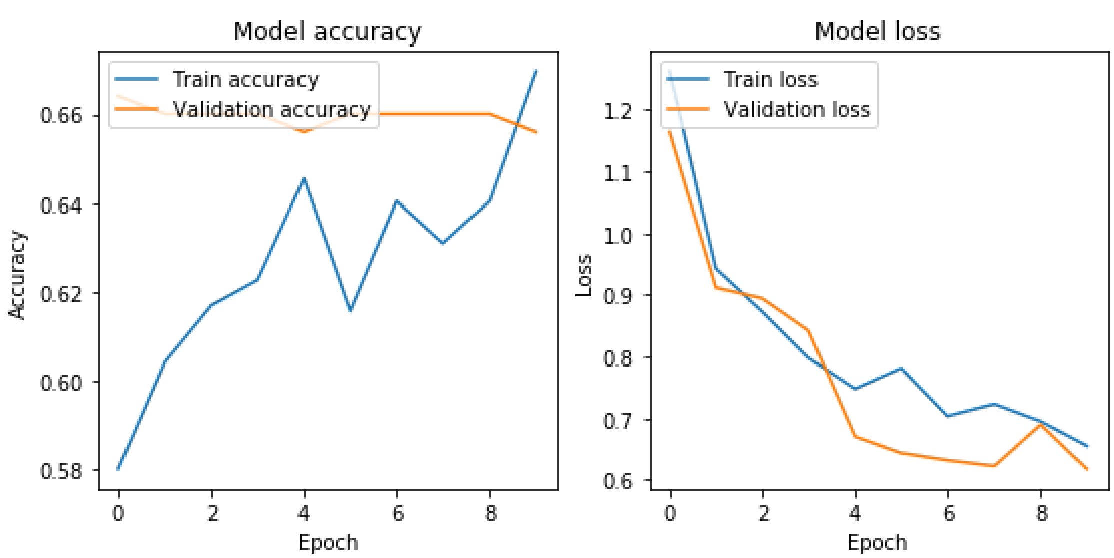

```
214/214 [==============================] - 13s 59ms/step
Test loss: 0.3548090073549859
Test accuracy: 84.11%
----------------------------------------------------
cymbals
        TRAIN
              precision    recall  f1-score   support

       False       0.94      0.65      0.77       468
        True       0.83      0.98      0.90       824

    accuracy                           0.86      1292
   macro avg       0.88      0.81      0.83      1292
weighted avg       0.87      0.86      0.85      1292

        TEST
              precision    recall  f1-score   support

       False       0.88      0.59      0.71        74
        True       0.82      0.96      0.88       140

    accuracy                           0.83       214
```

```
   macro avg       0.85      0.78      0.80       214
weighted avg       0.84      0.83      0.82       214

Train on 1284 samples, validate on 236 samples
Epoch 1/10
1284/1284 [==============================] - 206s 160ms/step - loss: 1.1872 -
acc: 0.5717 - val_loss: 0.6638 - val_acc: 0.6525
Epoch 2/10
1284/1284 [==============================] - 207s 161ms/step - loss: 0.8077 -
acc: 0.6706 - val_loss: 0.5718 - val_acc: 0.7458
Epoch 3/10
1284/1284 [==============================] - 204s 159ms/step - loss: 0.7469 -
acc: 0.7103 - val_loss: 0.6901 - val_acc: 0.7203
Epoch 4/10
1284/1284 [==============================] - 209s 163ms/step - loss: 0.7289 -
acc: 0.7173 - val_loss: 0.4587 - val_acc: 0.8178
Epoch 5/10
1284/1284 [==============================] - 211s 165ms/step - loss: 0.6601 -
acc: 0.7305 - val_loss: 0.4888 - val_acc: 0.7881
Epoch 6/10
1284/1284 [==============================] - 213s 166ms/step - loss: 0.6961 -
acc: 0.7204 - val_loss: 0.4940 - val_acc: 0.8051
Epoch 7/10
1284/1284 [==============================] - 205s 160ms/step - loss: 0.5670 -
acc: 0.7656 - val_loss: 0.4645 - val_acc: 0.8136
Epoch 8/10
1284/1284 [==============================] - 208s 162ms/step - loss: 0.5289 -
acc: 0.7773 - val_loss: 0.4360 - val_acc: 0.8475
Epoch 9/10
1284/1284 [==============================] - 207s 161ms/step - loss: 0.5629 -
acc: 0.7578 - val_loss: 0.5090 - val_acc: 0.7966
Epoch 10/10
1284/1284 [==============================] - 207s 161ms/step - loss: 0.4937 -
acc: 0.8022 - val_loss: 0.4521 - val_acc: 0.8305
```




```
227/227 [==============================] - 13s 57ms/step
Test loss: 0.5031188363414504
Test accuracy: 82.38%
----------------------------------------------------
drums
        TRAIN
              precision    recall  f1-score   support

       False       0.90      0.72      0.80       466
        True       0.86      0.95      0.90       818

    accuracy                           0.87      1284
   macro avg       0.88      0.84      0.85      1284
weighted avg       0.87      0.87      0.87      1284

        TEST
              precision    recall  f1-score   support

       False       0.91      0.67      0.77        91
        True       0.81      0.96      0.88       136

    accuracy                           0.84       227
   macro avg       0.86      0.81      0.83       227
weighted avg       0.85      0.84      0.84       227

Train on 1577 samples, validate on 250 samples
Epoch 1/10
1577/1577 [==============================] - 254s 161ms/step - loss: 1.2625 -
acc: 0.5802 - val_loss: 1.1643 - val_acc: 0.6640
```

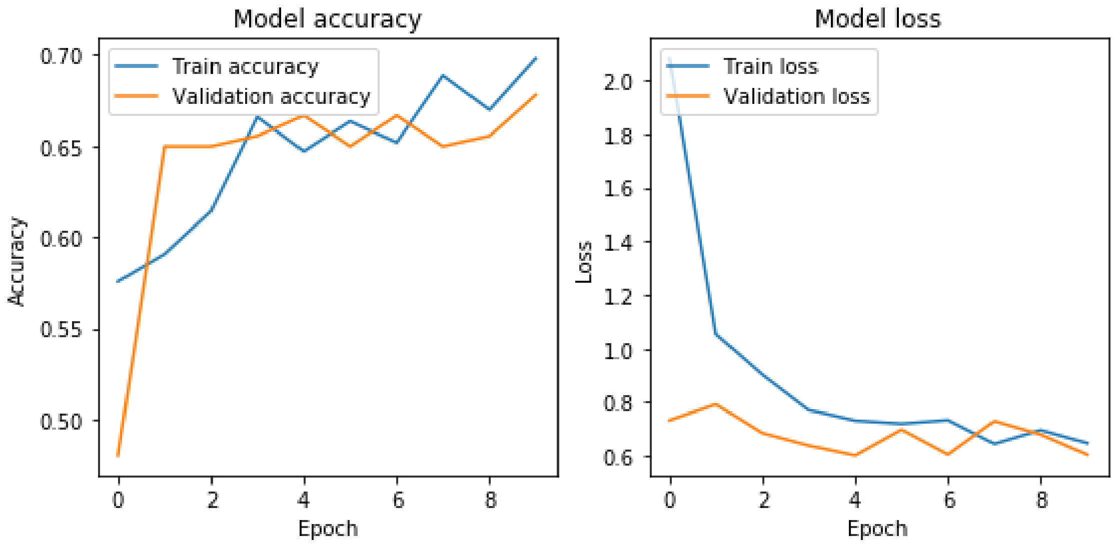
```
Epoch 2/10
1577/1577 [==============================] - 251s 159ms/step - loss: 0.9430 -
acc: 0.6043 - val_loss: 0.9120 - val_acc: 0.6600
Epoch 3/10
1577/1577 [==============================] - 249s 158ms/step - loss: 0.8731 -
acc: 0.6170 - val_loss: 0.8948 - val_acc: 0.6600
Epoch 4/10
1577/1577 [==============================] - 250s 158ms/step - loss: 0.7980 -
acc: 0.6227 - val_loss: 0.8425 - val_acc: 0.6600
Epoch 5/10
1577/1577 [==============================] - 245s 156ms/step - loss: 0.7479 -
acc: 0.6455 - val_loss: 0.6705 - val_acc: 0.6560
Epoch 6/10
1577/1577 [==============================] - 251s 159ms/step - loss: 0.7813 -
acc: 0.6157 - val_loss: 0.6440 - val_acc: 0.6600
Epoch 7/10
1577/1577 [==============================] - 247s 157ms/step - loss: 0.7038 -
acc: 0.6405 - val_loss: 0.6321 - val_acc: 0.6600
Epoch 8/10
1577/1577 [==============================] - 249s 158ms/step - loss: 0.7232 -
acc: 0.6309 - val_loss: 0.6230 - val_acc: 0.6600
Epoch 9/10
1577/1577 [==============================] - 248s 157ms/step - loss: 0.6951 -
acc: 0.6405 - val_loss: 0.6895 - val_acc: 0.6600
Epoch 10/10
1577/1577 [==============================] - 250s 159ms/step - loss: 0.6554 -
acc: 0.6696 - val_loss: 0.6181 - val_acc: 0.6560
```




```
257/257 [==============================] - 15s 56ms/step
```

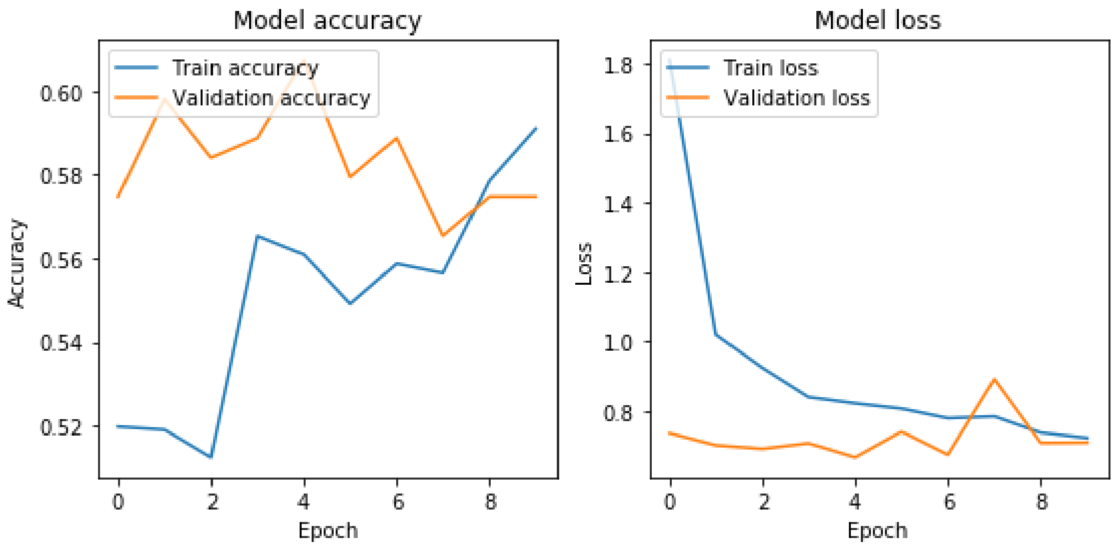
```
Test loss: 0.6230789652594333
Test accuracy: 66.15%
----------------------------------------------------
flute
        TRAIN
              precision    recall  f1-score   support

       False       0.73      0.84      0.78      1094
        True       0.44      0.30      0.36       483

    accuracy                           0.67      1577
   macro avg       0.59      0.57      0.57      1577
weighted avg       0.64      0.67      0.65      1577

        TEST
              precision    recall  f1-score   support

       False       0.71      0.80      0.76       177
        True       0.40      0.29      0.33        80

    accuracy                           0.64       257
   macro avg       0.56      0.54      0.54       257
weighted avg       0.61      0.64      0.62       257

Train on 1275 samples, validate on 177 samples
Epoch 1/10
1275/1275 [==============================] - 203s 159ms/step - loss: 2.0818 -
acc: 0.5757 - val_loss: 0.7326 - val_acc: 0.4802
Epoch 2/10
1275/1275 [==============================] - 201s 158ms/step - loss: 1.0547 -
acc: 0.5906 - val_loss: 0.7925 - val_acc: 0.6497
Epoch 3/10
1275/1275 [==============================] - 201s 157ms/step - loss: 0.9034 -
acc: 0.6141 - val_loss: 0.6851 - val_acc: 0.6497
Epoch 4/10
1275/1275 [==============================] - 203s 159ms/step - loss: 0.7712 -
acc: 0.6659 - val_loss: 0.6389 - val_acc: 0.6554
Epoch 5/10
1275/1275 [==============================] - 201s 157ms/step - loss: 0.7313 -
acc: 0.6471 - val_loss: 0.6019 - val_acc: 0.6667
Epoch 6/10
1275/1275 [==============================] - 201s 158ms/step - loss: 0.7199 -
acc: 0.6635 - val_loss: 0.6976 - val_acc: 0.6497
Epoch 7/10
1275/1275 [==============================] - 202s 158ms/step - loss: 0.7332 -
acc: 0.6518 - val_loss: 0.6054 - val_acc: 0.6667
Epoch 8/10
1275/1275 [==============================] - 222s 174ms/step - loss: 0.6452 -
```

```
acc: 0.6886 - val_loss: 0.7294 - val_acc: 0.6497
Epoch 9/10
1275/1275 [==============================] - 217s 170ms/step - loss: 0.6961 -
acc: 0.6698 - val_loss: 0.6796 - val_acc: 0.6554
Epoch 10/10
1275/1275 [==============================] - 217s 170ms/step - loss: 0.6479 -
acc: 0.6980 - val_loss: 0.6050 - val_acc: 0.6780
```



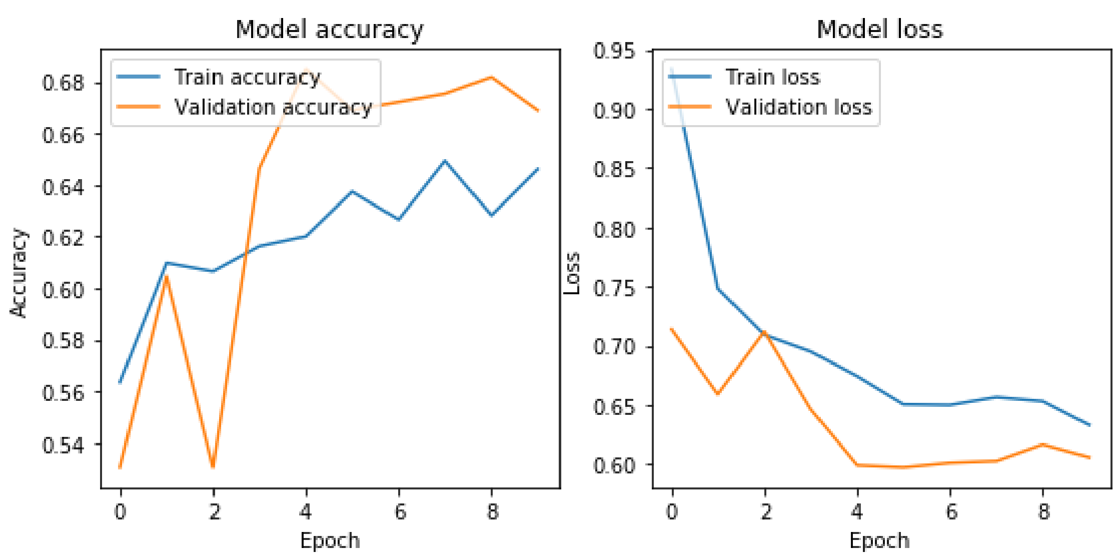

```
198/198 [==============================] - 13s 65ms/step
Test loss: 0.5209842715600524
Test accuracy: 73.74%
----------------------------------------------------
guitar
        TRAIN
              precision    recall  f1-score   support

       False       1.00      0.03      0.06       387
        True       0.70      1.00      0.83       888

    accuracy                           0.71      1275
   macro avg       0.85      0.52      0.44      1275
weighted avg       0.79      0.71      0.59      1275

        TEST
              precision    recall  f1-score   support

       False       1.00      0.05      0.09        63
        True       0.69      1.00      0.82       135
```

```
    accuracy                           0.70       198
   macro avg       0.85      0.52      0.45       198
weighted avg       0.79      0.70      0.59       198

Train on 1362 samples, validate on 214 samples
Epoch 1/10
1362/1362 [==============================] - 231s 170ms/step - loss: 1.8103 -
acc: 0.5198 - val_loss: 0.7355 - val_acc: 0.5748
Epoch 2/10
1362/1362 [==============================] - 234s 172ms/step - loss: 1.0193 -
acc: 0.5191 - val_loss: 0.7009 - val_acc: 0.5981
Epoch 3/10
1362/1362 [==============================] - 234s 172ms/step - loss: 0.9239 -
acc: 0.5125 - val_loss: 0.6909 - val_acc: 0.5841
Epoch 4/10
1362/1362 [==============================] - 235s 173ms/step - loss: 0.8397 -
acc: 0.5653 - val_loss: 0.7065 - val_acc: 0.5888
Epoch 5/10
1362/1362 [==============================] - 236s 173ms/step - loss: 0.8218 -
acc: 0.5609 - val_loss: 0.6664 - val_acc: 0.6075
Epoch 6/10
1362/1362 [==============================] - 235s 173ms/step - loss: 0.8066 -
acc: 0.5492 - val_loss: 0.7398 - val_acc: 0.5794
Epoch 7/10
1362/1362 [==============================] - 234s 172ms/step - loss: 0.7791 -
acc: 0.5587 - val_loss: 0.6739 - val_acc: 0.5888
Epoch 8/10
1362/1362 [==============================] - 234s 172ms/step - loss: 0.7842 -
acc: 0.5565 - val_loss: 0.8913 - val_acc: 0.5654
Epoch 9/10
1362/1362 [==============================] - 238s 174ms/step - loss: 0.7377 -
acc: 0.5786 - val_loss: 0.7077 - val_acc: 0.5748
Epoch 10/10
1362/1362 [==============================] - 237s 174ms/step - loss: 0.7216 -
acc: 0.5910 - val_loss: 0.7086 - val_acc: 0.5748
```




```
226/226 [==============================] - 15s 65ms/step
Test loss: 0.6503403548118287
Test accuracy: 63.72%
----------------------------------------------------
mallet_percussion
        TRAIN
              precision    recall  f1-score   support

       False       0.66      0.79      0.72       810
        True       0.57      0.41      0.48       552

    accuracy                           0.64      1362
   macro avg       0.62      0.60      0.60      1362
weighted avg       0.63      0.64      0.62      1362

        TEST
              precision    recall  f1-score   support

       False       0.65      0.75      0.70       137
        True       0.50      0.38      0.43        89

    accuracy                           0.61       226
   macro avg       0.58      0.57      0.57       226
weighted avg       0.59      0.61      0.59       226

Train on 1842 samples, validate on 311 samples
Epoch 1/10
1842/1842 [==============================] - 317s 172ms/step - loss: 0.9336 -
acc: 0.5635 - val_loss: 0.7137 - val_acc: 0.5305
```

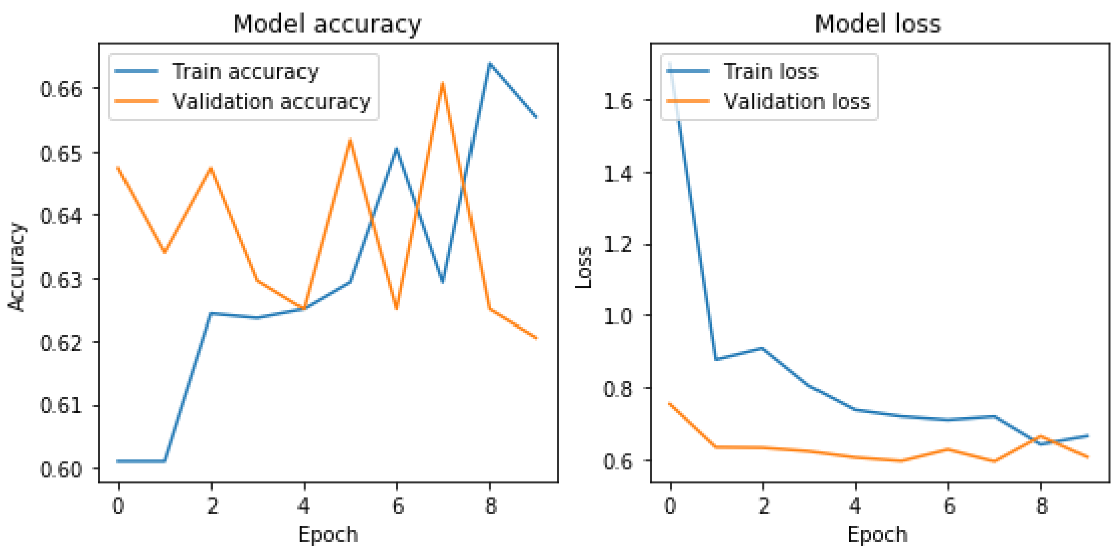
```
Epoch 2/10
1842/1842 [==============================] - 318s 173ms/step - loss: 0.7483 -
acc: 0.6097 - val_loss: 0.6588 - val_acc: 0.6045
Epoch 3/10
1842/1842 [==============================] - 321s 174ms/step - loss: 0.7091 -
acc: 0.6064 - val_loss: 0.7119 - val_acc: 0.5305
Epoch 4/10
1842/1842 [==============================] - 321s 174ms/step - loss: 0.6950 -
acc: 0.6162 - val_loss: 0.6463 - val_acc: 0.6463
Epoch 5/10
1842/1842 [==============================] - 315s 171ms/step - loss: 0.6742 -
acc: 0.6200 - val_loss: 0.5991 - val_acc: 0.6849
Epoch 6/10
1842/1842 [==============================] - 318s 172ms/step - loss: 0.6505 -
acc: 0.6374 - val_loss: 0.5973 - val_acc: 0.6688
Epoch 7/10
1842/1842 [==============================] - 316s 172ms/step - loss: 0.6499 -
acc: 0.6265 - val_loss: 0.6011 - val_acc: 0.6720
Epoch 8/10
1842/1842 [==============================] - 318s 173ms/step - loss: 0.6565 -
acc: 0.6493 - val_loss: 0.6026 - val_acc: 0.6752
Epoch 9/10
1842/1842 [==============================] - 321s 174ms/step - loss: 0.6531 -
acc: 0.6281 - val_loss: 0.6164 - val_acc: 0.6817
Epoch 10/10
1842/1842 [==============================] - 318s 173ms/step - loss: 0.6330 -
acc: 0.6460 - val_loss: 0.6057 - val_acc: 0.6688
```




```
311/311 [==============================] - 19s 61ms/step
```

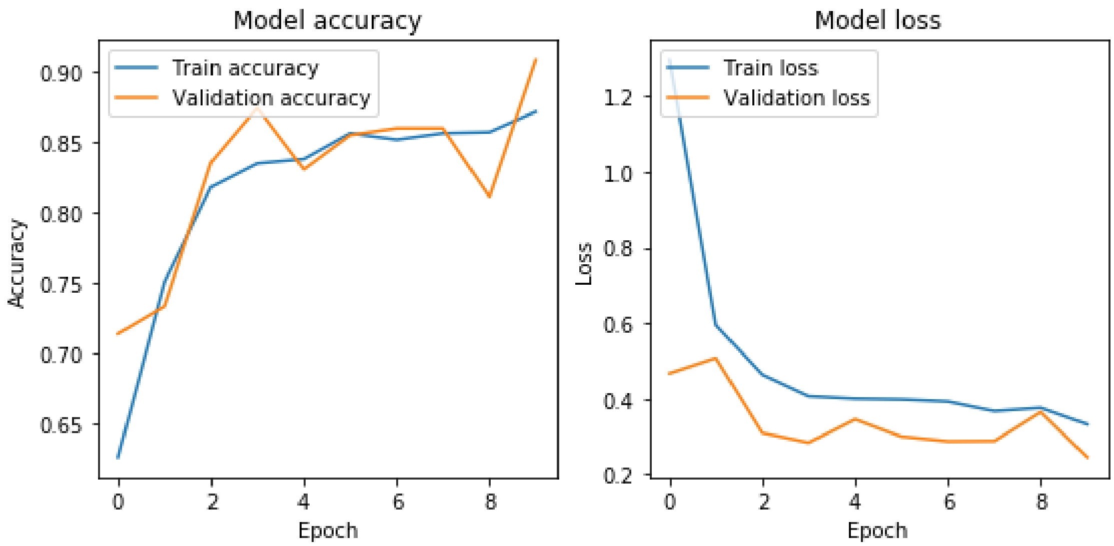
```
Test loss: 0.6496247246717717
Test accuracy: 60.13%
----------------------------------------------------
mandolin
        TRAIN
              precision    recall  f1-score   support

       False       0.78      0.65      0.71      1216
        True       0.48      0.63      0.55       626

    accuracy                           0.65      1842
   macro avg       0.63      0.64      0.63      1842
weighted avg       0.68      0.65      0.65      1842

        TEST
              precision    recall  f1-score   support

       False       0.67      0.62      0.65       192
        True       0.46      0.51      0.48       119

    accuracy                           0.58       311
   macro avg       0.57      0.57      0.57       311
weighted avg       0.59      0.58      0.59       311

Train on 1416 samples, validate on 224 samples
Epoch 1/10
1416/1416 [==============================] - 252s 178ms/step - loss: 1.7002 -
acc: 0.6010 - val_loss: 0.7526 - val_acc: 0.6473
Epoch 2/10
1416/1416 [==============================] - 246s 174ms/step - loss: 0.8764 -
acc: 0.6010 - val_loss: 0.6324 - val_acc: 0.6339
Epoch 3/10
1416/1416 [==============================] - 245s 173ms/step - loss: 0.9083 -
acc: 0.6243 - val_loss: 0.6312 - val_acc: 0.6473
Epoch 4/10
1416/1416 [==============================] - 246s 173ms/step - loss: 0.8037 -
acc: 0.6236 - val_loss: 0.6213 - val_acc: 0.6295
Epoch 5/10
1416/1416 [==============================] - 249s 176ms/step - loss: 0.7361 -
acc: 0.6250 - val_loss: 0.6040 - val_acc: 0.6250
Epoch 6/10
1416/1416 [==============================] - 244s 173ms/step - loss: 0.7180 -
acc: 0.6292 - val_loss: 0.5942 - val_acc: 0.6518
Epoch 7/10
1416/1416 [==============================] - 246s 174ms/step - loss: 0.7079 -
acc: 0.6504 - val_loss: 0.6264 - val_acc: 0.6250
Epoch 8/10
1416/1416 [==============================] - 246s 174ms/step - loss: 0.7177 -
```

```
acc: 0.6292 - val_loss: 0.5931 - val_acc: 0.6607
Epoch 9/10
1416/1416 [==============================] - 247s 174ms/step - loss: 0.6405 -
acc: 0.6638 - val_loss: 0.6633 - val_acc: 0.6250
Epoch 10/10
1416/1416 [==============================] - 246s 173ms/step - loss: 0.6638 -
acc: 0.6554 - val_loss: 0.6056 - val_acc: 0.6205
```



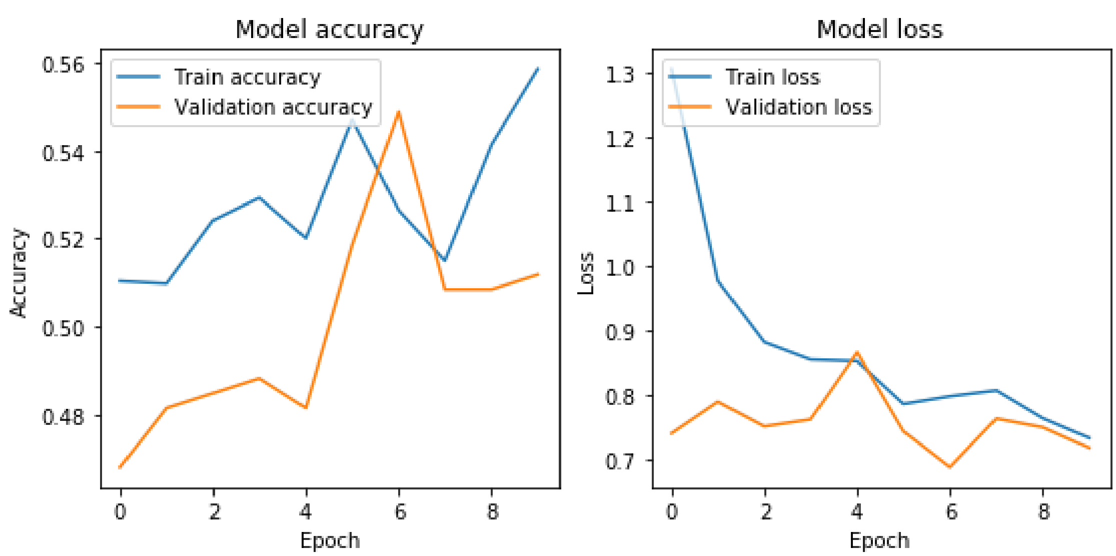

```
250/250 [==============================] - 16s 63ms/step
Test loss: 0.5691616973876953
Test accuracy: 70.40%
----------------------------------------------------
organ
        TRAIN
              precision    recall  f1-score   support

       False       0.92      0.44      0.60       968
        True       0.43      0.92      0.59       448

    accuracy                           0.59      1416
   macro avg       0.68      0.68      0.59      1416
weighted avg       0.77      0.59      0.59      1416

        TEST
              precision    recall  f1-score   support

       False       0.96      0.42      0.59       173
        True       0.43      0.96      0.59        77
```

```
    accuracy                           0.59       250
   macro avg       0.69      0.69      0.59       250
weighted avg       0.80      0.59      0.59       250

Train on 1310 samples, validate on 206 samples
Epoch 1/10
1310/1310 [==============================] - 224s 171ms/step - loss: 1.2964 -
acc: 0.6260 - val_loss: 0.4672 - val_acc: 0.7136
Epoch 2/10
1310/1310 [==============================] - 229s 175ms/step - loss: 0.5952 -
acc: 0.7504 - val_loss: 0.5076 - val_acc: 0.7330
Epoch 3/10
1310/1310 [==============================] - 222s 169ms/step - loss: 0.4637 -
acc: 0.8176 - val_loss: 0.3095 - val_acc: 0.8350
Epoch 4/10
1310/1310 [==============================] - 221s 169ms/step - loss: 0.4068 -
acc: 0.8344 - val_loss: 0.2838 - val_acc: 0.8738
Epoch 5/10
1310/1310 [==============================] - 211s 161ms/step - loss: 0.4003 -
acc: 0.8374 - val_loss: 0.3460 - val_acc: 0.8301
Epoch 6/10
1310/1310 [==============================] - 209s 159ms/step - loss: 0.3984 -
acc: 0.8557 - val_loss: 0.2999 - val_acc: 0.8544
Epoch 7/10
1310/1310 [==============================] - 211s 161ms/step - loss: 0.3929 -
acc: 0.8511 - val_loss: 0.2873 - val_acc: 0.8592
Epoch 8/10
1310/1310 [==============================] - 211s 161ms/step - loss: 0.3676 -
acc: 0.8557 - val_loss: 0.2881 - val_acc: 0.8592
Epoch 9/10
1310/1310 [==============================] - 209s 159ms/step - loss: 0.3763 -
acc: 0.8565 - val_loss: 0.3651 - val_acc: 0.8107
Epoch 10/10
1310/1310 [==============================] - 212s 162ms/step - loss: 0.3329 -
acc: 0.8710 - val_loss: 0.2453 - val_acc: 0.9078
```




```
204/204 [==============================] - 14s 71ms/step
Test loss: 0.26010071734587353
Test accuracy: 90.20%
----------------------------------------------------
piano
        TRAIN
              precision    recall  f1-score   support

       False       0.95      0.67      0.78       419
        True       0.86      0.98      0.92       891

    accuracy                           0.88      1310
   macro avg       0.90      0.83      0.85      1310
weighted avg       0.89      0.88      0.88      1310

        TEST
              precision    recall  f1-score   support

       False       0.92      0.67      0.77        66
        True       0.86      0.97      0.91       138

    accuracy                           0.87       204
   macro avg       0.89      0.82      0.84       204
weighted avg       0.88      0.87      0.87       204

Train on 1742 samples, validate on 297 samples
Epoch 1/10
1742/1742 [==============================] - 292s 168ms/step - loss: 1.3059 -
acc: 0.5103 - val_loss: 0.7396 - val_acc: 0.4680
```

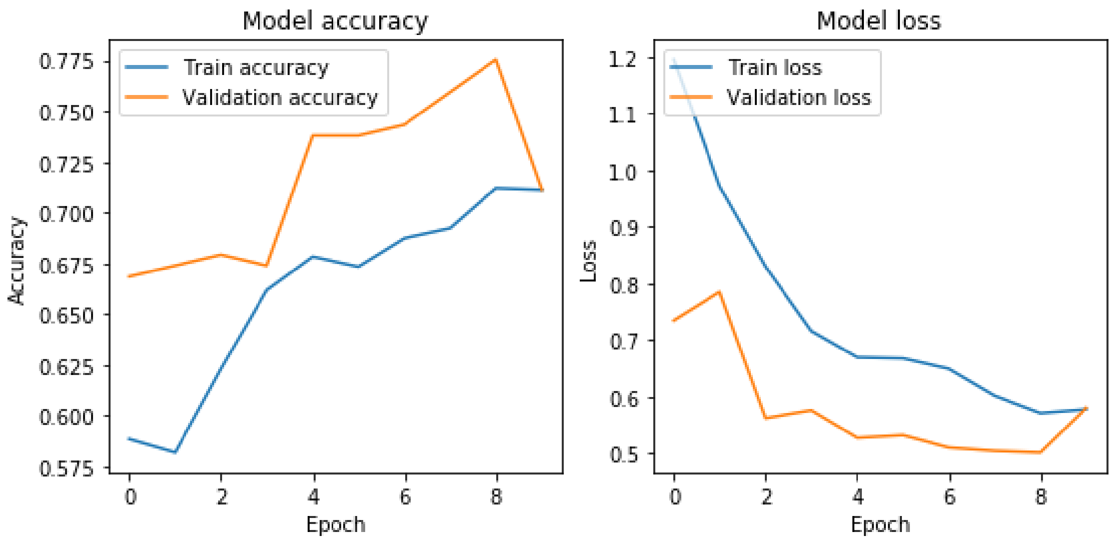
```
Epoch 2/10
1742/1742 [==============================] - 286s 164ms/step - loss: 0.9766 -
acc: 0.5098 - val_loss: 0.7886 - val_acc: 0.4815
Epoch 3/10
1742/1742 [==============================] - 289s 166ms/step - loss: 0.8812 -
acc: 0.5241 - val_loss: 0.7504 - val_acc: 0.4848
Epoch 4/10
1742/1742 [==============================] - 281s 162ms/step - loss: 0.8541 -
acc: 0.5293 - val_loss: 0.7612 - val_acc: 0.4882
Epoch 5/10
1742/1742 [==============================] - 279s 160ms/step - loss: 0.8517 -
acc: 0.5201 - val_loss: 0.8653 - val_acc: 0.4815
Epoch 6/10
1742/1742 [==============================] - 277s 159ms/step - loss: 0.7852 -
acc: 0.5471 - val_loss: 0.7427 - val_acc: 0.5185
Epoch 7/10
1742/1742 [==============================] - 279s 160ms/step - loss: 0.7971 -
acc: 0.5264 - val_loss: 0.6869 - val_acc: 0.5488
Epoch 8/10
1742/1742 [==============================] - 278s 160ms/step - loss: 0.8061 -
acc: 0.5149 - val_loss: 0.7625 - val_acc: 0.5084
Epoch 9/10
1742/1742 [==============================] - 274s 157ms/step - loss: 0.7630 -
acc: 0.5413 - val_loss: 0.7490 - val_acc: 0.5084
Epoch 10/10
1742/1742 [==============================] - 278s 160ms/step - loss: 0.7325 -
acc: 0.5586 - val_loss: 0.7169 - val_acc: 0.5118
```




```
326/326 [==============================] - 20s 63ms/step
```

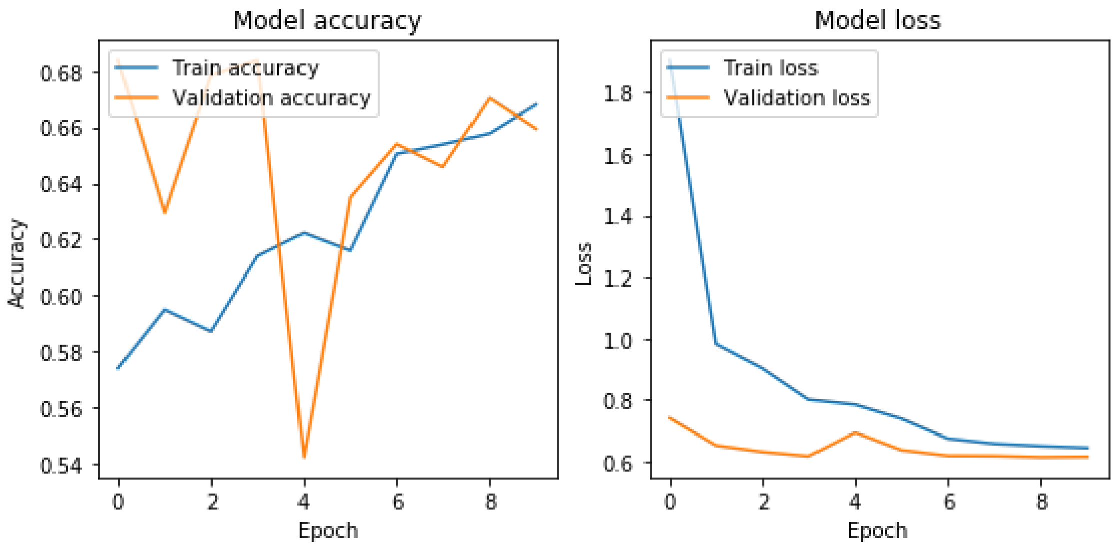
```
Test loss: 0.6907476643843153
Test accuracy: 55.83%
----------------------------------------------------
saxophone
        TRAIN
              precision    recall  f1-score   support

       False       0.65      0.35      0.46       914
        True       0.52      0.79      0.63       828

    accuracy                           0.56      1742
   macro avg       0.58      0.57      0.54      1742
weighted avg       0.59      0.56      0.54      1742

        TEST
              precision    recall  f1-score   support

       False       0.64      0.40      0.49       174
        True       0.52      0.74      0.61       152

    accuracy                           0.56       326
   macro avg       0.58      0.57      0.55       326
weighted avg       0.58      0.56      0.55       326

Train on 1212 samples, validate on 187 samples
Epoch 1/10
1212/1212 [==============================] - 193s 159ms/step - loss: 1.1954 -
acc: 0.5883 - val_loss: 0.7345 - val_acc: 0.6684
Epoch 2/10
1212/1212 [==============================] - 192s 158ms/step - loss: 0.9718 -
acc: 0.5817 - val_loss: 0.7849 - val_acc: 0.6738
Epoch 3/10
1212/1212 [==============================] - 192s 158ms/step - loss: 0.8296 -
acc: 0.6229 - val_loss: 0.5618 - val_acc: 0.6791
Epoch 4/10
1212/1212 [==============================] - 197s 163ms/step - loss: 0.7158 -
acc: 0.6617 - val_loss: 0.5758 - val_acc: 0.6738
Epoch 5/10
1212/1212 [==============================] - 192s 159ms/step - loss: 0.6693 -
acc: 0.6782 - val_loss: 0.5271 - val_acc: 0.7380
Epoch 6/10
1212/1212 [==============================] - 193s 160ms/step - loss: 0.6676 -
acc: 0.6733 - val_loss: 0.5318 - val_acc: 0.7380
Epoch 7/10
1212/1212 [==============================] - 195s 161ms/step - loss: 0.6492 -
acc: 0.6873 - val_loss: 0.5100 - val_acc: 0.7433
Epoch 8/10
1212/1212 [==============================] - 195s 161ms/step - loss: 0.6014 -
```

```
acc: 0.6922 - val_loss: 0.5042 - val_acc: 0.7594
Epoch 9/10
1212/1212 [==============================] - 194s 160ms/step - loss: 0.5708 -
acc: 0.7120 - val_loss: 0.5013 - val_acc: 0.7754
Epoch 10/10
1212/1212 [==============================] - 194s 160ms/step - loss: 0.5775 -
acc: 0.7112 - val_loss: 0.5801 - val_acc: 0.7112
```



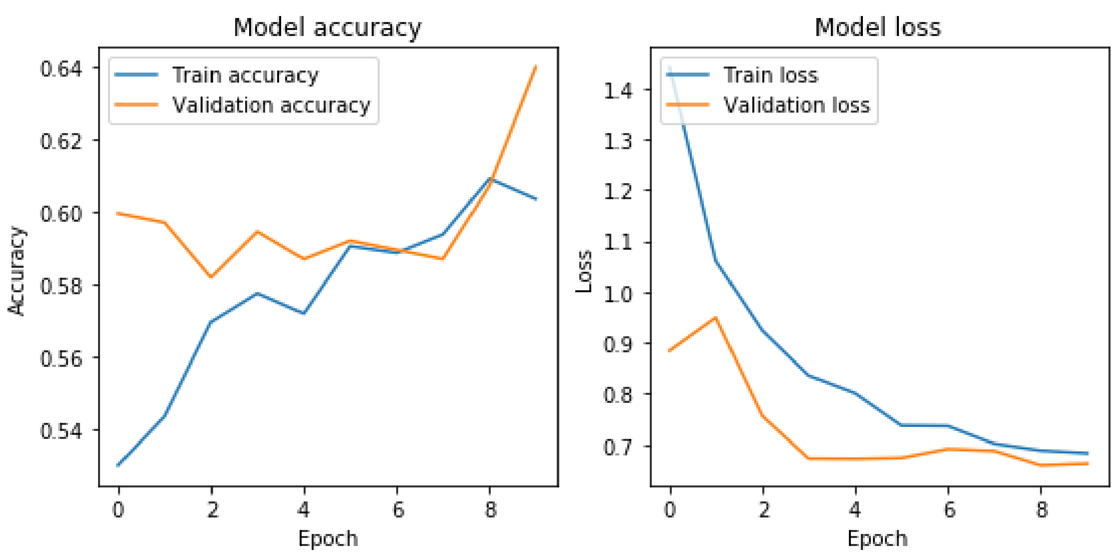

```
203/203 [==============================] - 12s 58ms/step
Test loss: 0.5563669988674483
Test accuracy: 70.94%
----------------------------------------------------
synthesizer
        TRAIN
              precision    recall  f1-score   support

       False       0.68      0.73      0.71       386
        True       0.87      0.84      0.85       826

    accuracy                           0.81      1212
   macro avg       0.78      0.79      0.78      1212
weighted avg       0.81      0.81      0.81      1212

        TEST
              precision    recall  f1-score   support

       False       0.61      0.67      0.64        64
        True       0.84      0.81      0.82       139
```

```
    accuracy                           0.76       203
   macro avg       0.73      0.74      0.73       203
weighted avg       0.77      0.76      0.77       203

Train on 2054 samples, validate on 367 samples
Epoch 1/10
2054/2054 [==============================] - 331s 161ms/step - loss: 1.9052 -
acc: 0.5740 - val_loss: 0.7412 - val_acc: 0.6839
Epoch 2/10
2054/2054 [==============================] - 330s 161ms/step - loss: 0.9819 -
acc: 0.5949 - val_loss: 0.6515 - val_acc: 0.6294
Epoch 3/10
2054/2054 [==============================] - 333s 162ms/step - loss: 0.9033 -
acc: 0.5871 - val_loss: 0.6305 - val_acc: 0.6785
Epoch 4/10
2054/2054 [==============================] - 331s 161ms/step - loss: 0.8008 -
acc: 0.6139 - val_loss: 0.6165 - val_acc: 0.6839
Epoch 5/10
2054/2054 [==============================] - 330s 161ms/step - loss: 0.7845 -
acc: 0.6222 - val_loss: 0.6934 - val_acc: 0.5422
Epoch 6/10
2054/2054 [==============================] - 331s 161ms/step - loss: 0.7389 -
acc: 0.6159 - val_loss: 0.6362 - val_acc: 0.6349
Epoch 7/10
2054/2054 [==============================] - 329s 160ms/step - loss: 0.6739 -
acc: 0.6504 - val_loss: 0.6179 - val_acc: 0.6540
Epoch 8/10
2054/2054 [==============================] - 329s 160ms/step - loss: 0.6571 -
acc: 0.6538 - val_loss: 0.6172 - val_acc: 0.6458
Epoch 9/10
2054/2054 [==============================] - 341s 166ms/step - loss: 0.6494 -
acc: 0.6577 - val_loss: 0.6133 - val_acc: 0.6703
Epoch 10/10
2054/2054 [==============================] - 432s 210ms/step - loss: 0.6436 -
acc: 0.6680 - val_loss: 0.6146 - val_acc: 0.6594
```




```
339/339 [==============================] - 23s 66ms/step
Test loss: 0.6200752998523655
Test accuracy: 66.96%
----------------------------------------------------
trombone
        TRAIN
              precision    recall  f1-score   support

       False       0.81      0.60      0.69      1415
        True       0.44      0.68      0.53       639

    accuracy                           0.63      2054
   macro avg       0.62      0.64      0.61      2054
weighted avg       0.69      0.63      0.64      2054

        TEST
              precision    recall  f1-score   support

       False       0.72      0.54      0.62       227
        True       0.38      0.58      0.46       112

    accuracy                           0.55       339
   macro avg       0.55      0.56      0.54       339
weighted avg       0.61      0.55      0.57       339

Train on 2156 samples, validate on 397 samples
Epoch 1/10
2156/2156 [==============================] - 425s 197ms/step - loss: 1.4412 -
acc: 0.5301 - val_loss: 0.8843 - val_acc: 0.5995
```

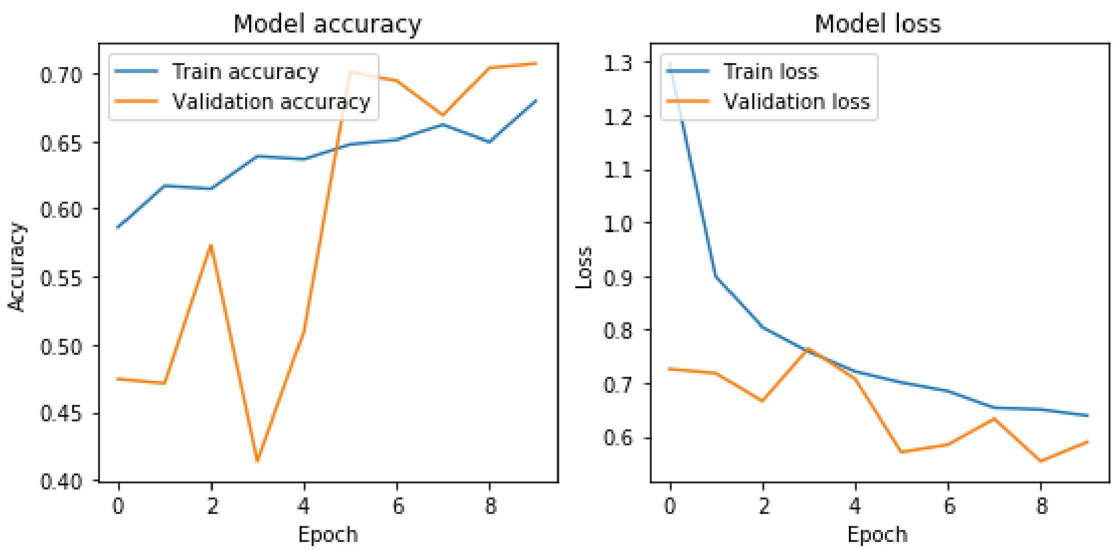
```
Epoch 2/10
2156/2156 [==============================] - 423s 196ms/step - loss: 1.0604 -
acc: 0.5436 - val_loss: 0.9492 - val_acc: 0.5970
Epoch 3/10
2156/2156 [==============================] - 451s 209ms/step - loss: 0.9239 -
acc: 0.5696 - val_loss: 0.7565 - val_acc: 0.5819
Epoch 4/10
2156/2156 [==============================] - 410s 190ms/step - loss: 0.8357 -
acc: 0.5775 - val_loss: 0.6728 - val_acc: 0.5945
Epoch 5/10
2156/2156 [==============================] - 343s 159ms/step - loss: 0.8010 -
acc: 0.5719 - val_loss: 0.6723 - val_acc: 0.5869
Epoch 6/10
2156/2156 [==============================] - 346s 161ms/step - loss: 0.7376 -
acc: 0.5904 - val_loss: 0.6739 - val_acc: 0.5919
Epoch 7/10
2156/2156 [==============================] - 363s 168ms/step - loss: 0.7367 -
acc: 0.5886 - val_loss: 0.6913 - val_acc: 0.5894
Epoch 8/10
2156/2156 [==============================] - 367s 170ms/step - loss: 0.7015 -
acc: 0.5937 - val_loss: 0.6875 - val_acc: 0.5869
Epoch 9/10
2156/2156 [==============================] - 370s 172ms/step - loss: 0.6882 -
acc: 0.6090 - val_loss: 0.6597 - val_acc: 0.6071
Epoch 10/10
2156/2156 [==============================] - 382s 177ms/step - loss: 0.6832 -
acc: 0.6034 - val_loss: 0.6629 - val_acc: 0.6398
```




```
363/363 [==============================] - 23s 63ms/step
```

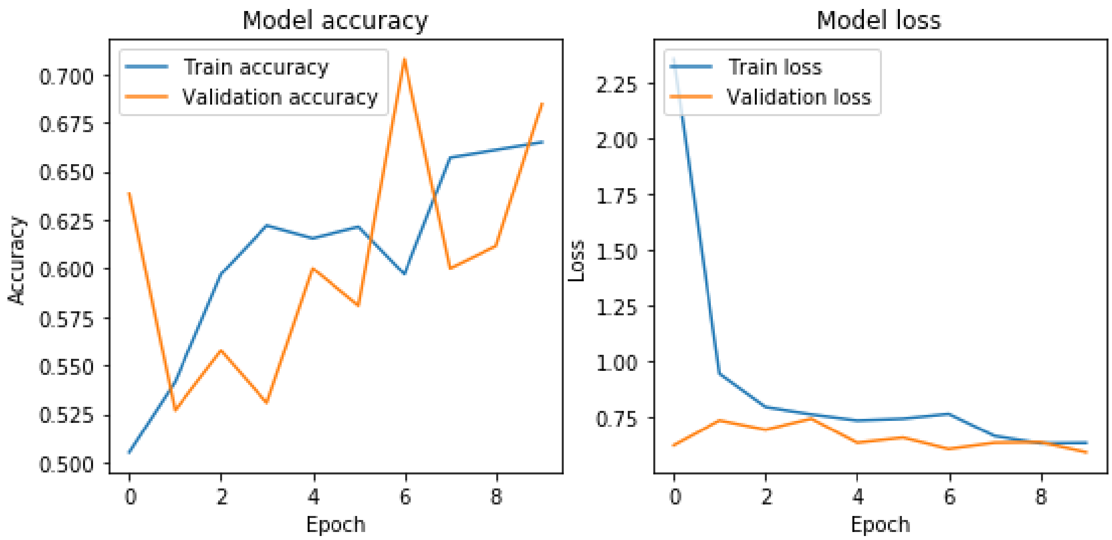
```
Test loss: 0.698793343604432
Test accuracy: 54.82%
----------------------------------------------------
trumpet
        TRAIN
              precision    recall  f1-score   support

       False       0.86      0.27      0.41      1324
        True       0.44      0.93      0.60       832

    accuracy                           0.52      2156
   macro avg       0.65      0.60      0.50      2156
weighted avg       0.70      0.52      0.48      2156

        TEST
              precision    recall  f1-score   support

       False       0.65      0.19      0.30       208
        True       0.44      0.86      0.58       155

    accuracy                           0.48       363
   macro avg       0.54      0.53      0.44       363
weighted avg       0.56      0.48      0.42       363

Train on 1827 samples, validate on 314 samples
Epoch 1/10
1827/1827 [==============================] - 332s 182ms/step - loss: 1.2970 -
acc: 0.5862 - val_loss: 0.7261 - val_acc: 0.4745
Epoch 2/10
1827/1827 [==============================] - 325s 178ms/step - loss: 0.8990 -
acc: 0.6169 - val_loss: 0.7184 - val_acc: 0.4713
Epoch 3/10
1827/1827 [==============================] - 321s 176ms/step - loss: 0.8046 -
acc: 0.6147 - val_loss: 0.6660 - val_acc: 0.5732
Epoch 4/10
1827/1827 [==============================] - 326s 179ms/step - loss: 0.7587 -
acc: 0.6388 - val_loss: 0.7648 - val_acc: 0.4140
Epoch 5/10
1827/1827 [==============================] - 323s 177ms/step - loss: 0.7216 -
acc: 0.6366 - val_loss: 0.7076 - val_acc: 0.5096
Epoch 6/10
1827/1827 [==============================] - 323s 177ms/step - loss: 0.7009 -
acc: 0.6475 - val_loss: 0.5712 - val_acc: 0.7006
Epoch 7/10
1827/1827 [==============================] - 321s 176ms/step - loss: 0.6846 -
acc: 0.6508 - val_loss: 0.5850 - val_acc: 0.6943
Epoch 8/10
1827/1827 [==============================] - 324s 177ms/step - loss: 0.6534 -
```

```
acc: 0.6617 - val_loss: 0.6333 - val_acc: 0.6688
Epoch 9/10
1827/1827 [==============================] - 322s 176ms/step - loss: 0.6502 -
acc: 0.6492 - val_loss: 0.5539 - val_acc: 0.7038
Epoch 10/10
1827/1827 [==============================] - 322s 176ms/step - loss: 0.6387 -
acc: 0.6793 - val_loss: 0.5899 - val_acc: 0.7070
```



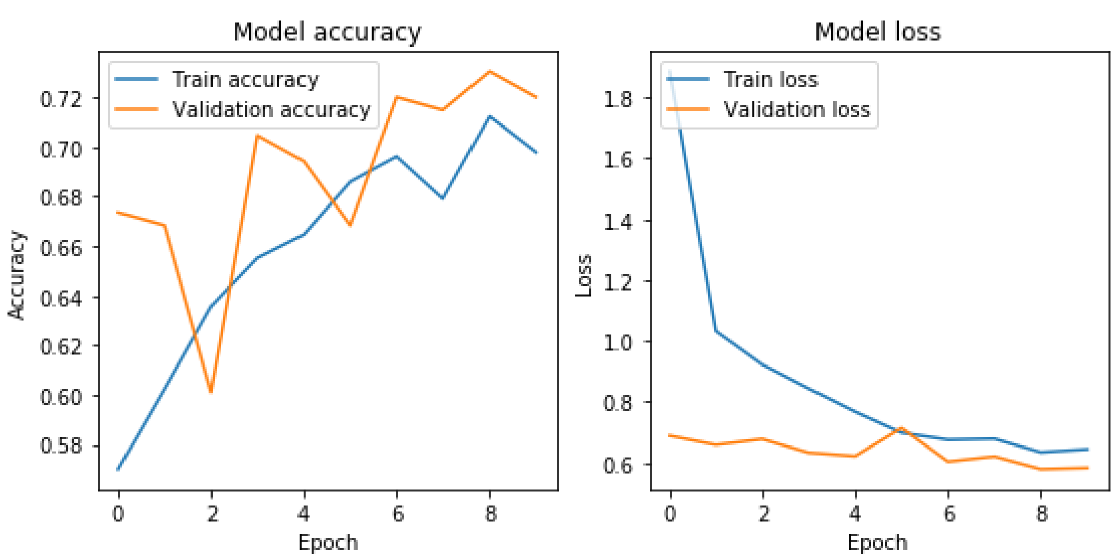

```
284/284 [==============================] - 18s 64ms/step
Test loss: 0.6137258754649633
Test accuracy: 65.85%
-----------------------------------------------------
ukulele
        TRAIN
              precision    recall  f1-score   support

       False       0.86      0.33      0.48      1260
        True       0.37      0.88      0.52       567

    accuracy                           0.50      1827
   macro avg       0.61      0.60      0.50      1827
weighted avg       0.71      0.50      0.49      1827

        TEST
              precision    recall  f1-score   support

       False       0.87      0.35      0.50       204
        True       0.34      0.86      0.49        80
```

```
    accuracy                           0.50       284
   macro avg       0.61      0.61      0.50       284
weighted avg       0.72      0.50      0.50       284

Train on 1516 samples, validate on 260 samples
Epoch 1/10
1516/1516 [==============================] - 269s 177ms/step - loss: 2.3592 -
acc: 0.5053 - val_loss: 0.6211 - val_acc: 0.6385
Epoch 2/10
1516/1516 [==============================] - 268s 177ms/step - loss: 0.9432 -
acc: 0.5416 - val_loss: 0.7335 - val_acc: 0.5269
Epoch 3/10
1516/1516 [==============================] - 272s 179ms/step - loss: 0.7936 -
acc: 0.5970 - val_loss: 0.6909 - val_acc: 0.5577
Epoch 4/10
1516/1516 [==============================] - 268s 177ms/step - loss: 0.7606 -
acc: 0.6220 - val_loss: 0.7413 - val_acc: 0.5308
Epoch 5/10
1516/1516 [==============================] - 276s 182ms/step - loss: 0.7329 -
acc: 0.6154 - val_loss: 0.6322 - val_acc: 0.6000
Epoch 6/10
1516/1516 [==============================] - 269s 177ms/step - loss: 0.7412 -
acc: 0.6214 - val_loss: 0.6556 - val_acc: 0.5808
Epoch 7/10
1516/1516 [==============================] - 271s 179ms/step - loss: 0.7628 -
acc: 0.5970 - val_loss: 0.6045 - val_acc: 0.7077
Epoch 8/10
1516/1516 [==============================] - 271s 179ms/step - loss: 0.6628 -
acc: 0.6570 - val_loss: 0.6324 - val_acc: 0.6000
Epoch 9/10
1516/1516 [==============================] - 267s 176ms/step - loss: 0.6298 -
acc: 0.6609 - val_loss: 0.6350 - val_acc: 0.6115
Epoch 10/10
1516/1516 [==============================] - 265s 175ms/step - loss: 0.6315 -
acc: 0.6649 - val_loss: 0.5894 - val_acc: 0.6846
```




```
257/257 [==============================] - 17s 66ms/step
Test loss: 0.548202626436137
Test accuracy: 72.37%
----------------------------------------------------
violin
        TRAIN
              precision    recall  f1-score   support

       False       0.92      0.46      0.61       640
        True       0.71      0.97      0.82       876

    accuracy                           0.76      1516
   macro avg       0.82      0.72      0.72      1516
weighted avg       0.80      0.76      0.73      1516

        TEST
              precision    recall  f1-score   support

       False       0.93      0.39      0.55       106
        True       0.69      0.98      0.81       151

    accuracy                           0.74       257
   macro avg       0.81      0.68      0.68       257
weighted avg       0.79      0.74      0.70       257

Train on 1172 samples, validate on 193 samples
Epoch 1/10
1172/1172 [==============================] - 196s 167ms/step - loss: 1.8857 -
acc: 0.5700 - val_loss: 0.6892 - val_acc: 0.6736
```

```
Epoch 2/10
1172/1172 [==============================] - 194s 166ms/step - loss: 1.0327 -
acc: 0.6024 - val_loss: 0.6594 - val_acc: 0.6684
Epoch 3/10
1172/1172 [==============================] - 191s 163ms/step - loss: 0.9228 -
acc: 0.6357 - val_loss: 0.6784 - val_acc: 0.6010
Epoch 4/10
1172/1172 [==============================] - 194s 165ms/step - loss: 0.8438 -
acc: 0.6553 - val_loss: 0.6328 - val_acc: 0.7047
Epoch 5/10
1172/1172 [==============================] - 205s 175ms/step - loss: 0.7678 -
acc: 0.6647 - val_loss: 0.6213 - val_acc: 0.6943
Epoch 6/10
1172/1172 [==============================] - 211s 180ms/step - loss: 0.6990 -
acc: 0.6860 - val_loss: 0.7145 - val_acc: 0.6684
Epoch 7/10
1172/1172 [==============================] - 208s 178ms/step - loss: 0.6762 -
acc: 0.6962 - val_loss: 0.6035 - val_acc: 0.7202
Epoch 8/10
1172/1172 [==============================] - 204s 174ms/step - loss: 0.6792 -
acc: 0.6792 - val_loss: 0.6201 - val_acc: 0.7150
Epoch 9/10
1172/1172 [==============================] - 210s 179ms/step - loss: 0.6339 -
acc: 0.7125 - val_loss: 0.5788 - val_acc: 0.7306
Epoch 10/10
1172/1172 [==============================] - 208s 178ms/step - loss: 0.6437 -
acc: 0.6980 - val_loss: 0.5828 - val_acc: 0.7202
```




```
199/199 [==============================] - 14s 69ms/step
```

```
Test loss: 0.5129846924513428
Test accuracy: 77.89%
----------------------------------------------------
voice
        TRAIN
              precision    recall  f1-score   support

       False       0.91      0.40      0.55       437
        True       0.73      0.98      0.84       735

    accuracy                           0.76      1172
   macro avg       0.82      0.69      0.70      1172
weighted avg       0.80      0.76      0.73      1172

        TEST
              precision    recall  f1-score   support

       False       0.86      0.37      0.52        65
        True       0.76      0.97      0.85       134

    accuracy                           0.77       199
   macro avg       0.81      0.67      0.68       199
weighted avg       0.79      0.77      0.74       199
```

[8]:
```
import matplotlib.pyplot as plt
from pylab import plot, show, figure, imshow, xlim, ylim, title


def plot_history():
    plt.figure(figsize=(9,4))
    plt.subplot(1,2,1)
    plt.plot(history.history['acc'])
    plt.plot(history.history['val_acc'])
    plt.title('Model accuracy')
    plt.ylabel('Accuracy')
    plt.xlabel('Epoch')
    plt.legend(['Train accuracy', 'Validation accuracy'], loc='upper left')
    plt.subplot(1,2,2)
    plt.plot(history.history['loss'])
    plt.plot(history.history['val_loss'])
    plt.title('Model loss')
    plt.ylabel('Loss')
    plt.xlabel('Epoch')
    plt.legend(['Train loss', 'Validation loss'], loc='upper left')
    plt.show()
```

[ ]:
```
"""
    # Step 3: simplify the data by averaging over time
    # 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)
    X_test_inst_sklearn = np.mean(X_test_inst, axis=1)
    X_train_inst_sklearn = X_train_inst_sklearn.astype('float32')
    X_train_inst_sklearn = lb.util.normalize(X_train_inst_sklearn)
"""

np.savez('models.npz',model=)
```