

best-4-3

September 21, 2021

```
[1]: base_transfer_set = ['01', '02', '04', '05', '08', '09', '12', '13', '16',  
    ↪ '17', '18', '20']  
target_transfer_set = ['03', '06', '07', '10', '11', '14', '15', '19']  
  
import random  
def random_combination(iterable, r):  
    "Random selection from itertools.combinations(iterable, r)"  
    pool = tuple(iterable)  
    n = len(pool)  
    indices = sorted(random.sample(range(n), r))  
    return tuple(pool[i] for i in indices)  
  
transfers_size_6 = []  
for i in range(4):  
    transfers_size_6.append(random_combination(target_transfer_set, 6))  
print(transfers_size_6)  
transfers_size_6 = [('03', '06', '07', '10', '11', '14'), ('03', '06', '07',  
    ↪ '10', '14', '15'), ('03', '06', '07', '10', '14', '15'), ('03', '07', '10',  
    ↪ '14', '15', '19')]  
for i, tmp in enumerate(transfers_size_6):  
    transfers_size_6[i] = list(transfers_size_6[i])  
print(transfers_size_6)  
  
transfers_size_4 = []  
for i in range(4):  
    transfers_size_4.append(random_combination(target_transfer_set, 4))  
print(transfers_size_4)  
transfers_size_4 = [('06', '10', '14', '15'), ('03', '10', '14', '19'), ('03',  
    ↪ '06', '10', '15'), ('03', '07', '10', '15')]  
for i, tmp in enumerate(transfers_size_4):  
    transfers_size_4[i] = list(transfers_size_4[i])  
print(transfers_size_4)  
  
transfers_size_3 = []  
for i in range(4):  
    transfers_size_3.append(random_combination(target_transfer_set, 3))  
print(transfers_size_3)
```

```

transfers_size_3 = [('07', '11', '14'), ('06', '07', '10'), ('03', '15', '19'),
                    ↪('06', '14', '19')]
for i, tmp in enumerate(transfers_size_3):
    transfers_size_3[i] = list(transfers_size_3[i])
print(transfers_size_3)

transfers_size_2 = []
for i in range(4):
    transfers_size_2.append(random_combination(target_transfer_set, 2))
print(transfers_size_2)
transfers_size_2 = [('06', '10'), ('07', '11'), ('06', '15'), ('14', '15')]
for i, tmp in enumerate(transfers_size_2):
    transfers_size_2[i] = list(transfers_size_2[i])
print(transfers_size_2)

```

```

[('03', '06', '07', '10', '14', '19'), ('03', '07', '10', '11', '15', '19'),
('06', '07', '10', '11', '14', '19'), ('03', '06', '07', '11', '14', '19')]
[['03', '06', '07', '10', '11', '14'], ['03', '06', '07', '10', '14', '15'],
['03', '06', '07', '10', '14', '15'], ['03', '07', '10', '14', '15', '19']]
[('03', '06', '11', '14'), ('06', '07', '15', '19'), ('10', '14', '15', '19'),
('06', '11', '15', '19')]
[['06', '10', '14', '15'], ['03', '10', '14', '19'], ['03', '06', '10', '15'],
['03', '07', '10', '15']]
[('10', '11', '15'), ('06', '14', '19'), ('03', '14', '19'), ('06', '14', '15')]
[['07', '11', '14'], ['06', '07', '10'], ['03', '15', '19'], ['06', '14', '19']]
[('10', '11'), ('15', '19'), ('11', '15'), ('06', '10')]
[['06', '10'], ['07', '11'], ['06', '15'], ['14', '15']]

```

```

[2]: import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns

```

```

def make_confusion_matrix(cf,
                          group_names=None,
                          categories='auto',
                          count=True,
                          percent=True,
                          cbar=True,
                          xyticks=True,
                          xyplotlabels=True,
                          sum_stats=True,
                          figsize=None,
                          cmap='Blues',
                          title=None):
    """

```

This function will make a pretty plot of an sklearn Confusion Matrix cm_
↪using a Seaborn heatmap visualization.

Arguments

```
-----  
cf:          confusion matrix to be passed in  
group_names: List of strings that represent the labels row by row to be  
→shown in each square.  
categories:  List of strings containing the categories to be displayed on  
→the x,y axis. Default is 'auto'  
count:       If True, show the raw number in the confusion matrix.  
→Default is True.  
normalize:   If True, show the proportions for each category. Default is  
→True.  
cbar:        If True, show the color bar. The cbar values are based off  
→the values in the confusion matrix.  
             Default is True.  
xyticks:     If True, show x and y ticks. Default is True.  
xyplotlabels: If True, show 'True Label' and 'Predicted Label' on the  
→figure. Default is True.  
sum_stats:   If True, display summary statistics below the figure.  
→Default is True.  
figsize:     Tuple representing the figure size. Default will be the  
→matplotlib rcParams value.  
cmap:        Colormap of the values displayed from matplotlib.pyplot.cm.  
→Default is 'Blues'  
             See http://matplotlib.org/examples/color/colormaps\_reference.html  
→html  
  
title:       Title for the heatmap. Default is None.  
'''  
  
# CODE TO GENERATE TEXT INSIDE EACH SQUARE  
blanks = ['' for i in range(cf.size)]  
  
if group_names and len(group_names)==cf.size:  
    group_labels = ["{}\n".format(value) for value in group_names]  
else:  
    group_labels = blanks  
  
if count:  
    group_counts = ["{0:0.0f}\n".format(value) for value in cf.flatten()]  
else:  
    group_counts = blanks  
  
if percent:  
    group_percentages = ["{0:.2%}".format(value) for value in cf.flatten()/  
→np.sum(cf)]
```

```

else:
    group_percentages = blanks

    box_labels = [f"{v1}-{v2}-{v3}".strip() for v1, v2, v3 in
→zip(group_labels,group_counts,group_percentages)]
    box_labels = np.asarray(box_labels).reshape(cf.shape[0],cf.shape[1])

# CODE TO GENERATE SUMMARY STATISTICS & TEXT FOR SUMMARY STATS
if sum_stats:
    #Accuracy is sum of diagonal divided by total observations
    accuracy = np.trace(cf) / float(np.sum(cf))

    #if it is a binary confusion matrix, show some more stats
    if len(cf)==2:
        #Metrics for Binary Confusion Matrices
        precision = cf[1,1] / sum(cf[:,1])
        recall    = cf[1,1] / sum(cf[1,:])
        f1_score  = 2*precision*recall / (precision + recall)
        stats_text = "\n\nAccuracy={:0.3f}\nPrecision={:0.3f}\nRecall={:0.
→3f}\nF1 Score={:0.3f}".format(
            accuracy,precision,recall,f1_score)
    else:
        stats_text = "\n\nAccuracy={:0.3f}".format(accuracy)
else:
    stats_text = ""

# SET FIGURE PARAMETERS ACCORDING TO OTHER ARGUMENTS
if figsize==None:
    #Get default figure size if not set
    figsize = plt.rcParams.get('figure.figsize')

if xyticks==False:
    #Do not show categories if xyticks is False
    categories=False

# MAKE THE HEATMAP VISUALIZATION
plt.figure(figsize=figsize)
sns.
→heatmap(cf,annot=box_labels,fmt="",cmap=cmap,cbar=cbar,xticklabels=categories,yticklabels=c

if xyplotlabels:
    plt.ylabel('True label')
    plt.xlabel('Predicted label' + stats_text)
else:

```

```
plt.xlabel(stats_text)

if title:
    plt.title(title)
```

```
-----
RuntimeError                                Traceback (most recent call last)
RuntimeError: module compiled against API version 0xe but this version of numpy
↳ is 0xd
```

```
-----
RuntimeError                                Traceback (most recent call last)
RuntimeError: module compiled against API version 0xe but this version of numpy
↳ is 0xd
```

```
[3]: import os
import pandas as pd
import warnings
warnings.filterwarnings("ignore")

def create_best_model(gesture_subset):
    gesture_subset.sort()
    print("Loadind Dataset for gestures: ", gesture_subset)
    path = 'gestures-dataset'
    dataset = None

    samples = 0
    for subject in os.listdir(path):
        if os.path.isfile(os.path.join(path, subject)):
            continue
        if subject in ('U01', 'U02', 'U03', 'U04', 'U05', 'U06', 'U07', 'U08'):
            for gesture in os.listdir(os.path.join(path, subject)):
                if os.path.isfile(os.path.join(path, subject, gesture)):
                    continue
                gesture = str(gesture)
                if gesture not in gesture_subset:
                    continue
                for samplefile in os.listdir(os.path.join(path, subject,
↳gesture)):
                    if os.path.isfile(os.path.join(path, subject, gesture,
↳samplefile)):
                        df = pd.read_csv(os.path.join(path, subject, gesture,
↳samplefile), \
                                        sep = ' ', \
                                        names = ['System.currentTimeMillis()', \
```

```

        'System.nanoTime()', \
        'sample.timestamp', \
        'X', \
        'Y', \
        'Z' \
    ])
df = df[["sample.timestamp", "X", "Y", "Z"]]

start = df["sample.timestamp"][0]
df["sample.timestamp"] -= start
df["sample.timestamp"] /= 10000000
df["subject"] = subject
df["gesture"] = gesture
df["sample"] = str(samplefile[:-4])
samples += 1
#print(df)
if dataset is None:
    dataset = df.copy()
else:
    dataset = pd.concat([dataset, df])

dataset = dataset.sort_values(by=['gesture', 'subject', 'sample', 'sample.
→timestamp'])
data = dataset
print(str(samples) + " samples loaded")

print("Scaling Dataset for gestures: ", gesture_subset)
from sklearn.preprocessing import StandardScaler

scaler = StandardScaler()
dataset_scaled = None

samples = 0
for i, gesture in enumerate(gesture_subset):
    df_gesture=data[data['gesture']==gesture]
    for j, subject in enumerate(df_gesture['subject'].unique()):
        df_subject=df_gesture[df_gesture['subject']==subject]
        for k, sample in enumerate(df_subject['sample'].unique()):
            df_sample=df_subject[df_subject['sample']==sample].copy()
            df_sample.sort_values(by=['sample.timestamp'])

            sc = scaler
            sc = sc.fit_transform(df_sample[["X", "Y", "Z"]])
            sc = pd.DataFrame(data=sc, columns=["X", "Y", "Z"])
            df_sample['X'] = sc['X']
            df_sample['Y'] = sc['Y']
            df_sample['Z'] = sc['Z']

```

```

        if dataset_scaled is None:
            dataset_scaled = df_sample.copy()
        else:
            dataset_scaled = pd.concat([dataset_scaled, df_sample])
        samples += 1
    print(str(samples) + " samples scaled")
    data = dataset_scaled

    print("Cleaning Dataset for gestures: ", gesture_subset)
    dataset_outliers = None
    dataset_cleaned = None

    samples = 0
    outliers = 0
    for i, gesture in enumerate(gesture_subset):
        df_gesture = data[data['gesture']==gesture]
        for j, subject in enumerate(df_gesture['subject'].unique()):
            df_subject = df_gesture[df_gesture['subject']==subject]

            time_mean = df_subject.groupby(["gesture", "subject", "sample"]).
→count().groupby(["gesture", "subject"]).agg({'sample.timestamp': ['mean']})
            time_std = df_subject.groupby(["gesture", "subject", "sample"]).
→count().groupby(["gesture", "subject"]).agg({'sample.timestamp': ['std']})
            time_max = time_mean['sample.timestamp'].iloc[0]['mean'] + 1.0 *
→time_std['sample.timestamp'].iloc[0]['std']
            time_min = time_mean['sample.timestamp'].iloc[0]['mean'] - 1.0 *
→time_std['sample.timestamp'].iloc[0]['std']
            for k, sample in enumerate(df_subject['sample'].unique()):
                df_sample=df_subject[df_subject['sample']==sample]
                df_sample_count = df_sample.count()['sample.timestamp']
                if df_sample_count < time_min or df_sample_count > time_max:
                    if dataset_outliers is None:
                        dataset_outliers = df_sample.copy()
                    else:
                        dataset_outliers = pd.concat([dataset_outliers,
→df_sample])
                outliers += 1
            else:
                if dataset_cleaned is None:
                    dataset_cleaned = df_sample.copy()
                else:
                    dataset_cleaned = pd.concat([dataset_cleaned,
→df_sample])
            samples += 1
    print(str(samples) + " samples cleaned")
    print(str(outliers) + " samples outliers")

```

```

data = dataset_cleaned

print("Time slicing Cleaned Dataset for gestures: ", gesture_subset)
dataset_timecut = None
samples = 0
damaged = 0
for i, gesture in enumerate(data['gesture'].unique()):
    df_gesture = data[data['gesture']==gesture]
    for j, subject in enumerate(df_gesture['subject'].unique()):
        df_subject = df_gesture[df_gesture['subject']==subject]
        time_max = 19 # 18 * 11 = 198
        for i, sample in enumerate(df_subject['sample'].unique()):
            df_sample = df_subject[df_subject['sample']==sample]
            df_sample_count = df_sample.count()['sample.timestamp']
            #print(df_sample_count)
            if df_sample_count >= time_max:
                df_sample = df_sample[df_sample['sample.timestamp'] <= (11_
↳* (time_max-1))]
                df_sample_count = df_sample.count()['sample.timestamp']
                #print(df_sample_count)
            elif df_sample_count < time_max:
                for tmp in range(df_sample_count * 11, (time_max) * 11, 11):
                    df = pd.DataFrame([[tmp, 0.0, 0.0, 0.0, gesture,
↳subject, sample]], columns=['sample.timestamp', 'X', 'Y', 'Z', 'gesture',
↳'subject', 'sample'])
                    df_sample = df_sample.append(df, ignore_index=True)
                    #print(df_sample)
                    df_sample_count = df_sample.count()['sample.timestamp']
                    #print(df_sample_count)
                    if df_sample_count != time_max:
                        damaged += 1
                        continue
                    if dataset_timecut is None:
                        dataset_timecut = df_sample.copy()
                    else:
                        dataset_timecut = pd.concat([dataset_timecut, df_sample])
                    samples += 1

dataset_cleaned = dataset_timecut
print(str(samples) + " cleaned samples sliced")
print(str(damaged) + " cleaned samples damaged")

data = dataset_outliers
print("Time slicing Outliers Dataset for gestures: ", gesture_subset)
dataset_timecut = None
samples = 0
damaged = 0

```



```

for i, gesture in enumerate(data['gesture'].unique()):
    df_gesture = data[data['gesture']==gesture]
    for j, subject in enumerate(df_gesture['subject'].unique()):
        df_subject = df_gesture[df_gesture['subject']==subject]
        time_max = 19 # 18 * 11 = 198
        for i, sample in enumerate(df_subject['sample'].unique()):
            df_sample = df_subject[df_subject['sample']==sample]
            df_sample_count = df_sample.count()['sample.timestamp']
            #print(df_sample_count)
            if df_sample_count >= time_max:
                df_sample = df_sample[df_sample['sample.timestamp'] <= (11
↳* (time_max-1))]
                df_sample_count = df_sample.count()['sample.timestamp']
                #print(df_sample_count)
            elif df_sample_count < time_max:
                for tmp in range(df_sample_count * 11, (time_max) * 11, 11):
                    df = pd.DataFrame([[tmp, 0.0, 0.0, 0.0, gesture,
↳subject, sample]], columns=['sample.timestamp', 'X', 'Y', 'Z', 'gesture',
↳'subject', 'sample'])
                    df_sample = df_sample.append(df, ignore_index=True)
                    #print(df_sample)
                    df_sample_count = df_sample.count()['sample.timestamp']
                    #print(df_sample_count)
                    if df_sample_count != time_max:
                        damaged += 1
                        continue
                    if dataset_timecut is None:
                        dataset_timecut = df_sample.copy()
                    else:
                        dataset_timecut = pd.concat([dataset_timecut, df_sample])
                    samples += 1

dataset_outliers = dataset_timecut
print(str(samples) + " outliers samples sliced")
print(str(damaged) + " outliers samples damaged")

data = dataset_cleaned

from keras.models import Sequential
from keras.layers import Bidirectional
from keras.layers import LSTM
from keras.layers import Dense
from keras.layers import Dropout
from keras.optimizers import adam_v2
from keras.wrappers.scikit_learn import KerasClassifier
# from scikeras.wrappers import KerasClassifier
from sklearn.model_selection import StratifiedGroupKFold

```

```

from sklearn.model_selection import cross_val_score
from sklearn.model_selection import GridSearchCV
from keras.utils import np_utils
from sklearn.preprocessing import LabelEncoder
from sklearn.pipeline import Pipeline
import numpy as np

# fix random seed for reproducibility
seed = 1000
np.random.seed(seed)
# create the dataset
def get_dataset(data):
    X_train = []
    Y_train = []
    groups = []
    for i, gesture in enumerate(data['gesture'].unique()):
        df_gesture = data[data['gesture']==gesture]
        for j, subject in enumerate(df_gesture['subject'].unique()):
            df_subject = df_gesture[df_gesture['subject']==subject]
            for k, sample in enumerate(df_subject['sample'].unique()):
                df_sample = df_subject[df_subject['sample']==sample]
                accel_vector = []
                for index, row in df_sample.sort_values(by='sample.
→timestamp').iterrows():
                    accel_vector.append([row['X'],row['Y'],row['Z']])
                accel_vector = np.asarray(accel_vector)
                X_train.append(accel_vector)
                Y_train.append(gesture)
                groups.append(subject)
    X_train = np.asarray(X_train)
    Y_train = LabelEncoder().fit_transform(Y_train)
    #print(Y_train)
    return X_train, Y_train, groups

# Function to create model, required for KerasClassifier
def create_model(dropout_rate=0.8, units=128, optimizer=adam_v2.
→Adam(learning_rate=0.001)):
    model = Sequential()
    model.add(
        Bidirectional(
            LSTM(
                units=units,
                input_shape=[19, 3]
            )
        )
    )
    model.add(Dropout(rate=dropout_rate))

```

```

        model.add(Dense(units=units, activation='relu'))
        model.add(Dense(len(gesture_subset), activation='softmax'))
        model.compile(loss='sparse_categorical_crossentropy',
→optimizer=optimizer, metrics=['accuracy'])
        #print(model.summary())
        return model

model = KerasClassifier(build_fn=create_model, verbose=0)
cv = StratifiedGroupKFold(n_splits=5, shuffle=True, random_state=1000)
# get the dataset
X, y, g = get_dataset(dataset_cleaned)
#cv = cv.split(X, y, g)
batch_size = [19]
epochs = [64, 128]
#epochs = [128]
units = [16, 32, 64, 128]
# units = [16]
dropout_rate = [0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9]
# dropout_rate = [0.5]
param_grid = dict(epochs=epochs, units=units, batch_size=batch_size,
→dropout_rate=dropout_rate)
print("Hyperparameter tuning started for Dataset for gestures: ",
→gesture_subset)
grid = GridSearchCV(estimator=model, param_grid=param_grid, n_jobs=1,
→cv=cv, verbose=1)
grid_result = grid.fit(X, y, groups=g)
# summarize results
print("Best: %f using %s" % (grid_result.best_score_, grid_result.
→best_params_))
means = grid_result.cv_results_['mean_test_score']
stds = grid_result.cv_results_['std_test_score']
train_mean = grid_result.cv_results_['mean_fit_time']
train_std = grid_result.cv_results_['std_fit_time']
score_mean = grid_result.cv_results_['mean_score_time']
score_std = grid_result.cv_results_['std_score_time']
params = grid_result.cv_results_['params']
for mean, stdev, train_mean, train_std, score_mean, score_std, param in
→zip(means, stds, train_mean, train_std, score_mean, score_std, params):
    print("accuracy: %f (%f) train time: %f (%f) score time: %f (%f) with:
→%r" % (mean, stdev, train_mean, train_std, score_mean, score_std, param))
    print("Hyperparameter tuning completed for Dataset: ", gesture_subset)

model = grid_result.best_estimator_
import pickle

```

```

def save_model(model, gesture_subset):
    gesture_subset.sort()
    name = '-'.join(gesture_subset)
    # saving model
    pickle.dump(model.classes_, open(name + '_model_classes.pkl', 'wb'))
    model.model.save(name + '_lstm')
    print("Saving model to disk started for Dataset gestures: ", gesture_subset)
    save_model(model, gesture_subset)
    print("Saving model to disk completed for Dataset gestures: ",
    ↪gesture_subset)

import tensorflow as tf
def load_model(gesture_subset):
    gesture_subset.sort()
    name = '-'.join(gesture_subset)
    # loading model
    build_model = lambda: tf.keras.models.load_model(name + '_lstm')
    classifier = KerasClassifier(build_fn=build_model, epochs=1,
    ↪batch_size=10, verbose=0)
    classifier.classes_ = pickle.load(open(name + '_model_classes.
    ↪pkl', 'rb'))
    classifier.model = build_model()
    return classifier
    print("Loading model to disk started for Dataset gestures: ",
    ↪gesture_subset)
    model = load_model(gesture_subset)
    #print(model.model.summary("Loading model to disk completed for Dataset
    ↪gestures: ", gesture_subset)

    print("Testing model against outliers for Dataset gestures: ",
    ↪gesture_subset)
    data = dataset_outliers
    X, y, g = get_dataset(dataset_outliers)
    y_pred = model.predict(X)

    from sklearn.metrics import classification_report
    print(classification_report(y, y_pred, target_names=gesture_subset))

    from sklearn.metrics import confusion_matrix
    cf_matrix = confusion_matrix(y, y_pred)
    make_confusion_matrix(cf_matrix, categories=gesture_subset, figsize=[8,8])
    return grid_result
base_transfer_set = ['01', '02', '04', '05', '08', '09', '12', '13', '16',
    ↪'17', '18', '20']
dataset = transfers_size_4[3]

```

```
results = create_best_model(dataset)
```

```
Loadind Dataset for gestures: ['03', '07', '10', '15']
653 samples loaded
Scaling Dataset for gestures: ['03', '07', '10', '15']
653 samples scaled
Cleaning Dataset for gestures: ['03', '07', '10', '15']
495 samples cleaned
158 samples outliers
Time slicing Cleaned Dataset for gestures: ['03', '07', '10', '15']
494 cleaned samples sliced
1 cleaned samples damaged
Time slicing Outliers Dataset for gestures: ['03', '07', '10', '15']
158 outliers samples sliced
0 outliers samples damaged

2021-09-21 12:19:38.798092: W
tensorflow/stream_executor/platform/default/dso_loader.cc:64] Could not load
dynamic library 'libcudart.so.11.0'; dlerror: libcudart.so.11.0: cannot open
shared object file: No such file or directory
2021-09-21 12:19:38.798118: I tensorflow/stream_executor/cuda/cudart_stub.cc:29]
Ignore above cudart dlerror if you do not have a GPU set up on your machine.

Hyperparameter tunning started for Dataset for gestures: ['03', '07', '10',
'15']
Fitting 5 folds for each of 72 candidates, totalling 360 fits

2021-09-21 12:19:43.258023: W
tensorflow/stream_executor/platform/default/dso_loader.cc:64] Could not load
dynamic library 'libcuda.so.1'; dlerror: libcuda.so.1: cannot open shared object
file: No such file or directory
2021-09-21 12:19:43.258054: W
tensorflow/stream_executor/cuda/cuda_driver.cc:269] failed call to cuInit:
UNKNOWN ERROR (303)
2021-09-21 12:19:43.258074: I
tensorflow/stream_executor/cuda/cuda_diagnostics.cc:156] kernel driver does not
appear to be running on this host (mqx-public): /proc/driver/nvidia/version does
not exist
2021-09-21 12:19:43.258695: I tensorflow/core/platform/cpu_feature_guard.cc:142]
This TensorFlow binary is optimized with oneAPI Deep Neural Network Library
(oneDNN) to use the following CPU instructions in performance-critical
operations: AVX2 AVX512F FMA
To enable them in other operations, rebuild TensorFlow with the appropriate
compiler flags.
2021-09-21 12:19:43.399564: I
tensorflow/compiler/mlir/mlir_graph_optimization_pass.cc:185] None of the MLIR
Optimization Passes are enabled (registered 2)

Best: 0.932148 using {'batch_size': 19, 'dropout_rate': 0.6, 'epochs': 64,
```

```

'units': 16}
accuracy: 0.900437 (0.076880) train time: 29.925908 (5.676378) score time:
2.083331 (0.266593) with: {'batch_size': 19, 'dropout_rate': 0.1, 'epochs': 64,
'units': 16}
accuracy: 0.881587 (0.108958) train time: 34.222636 (5.719787) score time:
2.918640 (0.791467) with: {'batch_size': 19, 'dropout_rate': 0.1, 'epochs': 64,
'units': 32}
accuracy: 0.858293 (0.111597) train time: 45.817595 (6.273369) score time:
5.283210 (1.439234) with: {'batch_size': 19, 'dropout_rate': 0.1, 'epochs': 64,
'units': 64}
accuracy: 0.887677 (0.067042) train time: 75.331881 (9.190818) score time:
6.806288 (1.294489) with: {'batch_size': 19, 'dropout_rate': 0.1, 'epochs': 64,
'units': 128}
accuracy: 0.871447 (0.090193) train time: 64.421398 (9.167876) score time:
4.905846 (1.360424) with: {'batch_size': 19, 'dropout_rate': 0.1, 'epochs': 128,
'units': 16}
accuracy: 0.889830 (0.091987) train time: 66.011446 (9.089509) score time:
5.261564 (0.476941) with: {'batch_size': 19, 'dropout_rate': 0.1, 'epochs': 128,
'units': 32}
accuracy: 0.863654 (0.079608) train time: 83.481520 (13.858943) score time:
6.183306 (0.777523) with: {'batch_size': 19, 'dropout_rate': 0.1, 'epochs': 128,
'units': 64}
accuracy: 0.895528 (0.075306) train time: 128.086164 (20.321135) score time:
6.113807 (1.392108) with: {'batch_size': 19, 'dropout_rate': 0.1, 'epochs': 128,
'units': 128}
accuracy: 0.907278 (0.074448) train time: 41.032224 (5.735573) score time:
4.887396 (1.402462) with: {'batch_size': 19, 'dropout_rate': 0.2, 'epochs': 64,
'units': 16}
accuracy: 0.885078 (0.086502) train time: 42.732735 (3.306724) score time:
2.937031 (0.151258) with: {'batch_size': 19, 'dropout_rate': 0.2, 'epochs': 64,
'units': 32}
accuracy: 0.862291 (0.101638) train time: 55.696650 (12.418311) score time:
4.835953 (1.031076) with: {'batch_size': 19, 'dropout_rate': 0.2, 'epochs': 64,
'units': 64}
accuracy: 0.841692 (0.081673) train time: 69.954213 (8.086484) score time:
3.831244 (0.694369) with: {'batch_size': 19, 'dropout_rate': 0.2, 'epochs': 64,
'units': 128}
accuracy: 0.855836 (0.101514) train time: 63.778155 (3.891571) score time:
5.149154 (0.698381) with: {'batch_size': 19, 'dropout_rate': 0.2, 'epochs': 128,
'units': 16}
accuracy: 0.879415 (0.075776) train time: 73.258273 (18.436235) score time:
4.790711 (1.412082) with: {'batch_size': 19, 'dropout_rate': 0.2, 'epochs': 128,
'units': 32}
accuracy: 0.877329 (0.097395) train time: 83.270574 (11.649810) score time:
4.908036 (0.955232) with: {'batch_size': 19, 'dropout_rate': 0.2, 'epochs': 128,
'units': 64}
accuracy: 0.852298 (0.118004) train time: 108.615655 (7.132567) score time:
4.630007 (1.238696) with: {'batch_size': 19, 'dropout_rate': 0.2, 'epochs': 128,

```

```

'units': 128}
accuracy: 0.880009 (0.103817) train time: 43.512692 (4.599273) score time:
3.469601 (0.896283) with: {'batch_size': 19, 'dropout_rate': 0.3, 'epochs': 64,
'units': 16}
accuracy: 0.901434 (0.099411) train time: 43.545901 (6.212829) score time:
5.139137 (1.550522) with: {'batch_size': 19, 'dropout_rate': 0.3, 'epochs': 64,
'units': 32}
accuracy: 0.890210 (0.092307) train time: 49.751404 (3.590541) score time:
3.681676 (0.562115) with: {'batch_size': 19, 'dropout_rate': 0.3, 'epochs': 64,
'units': 64}
accuracy: 0.838932 (0.077459) train time: 65.986879 (9.982110) score time:
5.744367 (1.310631) with: {'batch_size': 19, 'dropout_rate': 0.3, 'epochs': 64,
'units': 128}
accuracy: 0.874197 (0.106138) train time: 61.727243 (9.699572) score time:
6.298358 (0.492294) with: {'batch_size': 19, 'dropout_rate': 0.3, 'epochs': 128,
'units': 16}
accuracy: 0.907620 (0.061923) train time: 73.777014 (15.369106) score time:
5.753930 (1.828662) with: {'batch_size': 19, 'dropout_rate': 0.3, 'epochs': 128,
'units': 32}
accuracy: 0.856534 (0.073292) train time: 75.235379 (12.096932) score time:
5.128023 (1.135601) with: {'batch_size': 19, 'dropout_rate': 0.3, 'epochs': 128,
'units': 64}
accuracy: 0.879825 (0.087468) train time: 106.177438 (16.170638) score time:
4.794785 (0.894215) with: {'batch_size': 19, 'dropout_rate': 0.3, 'epochs': 128,
'units': 128}
accuracy: 0.884556 (0.038401) train time: 50.721738 (9.159083) score time:
5.843534 (1.843543) with: {'batch_size': 19, 'dropout_rate': 0.4, 'epochs': 64,
'units': 16}
accuracy: 0.846137 (0.134747) train time: 43.083370 (4.393811) score time:
4.795987 (1.264675) with: {'batch_size': 19, 'dropout_rate': 0.4, 'epochs': 64,
'units': 32}
accuracy: 0.890425 (0.052290) train time: 43.306775 (5.751725) score time:
5.484806 (1.541257) with: {'batch_size': 19, 'dropout_rate': 0.4, 'epochs': 64,
'units': 64}
accuracy: 0.868461 (0.086151) train time: 66.119879 (6.849083) score time:
5.381316 (1.192679) with: {'batch_size': 19, 'dropout_rate': 0.4, 'epochs': 64,
'units': 128}
accuracy: 0.916483 (0.083522) train time: 73.789673 (13.912724) score time:
4.827560 (0.960642) with: {'batch_size': 19, 'dropout_rate': 0.4, 'epochs': 128,
'units': 16}
accuracy: 0.892061 (0.077952) train time: 72.137717 (9.506126) score time:
4.950682 (0.777825) with: {'batch_size': 19, 'dropout_rate': 0.4, 'epochs': 128,
'units': 32}
accuracy: 0.879203 (0.066712) train time: 67.428534 (10.320606) score time:
4.318209 (0.427335) with: {'batch_size': 19, 'dropout_rate': 0.4, 'epochs': 128,
'units': 64}
accuracy: 0.884558 (0.092527) train time: 114.551649 (19.912904) score time:
4.407332 (1.194026) with: {'batch_size': 19, 'dropout_rate': 0.4, 'epochs': 128,

```

```

'units': 128}
accuracy: 0.899584 (0.053659) train time: 50.065105 (7.950339) score time:
5.610015 (1.867825) with: {'batch_size': 19, 'dropout_rate': 0.5, 'epochs': 64,
'units': 16}
accuracy: 0.885478 (0.099654) train time: 46.468902 (7.195022) score time:
5.054011 (1.519364) with: {'batch_size': 19, 'dropout_rate': 0.5, 'epochs': 64,
'units': 32}
accuracy: 0.927891 (0.078589) train time: 46.825308 (6.773892) score time:
3.962945 (1.338670) with: {'batch_size': 19, 'dropout_rate': 0.5, 'epochs': 64,
'units': 64}
accuracy: 0.875375 (0.082960) train time: 68.313393 (7.387582) score time:
4.631738 (1.083584) with: {'batch_size': 19, 'dropout_rate': 0.5, 'epochs': 64,
'units': 128}
accuracy: 0.908736 (0.062576) train time: 73.790948 (11.068430) score time:
5.312462 (0.486289) with: {'batch_size': 19, 'dropout_rate': 0.5, 'epochs': 128,
'units': 16}
accuracy: 0.882379 (0.117248) train time: 64.311689 (3.357163) score time:
3.889478 (1.230237) with: {'batch_size': 19, 'dropout_rate': 0.5, 'epochs': 128,
'units': 32}
accuracy: 0.885775 (0.111659) train time: 69.083720 (9.549086) score time:
4.830050 (1.073202) with: {'batch_size': 19, 'dropout_rate': 0.5, 'epochs': 128,
'units': 64}
accuracy: 0.898871 (0.084367) train time: 119.629602 (25.536268) score time:
4.909292 (1.463196) with: {'batch_size': 19, 'dropout_rate': 0.5, 'epochs': 128,
'units': 128}
accuracy: 0.932148 (0.042487) train time: 51.207886 (5.056629) score time:
4.736122 (1.003551) with: {'batch_size': 19, 'dropout_rate': 0.6, 'epochs': 64,
'units': 16}
accuracy: 0.887499 (0.094310) train time: 51.954100 (6.950839) score time:
5.603681 (1.545921) with: {'batch_size': 19, 'dropout_rate': 0.6, 'epochs': 64,
'units': 32}
accuracy: 0.880315 (0.083796) train time: 49.335445 (5.747802) score time:
5.397789 (1.172914) with: {'batch_size': 19, 'dropout_rate': 0.6, 'epochs': 64,
'units': 64}
accuracy: 0.890020 (0.086207) train time: 75.793433 (9.193106) score time:
6.360303 (1.303016) with: {'batch_size': 19, 'dropout_rate': 0.6, 'epochs': 64,
'units': 128}
accuracy: 0.917028 (0.085012) train time: 65.164346 (9.136642) score time:
4.356014 (0.808203) with: {'batch_size': 19, 'dropout_rate': 0.6, 'epochs': 128,
'units': 16}
accuracy: 0.899191 (0.122073) train time: 61.665110 (8.093458) score time:
4.213029 (1.094740) with: {'batch_size': 19, 'dropout_rate': 0.6, 'epochs': 128,
'units': 32}
accuracy: 0.886008 (0.081760) train time: 74.690525 (11.300004) score time:
5.486436 (1.746389) with: {'batch_size': 19, 'dropout_rate': 0.6, 'epochs': 128,
'units': 64}
accuracy: 0.897765 (0.055043) train time: 114.359126 (14.698073) score time:
4.676384 (0.922199) with: {'batch_size': 19, 'dropout_rate': 0.6, 'epochs': 128,

```



```

'units': 128}
accuracy: 0.884335 (0.101829) train time: 47.569819 (7.502455) score time:
5.911272 (1.732420) with: {'batch_size': 19, 'dropout_rate': 0.7, 'epochs': 64,
'units': 16}
accuracy: 0.892271 (0.050881) train time: 51.771283 (7.392998) score time:
5.534904 (1.623389) with: {'batch_size': 19, 'dropout_rate': 0.7, 'epochs': 64,
'units': 32}
accuracy: 0.910449 (0.057134) train time: 58.639985 (5.566624) score time:
5.867114 (1.071352) with: {'batch_size': 19, 'dropout_rate': 0.7, 'epochs': 64,
'units': 64}
accuracy: 0.878913 (0.092020) train time: 68.336628 (6.243376) score time:
3.918904 (0.848308) with: {'batch_size': 19, 'dropout_rate': 0.7, 'epochs': 64,
'units': 128}
accuracy: 0.876350 (0.098771) train time: 58.785006 (6.092515) score time:
4.913874 (1.354442) with: {'batch_size': 19, 'dropout_rate': 0.7, 'epochs': 128,
'units': 16}
accuracy: 0.885123 (0.068270) train time: 60.736710 (9.181203) score time:
5.308555 (0.566208) with: {'batch_size': 19, 'dropout_rate': 0.7, 'epochs': 128,
'units': 32}
accuracy: 0.890836 (0.063372) train time: 74.444985 (9.761698) score time:
4.664838 (1.508759) with: {'batch_size': 19, 'dropout_rate': 0.7, 'epochs': 128,
'units': 64}
accuracy: 0.914062 (0.053251) train time: 109.488767 (12.558439) score time:
6.152760 (1.376021) with: {'batch_size': 19, 'dropout_rate': 0.7, 'epochs': 128,
'units': 128}
accuracy: 0.913965 (0.067522) train time: 47.555478 (9.354448) score time:
5.549119 (1.341808) with: {'batch_size': 19, 'dropout_rate': 0.8, 'epochs': 64,
'units': 16}
accuracy: 0.915785 (0.065563) train time: 54.114048 (12.626399) score time:
7.948699 (1.264796) with: {'batch_size': 19, 'dropout_rate': 0.8, 'epochs': 64,
'units': 32}
accuracy: 0.925034 (0.060016) train time: 57.721693 (7.780804) score time:
6.546513 (0.957521) with: {'batch_size': 19, 'dropout_rate': 0.8, 'epochs': 64,
'units': 64}
accuracy: 0.908855 (0.049257) train time: 63.434842 (6.029799) score time:
4.780146 (0.467712) with: {'batch_size': 19, 'dropout_rate': 0.8, 'epochs': 64,
'units': 128}
accuracy: 0.854067 (0.137757) train time: 59.224992 (7.364666) score time:
4.820966 (0.895490) with: {'batch_size': 19, 'dropout_rate': 0.8, 'epochs': 128,
'units': 16}
accuracy: 0.906656 (0.075573) train time: 67.697570 (11.664566) score time:
6.400374 (1.235790) with: {'batch_size': 19, 'dropout_rate': 0.8, 'epochs': 128,
'units': 32}
accuracy: 0.884652 (0.076402) train time: 77.167600 (10.524143) score time:
6.167270 (2.672424) with: {'batch_size': 19, 'dropout_rate': 0.8, 'epochs': 128,
'units': 64}
accuracy: 0.902249 (0.038207) train time: 120.261101 (15.754278) score time:
6.379271 (1.736711) with: {'batch_size': 19, 'dropout_rate': 0.8, 'epochs': 128,

```

```

'units': 128}
accuracy: 0.851371 (0.125847) train time: 57.086769 (15.603497) score time:
6.606438 (1.140990) with: {'batch_size': 19, 'dropout_rate': 0.9, 'epochs': 64,
'units': 16}
accuracy: 0.919114 (0.056311) train time: 54.618171 (3.018825) score time:
5.137061 (0.891704) with: {'batch_size': 19, 'dropout_rate': 0.9, 'epochs': 64,
'units': 32}
accuracy: 0.879887 (0.097867) train time: 50.393857 (8.325274) score time:
6.460707 (2.566208) with: {'batch_size': 19, 'dropout_rate': 0.9, 'epochs': 64,
'units': 64}
accuracy: 0.900319 (0.076258) train time: 64.020380 (3.648576) score time:
4.589291 (1.490141) with: {'batch_size': 19, 'dropout_rate': 0.9, 'epochs': 64,
'units': 128}
accuracy: 0.865312 (0.133924) train time: 54.836777 (7.983391) score time:
3.537366 (0.703489) with: {'batch_size': 19, 'dropout_rate': 0.9, 'epochs': 128,
'units': 16}
accuracy: 0.886180 (0.059885) train time: 59.298305 (12.097153) score time:
4.819555 (1.697393) with: {'batch_size': 19, 'dropout_rate': 0.9, 'epochs': 128,
'units': 32}
accuracy: 0.901527 (0.077271) train time: 64.782819 (10.148060) score time:
4.062896 (0.669323) with: {'batch_size': 19, 'dropout_rate': 0.9, 'epochs': 128,
'units': 64}
accuracy: 0.914554 (0.060141) train time: 94.902497 (8.110197) score time:
4.869324 (1.872545) with: {'batch_size': 19, 'dropout_rate': 0.9, 'epochs': 128,
'units': 128}

```

Hyperparameter tuning completed for Dataset: ['03', '07', '10', '15']

Saving model to disk started for Dataset gestures: ['03', '07', '10', '15']

2021-09-21 19:29:19.508968: W tensorflow/python/util/util.cc:348] Sets are not currently considered sequences, but this may change in the future, so consider avoiding using them.

WARNING:absl:Found untraced functions such as lstm_cell_1081_layer_call_fn, lstm_cell_1081_layer_call_and_return_conditional_losses, lstm_cell_1082_layer_call_fn, lstm_cell_1082_layer_call_and_return_conditional_losses, lstm_cell_1081_layer_call_fn while saving (showing 5 of 10). These functions will not be directly callable after loading.

INFO:tensorflow:Assets written to: 03-07-10-15_lstm/assets

INFO:tensorflow:Assets written to: 03-07-10-15_lstm/assets

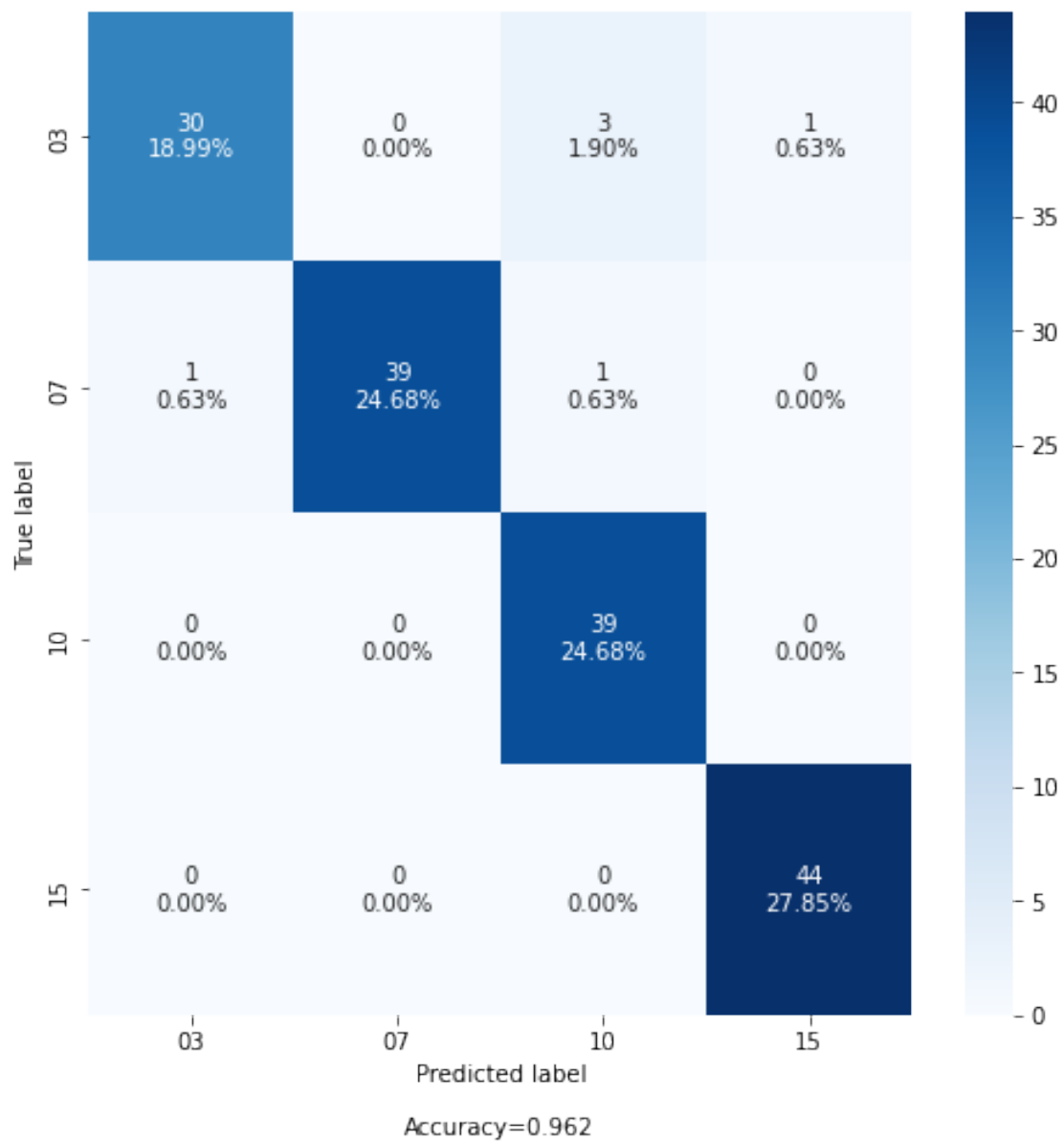
Saving model to disk completed for Dataset gestures: ['03', '07', '10', '15']

Loading model to disk started for Dataset gestures: ['03', '07', '10', '15']

Testing model against outliers for Dataset gestures: ['03', '07', '10', '15']

	precision	recall	f1-score	support
03	0.97	0.88	0.92	34
07	1.00	0.95	0.97	41
10	0.91	1.00	0.95	39

	15	0.98	1.00	0.99	44
accuracy				0.96	158
macro avg	0.96	0.96	0.96	0.96	158
weighted avg	0.96	0.96	0.96	0.96	158



```
[4]: import os
import pandas as pd
import warnings
```

```

warnings.filterwarnings("ignore")

baseset = dataset

def evaluate_model(baseset):
    print("Baseset: ", baseset)
    print("Loading Dataset: ", baseset)
    path = 'gestures-dataset'
    dataset = None

    samples = 0
    for subject in os.listdir(path):
        if os.path.isfile(os.path.join(path, subject)):
            continue
        if subject in ('U01', 'U02', 'U03', 'U04', 'U05', 'U06', 'U07', 'U08'):
            for gesture in os.listdir(os.path.join(path, subject)):
                if os.path.isfile(os.path.join(path, subject, gesture)):
                    continue
                gesture = str(gesture)
                if gesture not in baseset:
                    continue
                for samplefile in os.listdir(os.path.join(path, subject, gesture,
→gesture))):
                    if os.path.isfile(os.path.join(path, subject, gesture,
→samplefile)):
                        df = pd.read_csv(os.path.join(path, subject, gesture,
→samplefile), \
                                sep = ' ', \
                                names = ['System.currentTimeMillis()', \
                                'System.nanoTime()', \
                                'sample.timestamp', \
                                'X', \
                                'Y', \
                                'Z' \
                                ])
                        df = df[["sample.timestamp", "X", "Y", "Z"]]

                        start = df["sample.timestamp"][0]
                        df["sample.timestamp"] -= start
                        df["sample.timestamp"] /= 10000000
                        df["subject"] = subject
                        df["gesture"] = gesture
                        df["sample"] = str(samplefile[:-4])
                        samples += 1
                        #print(df)
                        if dataset is None:
                            dataset = df.copy()

```

```

        else:
            dataset = pd.concat([dataset, df])

    dataset = dataset.sort_values(by=['gesture', 'subject', 'sample', 'sample.
→timestamp'])
    data = dataset
    print(str(samples) + " samples loaded")

    print("Scaling Dataset: ", baseset)
    from sklearn.preprocessing import StandardScaler

    scaler = StandardScaler()
    dataset_scaled = None

    samples = 0
    for i, gesture in enumerate(baseset):
        df_gesture=data[data['gesture']==gesture]
        for j, subject in enumerate(df_gesture['subject'].unique()):
            df_subject=df_gesture[df_gesture['subject']==subject]
            for k, sample in enumerate(df_subject['sample'].unique()):
                df_sample=df_subject[df_subject['sample']==sample].copy()
                df_sample.sort_values(by=['sample.timestamp'])

                sc = scaler
                sc = sc.fit_transform(df_sample[["X", "Y", "Z"]])
                sc = pd.DataFrame(data=sc, columns=["X", "Y", "Z"])
                df_sample['X'] = sc['X']
                df_sample['Y'] = sc['Y']
                df_sample['Z'] = sc['Z']
                if dataset_scaled is None:
                    dataset_scaled = df_sample.copy()
                else:
                    dataset_scaled = pd.concat([dataset_scaled, df_sample])
                samples += 1
    print(str(samples) + " samples scaled")
    data = dataset_scaled

    print("Cleaning Dataset: ", baseset)
    dataset_outliers = None
    dataset_cleaned = None

    samples = 0
    outliers = 0
    for i, gesture in enumerate(baseset):
        df_gesture = data[data['gesture']==gesture]
        for j, subject in enumerate(df_gesture['subject'].unique()):
            df_subject = df_gesture[df_gesture['subject']==subject]

```

```

        time_mean = df_subject.groupby(["gesture", "subject", "sample"]).
→count().groupby(["gesture", "subject"]).agg({'sample.timestamp': ['mean']})
        time_std = df_subject.groupby(["gesture", "subject", "sample"]).
→count().groupby(["gesture", "subject"]).agg({'sample.timestamp': ['std']})
        time_max = time_mean['sample.timestamp'].iloc[0]['mean'] + 1.0 *
→time_std['sample.timestamp'].iloc[0]['std']
        time_min = time_mean['sample.timestamp'].iloc[0]['mean'] - 1.0 *
→time_std['sample.timestamp'].iloc[0]['std']
        for k, sample in enumerate(df_subject['sample'].unique()):
            df_sample=df_subject[df_subject['sample']==sample]
            df_sample_count = df_sample.count()['sample.timestamp']
            if df_sample_count < time_min or df_sample_count > time_max:
                if dataset_outliers is None:
                    dataset_outliers = df_sample.copy()
                else:
                    dataset_outliers = pd.concat([dataset_outliers,
→df_sample])
                    outliers += 1
            else:
                if dataset_cleaned is None:
                    dataset_cleaned = df_sample.copy()
                else:
                    dataset_cleaned = pd.concat([dataset_cleaned,
→df_sample])
                    samples += 1
        print(str(samples) + " samples cleaned")
        print(str(outliers) + " samples outliers")
        data = dataset_cleaned

        print("Time slicing Cleaned Dataset: ", baseset)
        dataset_timecut = None
        samples = 0
        damaged = 0
        for i, gesture in enumerate(data['gesture'].unique()):
            df_gesture = data[data['gesture']==gesture]
            for j, subject in enumerate(df_gesture['subject'].unique()):
                df_subject = df_gesture[df_gesture['subject']==subject]
                time_max = 19 # 18 * 11 = 198
                for i, sample in enumerate(df_subject['sample'].unique()):
                    df_sample = df_subject[df_subject['sample']==sample]
                    df_sample_count = df_sample.count()['sample.timestamp']
                    #print(df_sample_count)
                    if df_sample_count >= time_max:
                        df_sample = df_sample[df_sample['sample.timestamp'] <= (11
→* (time_max-1))]

```

```

        df_sample_count = df_sample.count()['sample.timestamp']
        #print(df_sample_count)
    elif df_sample_count < time_max:
        for tmp in range(df_sample_count * 11, (time_max) * 11, 11):
            df = pd.DataFrame([[tmp, 0.0, 0.0, 0.0, gesture,
↳subject, sample]], columns=['sample.timestamp', 'X', 'Y', 'Z', 'gesture',
↳'subject', 'sample'])

            df_sample = df_sample.append(df, ignore_index=True)
            #print(df_sample)
            df_sample_count = df_sample.count()['sample.timestamp']
            #print(df_sample_count)
            if df_sample_count != time_max:
                damaged += 1
                continue
            if dataset_timecut is None:
                dataset_timecut = df_sample.copy()
            else:
                dataset_timecut = pd.concat([dataset_timecut, df_sample])
            samples += 1

dataset_cleaned = dataset_timecut
print(str(samples) + " cleaned samples sliced")
print(str(damaged) + " cleaned samples damaged")

data = dataset_outliers
print("Time slicing Outliers Dataset: ", baseset)
dataset_timecut = None
samples = 0
damaged = 0
for i, gesture in enumerate(data['gesture'].unique()):
    df_gesture = data[data['gesture']==gesture]
    for j, subject in enumerate(df_gesture['subject'].unique()):
        df_subject = df_gesture[df_gesture['subject']==subject]
        time_max = 19 # 18 * 11 = 198
        for i, sample in enumerate(df_subject['sample'].unique()):
            df_sample = df_subject[df_subject['sample']==sample]
            df_sample_count = df_sample.count()['sample.timestamp']
            #print(df_sample_count)
            if df_sample_count >= time_max:
                df_sample = df_sample[df_sample['sample.timestamp'] <= (11
↳* (time_max-1))]

                df_sample_count = df_sample.count()['sample.timestamp']
                #print(df_sample_count)
            elif df_sample_count < time_max:
                for tmp in range(df_sample_count * 11, (time_max) * 11, 11):

```

```

        df = pd.DataFrame([[tmp, 0.0, 0.0, 0.0, gesture,
↪subject, sample]], columns=['sample.timestamp', 'X', 'Y', 'Z', 'gesture',
↪'subject', 'sample'])

        df_sample = df_sample.append(df, ignore_index=True)
        #print(df_sample)
        df_sample_count = df_sample.count()['sample.timestamp']
        #print(df_sample_count)
        if df_sample_count != time_max:
            damaged += 1
            continue
        if dataset_timecut is None:
            dataset_timecut = df_sample.copy()
        else:
            dataset_timecut = pd.concat([dataset_timecut, df_sample])
        samples += 1

dataset_outliers = dataset_timecut
print(str(samples) + " outliers samples sliced")
print(str(damaged) + " outliers samples damaged")

from keras import backend as K
data = dataset_cleaned
from keras.models import Sequential
from keras.layers import Bidirectional
from keras.layers import LSTM
from keras.layers import Dense
from keras.layers import Dropout
from keras.optimizers import adam_v2
from keras.wrappers.scikit_learn import KerasClassifier
from sklearn.model_selection import StratifiedGroupKFold
from sklearn.model_selection import cross_validate
from sklearn.model_selection import GridSearchCV
from keras.utils import np_utils
from sklearn.preprocessing import LabelEncoder
from sklearn.pipeline import Pipeline
from sklearn.metrics import accuracy_score
import numpy as np
import tensorflow as tf

# fix random seed for reproducibility
seed = 1000
np.random.seed(seed)
# create the dataset
def get_dataset(data, index=[]):
    X_train = []
    Y_train = []
    groups = []

```



```

samples_idx=0
for i, gesture in enumerate(data['gesture'].unique()):
    df_gesture = data[data['gesture']==gesture]
    for j, subject in enumerate(df_gesture['subject'].unique()):
        df_subject = df_gesture[df_gesture['subject']==subject]
        for k, sample in enumerate(df_subject['sample'].unique()):
            df_sample = df_subject[df_subject['sample']==sample]
            accel_vector = []
            for idx, row in df_sample.sort_values(by='sample.
↳timestamp').iterrows():
                accel_vector.append([row['X'],row['Y'],row['Z']])
            accel_vector = np.asarray(accel_vector)
            if len(index)==0:
                X_train.append(accel_vector)
                Y_train.append(gesture)
                groups.append(subject)
            else:
                if samples_idx in index:
                    X_train.append(accel_vector)
                    Y_train.append(gesture)
                    groups.append(subject)
                samples_idx+=1
X_train = np.asarray(X_train)
Y_train = LabelEncoder().fit_transform(Y_train)
#print(Y_train)
return X_train, Y_train, groups

def build_model(baseset):
    baseset.sort()
    basename = '-'.join(baseset)
    basemodel = tf.keras.models.load_model(basename + '_lstm')
    basemodel.build([None, 19, 3])
    #print(model.summary())
    basemodel.compile(loss='sparse_categorical_crossentropy',
↳optimizer=adam_v2.Adam(learning_rate=0.001), metrics=['accuracy'])
    return basemodel

# Function to create model, required for KerasClassifier
import pickle
def load_classifier(baseset):
    baseset.sort()
    basename = '-'.join(baseset)
    classifier = KerasClassifier(build_fn=build_model, baseset=baseset,
↳epochs=64, batch_size=19, verbose=0)
    classifier.classes_ = pickle.load(open(basename + '_model_classes.
↳pkl','rb'))

```

```

        classifier.model = build_model(baseset)
        return classifier

    #print(model.model.summary())
    #print(model.classes_)
    from sklearn.metrics import classification_report
    from sklearn.metrics import confusion_matrix

    for n_splits in [5]:
        for epoch in [[results.best_params_['epochs']]]:
            cv = StratifiedGroupKFold(n_splits=n_splits, shuffle=True,
→random_state=(1000+epoch[0]))
            X, y, g = get_dataset(dataset_cleaned)

            # Initialize the accuracy of the models to blank list. The accuracy
→of each model will be appended to this list
            accuracy_model = []
            best_estimator = None
            # Initialize the array to zero which will store the confusion matrix
            array = None
            outliers = None

            report_cleaned = None
            report_outliers = None

            print("Processing started for split estimator: " + str(n_splits) +
→", epochs: " + str(epoch))
            # Iterate over each train-test split
            fold = 1
            for train_index, test_index in cv.split(X, y, g):
                #print(test_index)
                if len(test_index) == 0 or len(train_index) == 0:
                    continue
                print("Processing ", fold, "-fold")
                fold += 1

                classifier = load_classifier(baseset)
                # Split train-test (Inverted)
                X_train, y_train, group_train = get_dataset(dataset_cleaned,
→train_index)
                X_test, y_test, group_test = get_dataset(dataset_cleaned,
→test_index)
                X_outliers, y_outliers, group_test =
→get_dataset(dataset_outliers)
                # Train the model
                History = classifier.fit(X_train, y_train, epochs=epoch[0])

```

```

        # Append to accuracy_model the accuracy of the model
        accuracy_model.append(accuracy_score(y_test, classifier.
→predict(X_test), normalize=True))
        if accuracy_model[-1] == max(accuracy_model):
            best_estimator = classifier
        # Calculate the confusion matrix
        c = confusion_matrix(y_test, classifier.predict(X_test))
        # Add the score to the previous confusion matrix of previous
→model

        if isinstance(array, np.ndarray) == False:
            array = c.copy()
        else:
            array = array + c

        # Calculate the confusion matrix
        c = confusion_matrix(y_outliers, classifier.predict(X_outliers))
        # Add the score to the previous confusion matrix of previous
→model

        if isinstance(outliers, np.ndarray) == False:
            outliers = c.copy()
        else:
            outliers = outliers + c

        #Accumulate for classification report
        if isinstance(report_cleaned, list) == False:
            report_cleaned = [y_test, classifier.predict(X_test)]
        else:
            report_cleaned[0] = np.append(report_cleaned[0],y_test)
            report_cleaned[1] = np.append(report_cleaned[1],classifier.
→predict(X_test))

        #Accumulate for classification report
        if isinstance(report_outliers, list) == False:
            report_outliers = [y_outliers, classifier.
→predict(X_outliers)]
        else:
            report_outliers[0] = np.
→append(report_outliers[0],y_outliers)
            report_outliers[1] = np.
→append(report_outliers[1],classifier.predict(X_outliers))

        # Print the accuracy
        print("At split estimator: " + str(n_splits) + ", epochs: " +
→str(epoch))
        print("Accurace mean(std): " + str(np.mean(accuracy_model)) + "(" +
→str(np.std(accuracy_model)) + ")")

```

```

        # To calculate the classification reports

        print("Classification report for all valid cross_validations_
↳against their tests sets")
        print(classification_report(report_cleaned[0], report_cleaned[1],
↳target_names=baseset))

        print("Classification report for all valid cross_validations_
↳against outliers")
        print(classification_report(report_outliers[0], report_outliers[1],
↳target_names=baseset))

        # To calculate the confusion matrix

        print("Confusion Matrix for all valid cross_validations against_
↳their tests sets")
        make_confusion_matrix(array, categories=baseset, figsize=[8,8])

        print("Confusion Matrix for all valid cross_validations against_
↳outliers")
        make_confusion_matrix(outliers, categories=baseset, figsize=[8,8])
def save_model(model, baseset):
    baseset.sort()
    name = '-'.join(baseset)
    # saving model
    pickle.dump(model.classes_, open(name + '_model_classes.pkl','wb'))
    model.model.save(name + '_lstm')
    save_model(best_estimator, baseset)

model = evaluate_model(baseset)

```

```

Baseset:  ['03', '07', '10', '15']
Loadind Dataset:  ['03', '07', '10', '15']
653 samples loaded
Scaling Dataset:  ['03', '07', '10', '15']
653 samples scaled
Cleaning Dataset:  ['03', '07', '10', '15']
495 samples cleaned
158 samples outliers
Time slicing Cleaned Dataset:  ['03', '07', '10', '15']
494 cleaned samples sliced
1 cleaned samples damaged
Time slicing Outliers Dataset:  ['03', '07', '10', '15']
158 outliers samples sliced
0 outliers samples damaged
Processing started for split estimator: 5, epochs: [64]

```

```

Processing 1 -fold
Processing 2 -fold
Processing 3 -fold
Processing 4 -fold
Processing 5 -fold
At split estimator: 5, epochs: [64]
Accurace mean(std): 0.9953125(0.009375)
Classification report for all valid cross_validations against their tests sets

```

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

03	0.98	1.00	0.99	132
07	1.00	1.00	1.00	119
10	1.00	0.98	0.99	125
15	1.00	1.00	1.00	118
accuracy			0.99	494
macro avg	0.99	0.99	0.99	494
weighted avg	0.99	0.99	0.99	494

```

Classification report for all valid cross_validations against outliers

```

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

03	0.97	0.90	0.93	170
07	1.00	0.95	0.97	205
10	0.92	1.00	0.96	195
15	0.98	1.00	0.99	220
accuracy			0.97	790
macro avg	0.97	0.96	0.96	790
weighted avg	0.97	0.97	0.97	790

```

Confusion Matrix for all valid cross_validations against their tests sets
Confusion Matrix for all valid cross_validations against outliers

```

```

WARNING:absl:Found untraced functions such as lstm_cell_1114_layer_call_fn,
lstm_cell_1114_layer_call_and_return_conditional_losses,
lstm_cell_1115_layer_call_fn,
lstm_cell_1115_layer_call_and_return_conditional_losses,
lstm_cell_1114_layer_call_fn while saving (showing 5 of 10). These functions
will not be directly callable after loading.

```

```

INFO:tensorflow:Assets written to: 03-07-10-15_lstm/assets

```

```

INFO:tensorflow:Assets written to: 03-07-10-15_lstm/assets

```

