In [1]:
```
import glob
import numpy as np
import matplotlib as mpl
import matplotlib.pyplot as plt
import pandas as pd
import os
```

In [2]:
```
# Create datacsv Directory for all the dataset (A,B,C,D, and E) if don't exist

os.makedirs('H:/Final_project/Data/Data_csv/A/',exist_ok = True)
os.makedirs('H:/Final_project/Data/Data_csv/B/',exist_ok = True)
os.makedirs('H:/Final_project/Data/Data_csv/C/',exist_ok = True)
os.makedirs('H:/Final_project/Data/Data_csv/D/',exist_ok = True)
os.makedirs('H:/Final_project/Data/Data_csv/E/',exist_ok = True)
```

In [ ]:
```
# Create chunk Directory for all the dataset (A,B,C,D, and E) if don't exist

for num_files in range(1,101):
    os.makedirs(os.path.join('H:/Final_project/Data/chunk/A/' + str(num_files)),exist_o
    os.makedirs(os.path.join('H:/Final_project/Data/chunk/B/' + str(num_files)),exist_o
    os.makedirs(os.path.join('H:/Final_project/Data/chunk/C/' + str(num_files)),exist_o
    os.makedirs(os.path.join('H:/Final_project/Data/chunk/D/' + str(num_files)),exist_o
    os.makedirs(os.path.join('H:/Final_project/Data/chunk/E/' + str(num_files)),exist_o
```

In [3]:
```
# Create merged Directory for all the dataset (A,B,C,D, and E) if don't exist

os.makedirs('H:/Final_project/Data/merged/A/',exist_ok = True)
os.makedirs('H:/Final_project/Data/merged/B/',exist_ok = True)
os.makedirs('H:/Final_project/Data/merged/C/',exist_ok = True)
os.makedirs('H:/Final_project/Data/merged/D/',exist_ok = True)
os.makedirs('H:/Final_project/Data/merged/E/',exist_ok = True)
```

In [4]:
```
```

In [5]:
```
# We downloaded the dataset using the given link https://repositori.upf.edu/handle/1023
# The datasets contain file name A, B, C, D, and E

# We changed the format file txt to csv
#Each dataset contains 100-single channel EEG segments with duration of 23.6 seconds an
#the corresponding time series is sampled into 4097 data points.

# Folder A
path1A = r'H:/Final_project/Data/Raw_data/A'
path2A = r'H:/Final_project/Data/Data_csv/A'
for i in range(1,101):
    fileA = pd.read_csv(os.path.join(path1A,(str(i)+ ".txt")))
    fileA = fileA.drop(labels=[0,1])
    new_csv_fileA = fileA.to_csv(os.path.join(path2A, (str(i) + ".csv")),index=False)
```

```
# Folder B

path1B = r'H:/Final_project/Data/Raw_data/B'
path2B = r'H:/Final_project/Data/Data_csv/B'
for i in range(1,101):
    fileB = pd.read_csv(os.path.join(path1B,(str(i)+ ".txt")))
    fileB = fileB.drop(labels=[0,1])
    new_csv_fileB = fileB.to_csv(os.path.join(path2B, (str(i) + ".csv")),index=False)


# Folder C

path1C = r'H:/Final_project/Data/Raw_data/C'
path2C = r'H:/Final_project/Data/Data_csv/C'
for i in range(1,101):
    fileC = pd.read_csv(os.path.join(path1C,(str(i)+ ".txt")))
    fileC = fileC.drop(labels=[4094,4095])
    new_csv_fileC = fileC.to_csv(os.path.join(path2C, (str(i) + ".csv")),index=False)



# Folder D

path1D = r'H:/Final_project/Data/Raw_data/D'
path2D = r'H:/Final_project/Data/Data_csv/D'
for i in range(1,101):
    fileD = pd.read_csv(os.path.join(path1D,(str(i)+ ".txt")))
    fileD = fileD.drop(labels=[0,1])
    new_csv_fileD = fileD.to_csv(os.path.join(path2D, (str(i) + ".csv")),index=False)



# Folder E

path1E = r'H:/Final_project/Data/Raw_data/E'
path2E = r'H:/Final_project/Data/Data_csv/E'
for i in range(1,101):
    fileE = pd.read_csv(os.path.join(path1E,(str(i)+ ".txt")))
    fileE = fileE.drop(labels=[0,1])
    new_csv_fileE = fileE.to_csv(os.path.join(path2E, (str(i) + ".csv")),index=False)
```

In [16]:

```
df1 = pd.read_csv(r'H:/Final_project/Data/Data_csv/A/1.csv')
df1.shape
#After reshaping each file contain 4094 samples
```

Out[16]: (4094, 1)

In [6]:

```
# We saved the data into a folder and then divide each files into 23 chunk and each con
#finally merged the 23 chunk file into a single file


# Merged folder A

filenameA =r'H:/Final_project/Data/Data_csv/A'
out_pathA = r'H:/Final_project/Data/Chunk/A/'
```

```
merged_pathA = r'H:/Final_project/Data/merged/A'
for num_files in range (1,101):
    fileA = pd.read_csv(os.path.join(filenameA,(str(num_files)+ ".csv")))
    dir_pathA = os.path.join(out_pathA +str(num_files)+'/')
    j =0
    for i in range(len(fileA)):
        if i % 178 == 0:
            j = j+1
            fileA[i:i+178].to_csv(os.path.join(dir_pathA, (str(j) + ".csv")),index=Fals

    dir_mergedpathA = os.path.join(merged_pathA +'/')
    allmerged_filesA =glob.glob(out_pathA + str(num_files)+'/'+ "/*.csv")
    df_from_each_fileA = (pd.read_csv(f, sep=',').transpose() for f in allmerged_filesA
    df_mergedA   = pd.concat(df_from_each_fileA, ignore_index=True)
    df_mergedA.to_csv(os.path.join(merged_pathA, (str(num_files) + ".csv")),index=False
```

In [7]:

```
# Merged folder B
filenameB =r'H:/Final_project/Data/Data_csv/B'
out_pathB = r'H:/Final_project/Data/Chunk/B/'
merged_pathB = r'H:/Final_project/Data/merged/B'
for num_files in range (1,101):
    fileB = pd.read_csv(os.path.join(filenameB,(str(num_files)+ ".csv")))
    #fileB = fileB.drop(columns=['Unnamed: 0'])
    dir_pathB = os.path.join(out_pathB +str(num_files)+'/')
    j =0
    for i in range(len(fileB)):
        if i % 178 == 0:
            j = j+1
            fileB[i:i+178].to_csv(os.path.join(dir_pathB, (str(j) + ".csv")),index=Fals

    dir_mergedpathB = os.path.join(merged_pathB +'/')
    allmerged_filesB =glob.glob(out_pathB + str(num_files)+'/'+ "/*.csv")
    df_from_each_fileB = (pd.read_csv(f, sep=',').transpose() for f in allmerged_filesB
    df_mergedB   = pd.concat(df_from_each_fileB, ignore_index=True)
    df_mergedB.to_csv(os.path.join(merged_pathB, (str(num_files) + ".csv")),index=False
```

In [8]:

```
# Merged folder C
filenameC =r'H:/Final_project/Data/Data_csv/C'
out_pathC = r'H:/Final_project/Data/Chunk/C/'
merged_pathC = r'H:/Final_project/Data/merged/C'
for num_files in range (1,101):
    fileC = pd.read_csv(os.path.join(filenameC,(str(num_files)+ ".csv")))
    #fileC = fileC.drop(columns=['Unnamed: 0'])
    dir_pathC = os.path.join(out_pathC +str(num_files)+'/')
    j =0
    for i in range(len(fileC)):
        if i % 178 == 0:
            j = j+1
            fileC[i:i+178].to_csv(os.path.join(dir_pathC, (str(j) + ".csv")),index=Fals

    dir_mergedpathC = os.path.join(merged_pathC +'/')
    allmerged_filesC =glob.glob(out_pathC + str(num_files)+'/'+ "/*.csv")
    df_from_each_fileC = (pd.read_csv(f, sep=',').transpose() for f in allmerged_filesC
    df_mergedC   = pd.concat(df_from_each_fileC, ignore_index=True)
    df_mergedC.to_csv(os.path.join(merged_pathC, (str(num_files) + ".csv")),index=False
```

In [9]:
```
# Merged folder D

filenameD =r'H:/Final_project/Data/Data_csv/D'
out_pathD = r'H:/Final_project/Data/Chunk/D/'
merged_pathD = r'H:/Final_project/Data/merged/D'
for num_files in range (1,101):
    fileD = pd.read_csv(os.path.join(filenameD,(str(num_files)+ ".csv")))
    #fileD = fileD.drop(columns=['Unnamed: 0'])
    dir_pathD = os.path.join(out_pathD +str(num_files)+'/')
    j =0
    for i in range(len(fileD)):
        if i % 178 == 0:
            j = j+1
            fileD[i:i+178].to_csv(os.path.join(dir_pathD, (str(j) + ".csv")),index=Fals

    dir_mergedpathD = os.path.join(merged_pathD +'/')
    allmerged_filesD =glob.glob(out_pathD + str(num_files)+'/'+ "/*.csv")
    df_from_each_fileD = (pd.read_csv(f, sep=',').transpose() for f in allmerged_filesD
    df_mergedD   = pd.concat(df_from_each_fileD, ignore_index=True)
    df_mergedD.to_csv(os.path.join(merged_pathD, (str(num_files) + ".csv")),index=False
```

In [10]:
```
# Merged folder E

filenameE =r'H:/Final_project/Data/Data_csv/E'
out_pathE = r'H:/Final_project/Data/Chunk/E/'
merged_pathE = r'H:/Final_project/Data/merged/E'
for num_files in range (1,101):
    fileE = pd.read_csv(os.path.join(filenameE,(str(num_files)+ ".csv")))
    #fileE = fileE.drop(columns=['Unnamed: 0'])
    dir_pathE = os.path.join(out_pathE +str(num_files)+'/')
    j =0
    for i in range(len(fileE)):
        if i % 178 == 0:
            j = j+1
            fileE[i:i+178].to_csv(os.path.join(dir_pathE, (str(j) + ".csv")),index=Fals

    dir_mergedpathE = os.path.join(merged_pathE +'/')
    allmerged_filesE =glob.glob(out_pathE + str(num_files)+'/'+ "/*.csv")
    df_from_each_fileE = (pd.read_csv(f, sep=',').transpose() for f in allmerged_filesE
    df_mergedE   = pd.concat(df_from_each_fileE, ignore_index=True)
    df_mergedE.to_csv(os.path.join(merged_pathE, (str(num_files) + ".csv")),index=False
```

In [ ]:

In [ ]: