

§2 Второй отборочный этап

В следующих двух задачах используется один и тот же файл, расположенный по адресу: <https://drive.google.com/file/d/0Bz6hKL3efzRwajVtS1V4aDhCTjQ/view?usp=sharing>

В течение нескольких секунд происходила запись ЭЭГ с помощью восьмиканального энцефалографа. Результат представлен в виде csv-файла, каждая строка которого представляет собой данные с одного из каналов. Каналом будем называть электрод, с которого снимается физиологический сигнал, т.е. создающийся преимущественно колонками нейронов. Референтным электродом будем называть электрод, потенциал на котором мы будем принимать за ноль (обычно референтный электрод располагается на мочке уха). В строке файла через запятую записаны значения разности потенциалов между референтным электродом и каналом, выраженные в мВ, снимаемые через каждые 8.26 мс. Дельта-диапазоном будем называть диапазон частот 4-8

Гц, альфа-диапазоном – 8-14 Гц, бета-диапазоном – 14-20 Гц.

Для решения кейса на языке Python вам может понадобиться следующая информация

<https://docs.scipy.org/doc/numpy-dev/user/quickstart.html>

<https://docs.scipy.org/doc/numpy-dev/user/quickstart.html>

<https://docs.python.org/2/library/csv.html>

<https://docs.scipy.org/doc/scipy-0.18.1/reference/tutorial/index.html>

https://ru.wikipedia.org/wiki/%D0%9A%D0%BE%D0%BC%D0%BF%D0%BB%D0%B5%D0%BA%D1%81%D0%BD%D0%BE%D0%B5_%D1...

<https://ru.wikipedia.org/wiki/%D0%9F%D1%80%D0%B5%D0%BE%D0%B1%D1%80%D0%B0%D0%B7%D0%BE%D0%B2%D0%B0%D0%...>

Задача 2.1 (2 балла)

Условие:

Перечислите через запятую любые три момента начала морганий или шевеления челюстями, ушами, бровями с точностью до 100 мс. Сигнал в эти моменты называется артефактным.

Решение:

Достаточно посмотреть на график напряжения с одного или нескольких каналов ЭЭГ и отметить резкие скачки. В этом может помочь код:

```
import numpy as np
import csv
import pylab

eeg = []

with open('...\data.csv', 'rb') as csvfile:
    eegreader = csv.reader(csvfile, delimiter=',',
quotechar='|')
    for row in eegreader:
        eeg_row = []
        for element in row:
            if len(element.split(' '))>1:
                eeg_row.append(float(element.split('
')[len(element.split(' '))-1]))
        eeg.append(eeg_row)
eeg = np.array(eeg)

print np.shape(eeg)
```

```

time = [t*1./121 for t in range(len(eeg[0]))]
print max(time)
n_channels = np.shape(eeg)[0]

pylab.figure()
for ch in range(n_channels):
    pylab.subplot(n_channels,1,ch+1)
    pylab.plot(time,eeg[ch])
pylab.show()

```

Ответ:

1.16, 1.71, 1.63, 2.58, 3.45, 5.625, 6.57, 7.11

Задача 2.2 (8 баллов)

Условие:

Разделите запись со второго канала на фрагменты с начала записи до конца третьей секунды, с начала четвертой до конца шестой, с начала седьмой до конца девятой. Для каждого из фрагментов рассчитайте отношения мощности частот каждого из диапазонов к мощности ЭЭГ в диапазоне от 4 до 20 Гц. Ответ запишите в виде строки с числами через запятую в порядке дельта1-альфа1-бетта1-...- дельта3-альфа3-бетта3 . Всего в ответе должно быть девять чисел. Мощность сигнала на частоте - квадрат модуля амплитуды сигнала на этой частоте.

Решение :

```

import numpy as np
import csv
import pylab

eeg = []

with open('C:\NVR\Zonov\states\state31.csv', 'rb') as csvfile:
    eegreader = csv.reader(csvfile, delimiter=',',
quotechar='|')
    for row in eegreader:
        eeg_row = []
        for element in row:
            if len(element.split(' '))>1:
                eeg_row.append(float(element.split('
')[len(element.split(' '))-1]))
        eeg.append(eeg_row)
eeg = np.array(eeg)

print np.shape(eeg)

time = [t*1./121 for t in range(len(eeg[0]))]
print max(time)
n_channels = np.shape(eeg)[0]

pylab.figure()
for ch in range(n_channels):
    pylab.subplot(n_channels,1,ch+1)
    pylab.plot(time,eeg[ch])
pylab.show()

```

Ответ:

0.69, 0.22, 0.09, 0.82, 0.14, 0.05, 0.6, 0.19, 0.21

Задача 3 (20 баллов)**Условие:**

Мыслительная активность человека хорошо проявляется на ЭЭГ, на этом основано нейропилотирование - управление предметами с помощью мысленных команд. В архиве, расположенном по адресу

<https://drive.google.com/file/d/0Bz6hKL3efzRwZmd5eHpfOFBLNzg/view?usp=sharing>, есть файлы с записями ЭЭГ одного из нейропилотов. Число d в названии файла `state_s_day_d.csv` означает день, в который проводилась запись, число s означает состояние, в котором находился нейропилот: нулю соответствует нормальное состояние, единице - расслабление, двойке - концентрация. Как и в первой части этого этапа, в строке файла через запятую записаны значения разности потенциалов между референтным электродом и каналом, выраженные в мВ, снимаемые через каждые 8.26 мс. В каждой задаче этого этапа вам будет дан один файл того же формата с записью ЭЭГ этого же нейропилота и вам нужно будет сказать, в каком состоянии он находился в каждую секунду записи.

Необходимо через запятую перечислить состояния (0,1 или 2), в котором пилот находился в каждую секунду записи. В записи девять секунд, в ответе должны быть написаны девять чисел через запятую. Допускаются ошибки в трёх секундах.

Решение:

Участники могут использовать различные признаки и классификаторы. Вот пример одного из возможных решений.

Как и во второй задаче, вычисляем отношения мощностей в диапазонах дельта-альфа-бета.

```
data = []
with open('test.csv', 'rb') as csvfile:
    eegreader = csv.reader(csvfile, delimiter=',',
quotechar='|')
    for row in eegreader:
        eeg_row = []
        #for element in row:
        #    if len(element.split(' '))>1:
        #        eeg_row.append(float(element.split('
')[len(element.split(' '))-1]))
        data.append(    row)
data_array = array(data)

FD=121
N=363
X = []

for n in range(1,8):
    spectrum=rfft(data_array[n, FD*0:3*FD])
    power=(np_abs(spectrum)/N)**2
    freqs=rfftfreq(N, 1./FD)
    X.append(sum(power[12:24])/sum(power[12:60]))
    X.append(sum(power[24:42])/sum(power[12:60]))
    X.append(sum(power[42:60])/sum(power[12:60]))
```

Далее используем один из стандартных классификаторов (в x заносим аналогичные параметры для фрагмента из файла, который мы распознаем):

```

clf = GaussianNB()
clf.fit(X, Y)
print(clf.predict(x))

```

Задача 4 (12 баллов)

Условие:

В следующей задаче вам будет дан файл, склеенный из записей ЭЭГ одного нейропилота. Известно, что записи, из которых склеен файл, относятся к четырём различным состояниям и длина каждой из них не меньше секунды. Частота записи - 121 Гц (то есть каждой секунде записей соответствует последовательность из 121 числа в строке). Для записи использовалось 8 каналов. При склейке записи одного канала приклеивались к записям того же самого канала.

Файл находится по адресу:

<https://drive.google.com/file/d/0Bz6hKL3efzRwZnktbFJqa1NKQzA/view?usp=sharing>

Для каждой секунды файла определить, к какому состоянию она относится. В ответе через запятую укажите последовательность из 12 цифр от 0 до 3 через запятую.

Одинаковые цифры должны обозначать одинаковое состояние, разные - разные состояния. Какие конкретно будут цифры - не важно, то есть ответ "1,2,3,1,2,3,0,2,3,0,2,3" равнозначен ответу "0,1,2,0,1,2,3,1,2,3,1,2" так как в этих ответах на одинаковые номера соответствуют одним и тем же секундам (первой и четвертой секунде в первом ответе соответствует одинаковая цифра '1', а во втором ответе одинаковая цифра '0'; второй, пятой, восьмой и одиннадцатой секунде в первом ответе соответствует цифра '2', а во втором цифра '1').

Решение:

Построим график для любого из каналов, воспользовавшись кодом из первой задачи. На нем видны повторяющиеся фрагменты длиной 2 секунды. Делаем предположение, что весь файл разбит на фрагменты по 2 секунды из каждого состояния. Далее, как и в задаче 2, считаем отношение мощности частот к мощности ЭЭГ в диапазонах (дельта-альфа-бетта) для каждого из двухсекундных фрагментов и заносим каждые 3 числа в отдельный элемент массива X.

```

for FD in range(6):
    x = []
    for n in range(1, 8):
        spectrum=rfft(data_array[n, FD*0:2*FD])
        power=(np_abs(spectrum)/N)**2
        freqs=rfftfreq(N, 1./FD)
        delta = sum(power[8:16])/sum(power[8:40])
        alpha = sum(power[16:28])/sum(power[8:40])
        beta = sum(power[28:40])/sum(power[8:40])
        x.append(delta)
        x.append(alpha)
        x.append(beta)
    X.append(x)

```

Используем кластеризацию методом k-средних

```

from sklearn.cluster import KMeans
import numpy as np
kmeans = KMeans(n_cluster=4).fit(X)
print(kmeans.labels)

```

На забываем в ответе «удваивать» числа, т.к. нужно записать номера состояний для каждой секунды

Ответ: 1,1,0,0,2,2,1,1,3,3,0,0