

数据集 1 <运动图像，无提示分类器应用程序>

数据集由柏林 BCI 小组提供：柏林理工学院（机器学习实验室）和 Fraunhofer FIRST（智能数据分析小组）（Klaus-Robert Müller、Benjamin Blankertz、Carmen Vidaurre、Guido Nolte）和柏林夏里特医学院神经病学系神经物理学小组（Gabriel Curio）的本杰明·富兰克林校园

与 Benjamin Blankertz 的通信 <benjamin.blankertz@tu-berlin.de>

#### 惊觉刺激

BCI 数据上的大多数算法演示是评估 EEG 试验的分类，即固定长度的窗口化 EEG 信号，其中每个试验对应于特定的精神状态。但是在具有异步反馈的 BCI 应用中，人们面临着一个问题，即分类器必须连续应用于输入的脑电图，而没有提示受试者何时切换她 / 他的意图。该数据集带来了将分类器应用于没有给出提示信息的连续 EEG 的挑战。

此数据集中解决的另一个问题是，评估数据包含用户没有控制意图的时间段。在这些间隔期间，分类器应该返回 0（不隶属于目标类之一）。

#### 实验装置

这些数据是从健康受试者中记录的。在整个会话中，运动意象是在没有反馈的情况下进行的。对于每个主题，从左手、右手和脚三类中选择两类运动意象（由受试者选择一侧；也可以选择双脚）。

校准数据：在前两次运行中，指向左侧、右侧或向下的箭头在计算机屏幕上显示为视觉提示。提示显示 4 秒，在此期间指示受试者执行提示运动意象任务。这些周期与 2 秒的空白屏幕和 2 秒的屏幕中央显示注视叉交错。注视十字叠加在提示上，即显示 6 秒。这些数据提供了完整的标记信息。

评估数据：然后是 4 次运行，用于评估提交给比赛的提交。在这里，运动意象任务由柔和的声学刺激（单词 left、right 和 foot）提示，持续时间在 1.5 到 8 秒之间。运动意象时期的结束由单词 stop 表示。间歇期也有 1.5 到 8 秒的不同持续时间。请注意，在评估数据中，每个条件的试验数量不一定相等。

特点：一些数据集是人工生成的。这个想法是有一种方法来生成具有指定特性的人工脑电图信号，这种信号非常逼真，可以用来评估和比较分析技术。比赛的结果将显示应用的方法在人工数据和真实数据上的表现是否相当。我们提供的唯一信息是，至少有一个真实数据集和至少一个人工数据集，而真实分布情况在提交截止日期之前仍未披露。出于比赛目的，仅考虑真实数据集的结果。生成人工数据的函数由 Guido Nolte 和 Carmen Vidaurre 提供。

#### 数据格式

给定 59 个 EEG 通道的连续信号，对于校准数据，标记

指示提示呈现的时间点和相应的目标类。

数据以 Matlab 格式 (\*.mat) 提供, 其中包含变量:

- cnt: 连续的脑电信号, 大小为 [时间 x 通道]。该数组以数据类型 INT16 存储。要将其转换为 uV 值, 请使用 `cnt = 0.1*double (cnt)` ; 在 Matlab 中。
- mrk: 带字段的 Target cue 信息结构 (评估数据文件不包含此变量)

- o 位置: 以单位样本给出的脑电图信号中提示位置的向量, 长度 #cues

- o y: 目标类的向量 (第一类为 -1, 第二类为 1), 长度 #cues

- nfo: 通过字段提供附加信息的结构

- o fs: 采样率,

- o clab: 通道标签的元胞数组,

- o 类: 运动图像类名称的元胞数组,

- o Xpos: 电极在 2D 投影中的 x 位置,

- o YPOS: 电极在 2D 投影中的 y 位置。

作为替代方法, 数据也以压缩的 ASCII 格式提供:

- \*\_cnt.txt: 连续的脑电图信号, 其中每行保存特定时间点所有通道的值

- \*\_mrk.txt: 目标提示信息, 每行代表一个提示, 其中第一个值定义时间点 (以单位样本为单位), 第二个值定义目标类 (1 表示第一类, 1 表示第二类)。对于评估数据, 不提供 \*\_mrk.txt 文件。

- \*\_nfo.txt: 包含与 MATLAB 格式描述的其他信息。

#### 要求和评估

请提供一个 ASCII 文件 (名为 "Result\_BCIC\_IV\_ds1.txt"), 其中包含评估信号的每个采样点的分类器输出 (-1 和 1 之间的实数), 每行一个值。根据范围从 -1 到 1 的一维光标控制应用程序对提交的内容进行评估。类 1 的心理状态用于将光标定位在 -1, 类 2 的心理状态用于将光标定位在 1 附近。在没有这些心理状态 (间歇间隔) 的情况下, 光标应该在位置 0。请注意, 参赛者不知道受试者在多长时间内处于确定的精神状态。参赛者提交所有时间点的分类器输出。评估函数计算相对于目标向量的平方误差, 第一类为 -1, 第二类为 1, 否则为 0, 跨时间点平均。在平均中, 我们将忽略瞬态期间的时间点 (从每个提示开始 1 秒)。出于比赛目的, 仅考虑真实数据集的结果, 但也报告人工数据的结果以进行比较。(可选) 请报告您认为哪些数据集是人工生成的。

您还必须提供所用算法的描述 (ASCII、HTML 或 PDF 格式), 以便在结果网页上发布。

#### 技术信息

使用 BrainAmp MR plus 放大器和 Ag/AgCl 电极帽进行记录。测量了来自 59 个脑电图位置的信号, 这些信号最密集地分布在感觉运动区域。信号在 0.05 到 200 Hz 之间进行带通滤波, 然后以 1000 Hz 的精度以 16 位 (0.1 uV) 的精度进行数字化。我们还提供了一个以 100 Hz 进行下采样的数据版本 (第一个低通滤波原始数据 (切比雪夫 II 型滤波器

10 阶，阻带纹波下调 50dB，阻带边缘频率 49Hz)，然后计算 10 个样本的块的平均值)。

#### 引用

任何分析此数据集的出版物都应引用以下论文作为录音的参考：

本杰明·布兰克茨、圭多·多恩赫格、马蒂亚斯·克劳莱达特、克劳斯·罗伯特·穆勒和加布里埃尔·库里奥。  
非侵入性 Berlin 脑机接口：在未经训练的受试者中快速获取有效性能。神经影像，37（2）：539-550,2007 年。  
[PDF 格式]