



Анализ данных ЭЭГ (тестовое задание)

Status Not started

Где лежат данные?

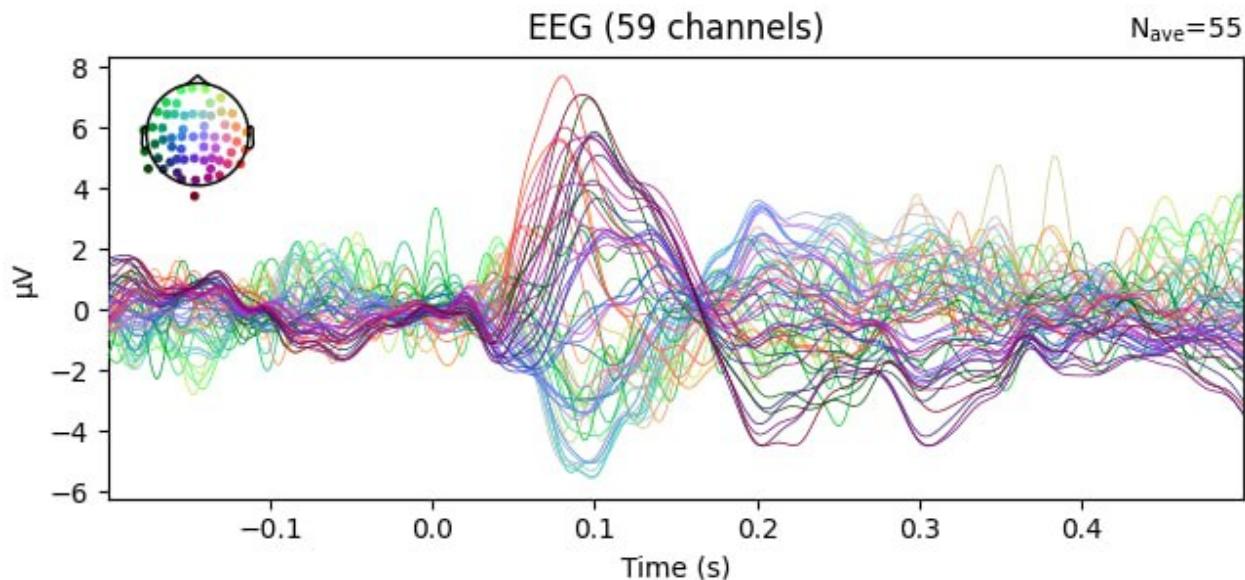
- на ЯД: <https://disk.yandex.ru/d/8gldDixWbvT6Hg>
- в этой папке есть три файла для каждого респондента:
 - 1_RAW DATA** - это файл с сырьими данными с момента, как мы включили запись ЭЭГ до момента, когда мы её выключили. Он нам важен, потому что там есть данные по эмоциональным метрикам (stress, interest, engagement, и т.д.).
 - 1_Восприятие инфоповодов_RAW** - это файл с сырьими данными самого эксперимента (то есть он обрезан с момента, как мы включили сам эксперимент и до момента, когда мы его выключили)
 - 1_Восприятие инфоповодов_intervalMarker.csv** - это файл с маркерами событий, в том числе и самих стимулов.
- ВАЖНО:** для пилотного респондента есть только два файла: **0_RAW DATA** и **0_Восприятие инфоповодов (basic pre-processing)**: второй файл - это файл, полученный с предобработки Emotiv, из него можно взять таймстемпы для сырых данных.

Что нужно сделать?

Задача 1: ERP-графики

Что должно получиться? 24 графика вызванных потенциалов для каждого стимула **для всех респондентов** по всем отведениям

Как это должно выглядеть:



Какие подзадачи могут входить:

1. Предобработка ЭЭГ-сигналов (фильтрация, удаление артефактов, референсирование и т.д.)
2. ERP-обработка
 - Указание интервала усреднения:
 - До стимула (предстимульный интервал, baseline) — 200 мс
 - После стимула — 1000 мс
 - Базовая коррекция по предстимульному интервалу
 - усреднение по **каждому типу стимула**
 - проверка потенциалов по каждому респонденту (в итоговом варианте графики по каждому респонденту не нужны)
3. **визуализация: итоговые графики вызванных потенциалов**
 - в итоговом варианте **нужны только графики по всем респондентам** (т.е. необходимо сделать усреднение по респондентам) **для каждого стимула** (всего 24 стимула: среди них инфоповоды, комментарии, ответы к комментариям на платформах ВКонтакте и Телеграм)

Задача 2: графики по эмоциям

Что нужно сделать? построить по каждому стимулу по всем респондентам графики по эмоциональным метрикам (stress, interest, engagement, attention и т.д.) от ПО Emotiv.

Что нужно взять?

- **сырые данные с emotiv.** Они выглядят так: ***номер респондента*_RAW DATA** - это файл с сырыми данными с момента, как мы включили запись ЭЭГ до момента, когда мы её выключили. Он нам важен, потому что там есть данные по эмоциональным метрикам (stress, interest, engagement, и т.д.).
- **файлы с интервальными маркерами:** ***номер респондента*_Восприятие инфоповодов_intervalMarker.csv** - это файл с маркерами событий, в том числе и самих стимулов.

Далее:

- нужно синхронизировать эти файлы по событиям: в сырых данных нет маркеров по стимулам (в какие моменты времени какие стимулы демонстрировались). Поэтому нужно взять эти маркеры из файлов с интервальными маркерами и внести их в сырье данные.
- усреднить данные по респондентам
- построить **общие графики** по изменению эмоциональных метрик по каждому стимулу: всего получится 24 графика (не забудь взять время до демонстрации стимула, чтобы сравнить изменение эмоциональных метрик с бэйслайном).

Задача со звёздочкой:

- после того, как удастся построить общие графики, таким же образом нужно построить **графики по полу:**

1.1. то есть 24 графика по женщинам. Номера респондентов-женщин: 1, 7, 9, 11, 12, 13, 15, 16

1.2. и 24 графика по мужчинам. Номера респондентов-мужчин: пилот, 2, 3, 4, 5, 6, 8, 10, 14

Таймлайн (порядок стимулов на графике):

1. VK_JAPAN_INFO
2. VK_JAPAN_COM
3. VK_JAPAN_THR
4. TG_JAPAN_INFO
5. TG_JAPAN_COM
6. TG_JAPAN_THR
7. TG_MUSK_INFO
8. TG_MUSK_COM
9. TG_MUSK_THR
10. VK_MUSK_INFO
11. VK_MUSK_COM
12. VK_MUSK_THR
13. VK_BORISOV_INFO
14. VK_BORISOV_COM
15. VK_BORISOV_THR
16. TG_BORISOV_INFO
17. TG_BORISOV_COM
18. TG_BORISOV_THR
19. TG_EGE_INFO
20. TG_EGE_COM

Как должен выглядеть результат?

Полученные графики загрузите на диск и откройте доступ к файлу по ссылке.

Графики могут быть в формате (pdf/png/jpeg). Каждый график - это отдельный файл/документ.

Полученный результат следует сопроводить документом в формате word с техническим описанием процесса работы с данными, возможными багами и

их решением. В документ нужно вставить ссылку на github на код обработки данных. Предварительно проверьте доступ по ссылке.

Результаты предобработки сырых данных нужно добавить отдельным документом в формате xlsx таблицы.

Дедлайн: два рабочих дня с момента получения ТЗ, 18:00 по мск