

Аннотация

Анализ электроэнцефалографии (ЭЭГ) является важным инструментом в нейробиологии, нейроинженерии (например, интерфейсы мозг - компьютер (ИМК)), а также имеет промышленное применение. С развитием этих областей появляется всё большая потребность в анализе ЭЭГ сигналов. Получение информации об электрической активности зон головного мозга и о том, каким образом эта активность влияет на классификацию типа решаемой человеком задачи является важным шагом на пути к тому, чтобы извлекать из ЭЭГ больше информации о функционировании мозга. Для достижения данной цели на имеющихся ЭЭГ данных мы рассмотрели каждый электрод в отдельности, чтобы ответить на следующие вопросы: (1) Насколько активность каждого электрода в отдельности влияет на результат? (2) Какие электроды наибольшим образом влияют на предсказание типа решаемой задачи? Затем была выдвинута следующая гипотеза: рассматривая электроды отдельно друг от друга, можно получить результат, точность которого $> 60\%$.

В ходе данной работы, используя методы машинного обучения, была получена информация о том, как каждый электрод в отдельности влияет на результат, и установлено, какие из электродов наибольшим образом влияют на качество натренированного классификатора. А также, получив среднее p -value по всем электродам $p_{val} \sim 0.064$, сделан вывод о том, что выдвинутую гипотезу следует отвергнуть.

Abstract

Electroencephalography (EEG) analysis has been an important tool in neuroscience with applications in neuroscience, neural engineering (e.g. Brain-computer interfaces, BCI's), and even commercial applications. With the development of these areas, there is an increasing need for the analysis of EEG signals. Obtaining information about the electrical activity of brain areas and how this activity affects the classification of the type of task being solved by a person is an important step towards extracting more information about the functioning of the brain from the EEG. To achieve this goal, based on the available EEG data, we examined each electrode separately to answer the following questions: (1) How much does the activity of each electrode individually affect the result? (2) Which electrodes have the greatest influence on the prediction of the type of problem being solved? Then the following hypothesis was put forward: considering the electrodes separately from each other, it is possible to obtain a result with an accuracy more than 60%. Using machine learning methods, information was obtained on how each electrode individually affects the result, and it was established which of the electrodes most affect the quality of the trained classifier. And also, having obtained the average p -value for all electrodes $p_{val} \sim 0.064$, it was concluded that the hypothesis put forward should be rejected.

Содержание

1	Введение	4
2	Использованные данные	6
2.1	Теоретическое введение	6
2.2	Процедура снятия ЭЭГ данных	7
3	Формат данных	9
3.1	Проблемы при регистрации ЭЭГ	9
3.2	Предварительная обработка данных	9
3.3	Вид полученных данных	10
4	Постановка задачи	12
4.1	Задача обучения классификатора	12
4.2	Цели и задачи	12
4.3	Выбор метода	12

1 Введение

Электроэнцефалография (ЭЭГ) — метод исследования головного мозга, основанный на регистрации его электрических потенциалов. ЭЭГ измеряет колебания напряжения в результате ионного тока в нейронах головного мозга. Клинически электроэнцефалограмма является графическим изображением спонтанной электрической активности мозга в течение определенного периода времени, записанной с нескольких электродов мозга или поверхности скальпа [1], (т.е. каждый электрод соответствует определённой области мозга).

ЭЭГ широко используется в исследованиях, связанных с нейронной инженерией, неврологией и биомедицинской инженерией (например, интерфейсы мозг-компьютер [2], анализ сна [3], обнаружение приступов эпилепсии [4]) из-за её неинвазивности и относительно низкой финансовой стоимости. Автоматическая классификация этих сигналов является важным шагом на пути к тому, чтобы сделать использование ЭЭГ более практичным в применении и менее зависимым от подготовленных специалистов. Типичный процесс подготовки классификации ЭЭГ включает в себя удаление глазодвигательных и мышечных артефактов, отбор признаков и классификацию.

На самом базовом уровне набор данных ЭЭГ состоит из объектов — векторов действительных значений, которые представляют генерируемые мозгом потенциалы, снятые с кожи головы. Размерность каждого такого вектора, характеризуется числом электродов и количеством диапазонов спектра каждого электрода.

На данный момент существует большое количество информации о применении традиционных алгоритмов машинного обучения для распознавания типа решаемой человеком задачи (например, [5], [6]), в которых для предсказания учитываются данные всех электродов в совокупности. В то же время, исследований, в которых электроды рассматриваются по отдельности, не достаточно для того, чтобы располагать полной информацией об электрической активности мозга.

В данной работе электроды были рассмотрены отдельно друг от друга, чтобы ответить на следующие вопросы: (1) Как активность каждого электрода (и соответствующей ему зоны мозга) в отдельности влияет на результат классификации типа решаемой задачи? (2) Какие электроды наибольшим образом влияют на результат? (3) Можно ли получить результат, точность которого выше 60%, рассматривая таким образом электроды?

Для получения ответов на вышеперечисленные вопросы были использованы стратегии для классификации ЭЭГ с использованием линейных методов машинного обучения, а также методы предварительной обработки данных ЭЭГ. Полученная информация может послужить отправной точкой для начального этапа проектирования архитектуры в будущих приложениях машинного обучения для классификации ЭЭГ.

2 Используемые данные

2.1 Теоретическое введение

Тест Стернберга — классический эксперимент, проведенный в 1966 году психологом Солом Стернбергом, позволивший сделать вывод о том, что информация извлекается из кратковременной памяти путём последовательного исчерпывающего сканирования. Оригинальные и модифицированные схемы теста (Sternberg item recognition paradigm, SIRP), описанного в статье [7], используются для изучения особенностей кратковременной и рабочей памяти.

Оригинальный тест состоял из 24 тренировочных и 144 тестовых проб. В каждой испытуемым предъявлялся случайный набор цифр, который требовалось запомнить. Длина набора варьировалась от 1 до 6 цифр, каждая из которых предъявлялась отдельно в течение 1-2 секунды. После этого следовала пауза длиной в 2 секунды, а за ней контрольная цифра. Испытуемые должны были потянуть один из двух рычагов в качестве ответа «Да, это одна из запомненных цифр» или «Нет, это новая цифра» (требовавшиеся с одинаковой вероятностью), после чего контрольный стимул исчезал, а загоравшаяся лампочка давала обратную связь о правильности ответа. В конце испытуемых также просили произнести запомненную последовательность (схема экспериментальной установки изображена на рис. 1).

Подробный сценарий каждой тестовой пробы:

1. предупреждающий сигнал;
2. запоминаемый стимул-образец;
3. отсрочка;
4. предупреждающий сигнал;
5. контрольный стимул;
6. ответ испытуемого;
7. обратная связь о правильности выполнения задания;
8. просьба к испытуемому вспомнить стимул-образец, после чего начинается следующая проба.

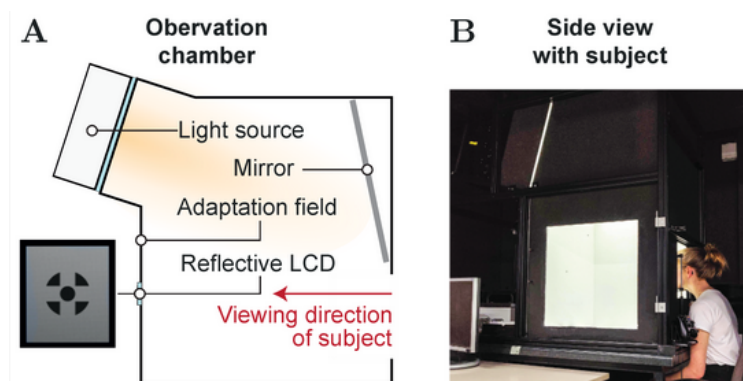


Рисунок 1 — Экспериментальная установка. (А) Схематическое изображение камеры наблюдения и направления обзора. (В) Изображение камеры наблюдения с видом сбоку (дверь для обслуживания снята), на котором изображен объект, фокусирующий цель фиксации, которая отображается на отражающем жидкокристаллическом дисплее, встроенном в центр задней стенки камеры. Подставка для подбородка используется для соответствующего позиционирования и выравнивания головы субъекта.[8]

Парадигма Стернберга была успешно использована в рамках изучения индивидуальных различий в процессах памяти у здоровых участников [9], [10], в исследованиях посвященных изучению дефицитарности и изменений кратковременной памяти при старении [11], в исследованиях шизофрении и болезни Альцгеймера [12], депрессии [13], множественного склероза [14], в исследованиях изучающих воздействия различных медикаментов на процессы памяти [15].

2.2 Процедура снятия ЭЭГ данных

В настоящем исследовании для получения ЭЭГ данных использовался "Тест Стернберга". В эксперименте на голову участника исследования одевается специальная шапочка со специальными металлическими электродами, которые регистрирует биоэлектрическую активность мозга (см. рис. 2). Каждый электрод с какой-то частотой (обычно 500-1000 Гц) регистрирует амплитуду колебаний электромагнитной активности изменения электрического потенциала с поверхности головы.

Затем участнику эксперимента последовательно предъявляются наборы цифр. В каждом наборе цифры от 1 до 9 (без повторений одной цифры дважды) были представлены в случайной последовательности. При этом размеры наборов могут отличаться (в оригинальных исследованиях Стернберга, как было описано выше, это наборы объемом от 1 до 6 цифр). В зависимости от того, какой длины последовательность была предъявлена участнику, определялся тип решаемой задачи: лёгкая, средняя, повышенной сложности и тяжёлая (3, 4, 5 или 6 и более цифр для запоминания соответственно)— всего 4 типа задачи. Цифры показываются одна за другой, участнику необходимо запомнить их последовательность. После контрольного сигнала (появление определенной цифры на экране) участнику необходимо как можно быстрее ответить, присутствовала ли цифра в предъявленном до этого наборе. После этого участника просят вспомнить порядок представления цифр, для того чтобы убедиться, что он действительно запомнил последовательность.



Рисунок 2 — Расположение электродов на поверхности головы

Перед предъявлением набора стимулов участникам эксперимента предъявляется фиксационный крест на 1 секунду. Предъявление набора стимулов начинается через 1.2 секунды после предъявления фиксационного креста и длится 1.5 секунды. Через 2 секунды после окончания предъявления набора предъявляется тестовый стимул (задача участника - сказать был ли тестовый стимул в наборе). Тестовый стимул предъявляется на 2 секунды. С момента начала предъявления тестового стимула участник может давать ответ.

В исследовании принимал участие 101 человек. Каждому человеку предлагалось для решения порядка 30 задач на каждый из 4 типов.

3 Формат данных

3.1 Проблемы при регистрации ЭЭГ

После того, как ЭЭГ снято, необходимо отфильтровать сигнал для получения более точных показателей. Сигнал плохо виден на фоне шума, который создают различные артефакты. В связи с этим возникает проблема обработки записи — удаление артефактов для получения более точных показателей.

При регистрации ЭЭГ артефактом является любая активность, не связанная с электрической активностью мозга. Все артефакты и помехи при регистрации ЭЭГ могут быть разделены на две большие группы:

1. аппаратные артефакты и внешние помехи физической природы;
2. физиологические артефакты, регистрируемые от больного.

Наиболее частыми являются электродные артефакты и физиологические артефакты, связанные с регистрацией потенциалов, возникающих при моргании и движении глаз, миографических потенциалов при мышечном напряжении.

3.2 Предварительная обработка данных

Для удаления артефактов был использованный автоматический метод, реализованный для среды обработки ЭЭГ (MNE) [16]. После предварительной очистки данные ЭЭГ были отфильтрованы в диапазоне частот 1-30 Гц. Далее с помощью метода Уэлча для стандартных узких частотных диапазонов ЭЭГ

- δ : 1-4 Гц
- θ : 4-8 Гц
- α_1 : 8-10 Гц
- α_2 : 10-13 Гц
- β_1 : 13-20 Гц
- β_2 : 20-30 Гц

спектральная мощность сигнала была проанализирована отдельно для тестов Стернберга разной сложности (3, 4, 5 или 6 цифр и более для запоминания).

После удаления артефактов данные были получены в виде набора .csv файлов. Каждый файл содержал в себе данные, относящиеся к конкретному человеку и типу задачи (см. главу 2.2), то есть всего 404 файла. Данные из файлов были объединены в один набор данных, пригодных для обработки алгоритмами машинного обучения.

3.3 Вид полученных данных

Полученный набор данных устроен следующим образом (см. рис. 3)

	1_1	1_2	1_3	1_4	...	63_5	63_6	name	task
0	2.148821e-12	1.762882e-12	1.459278e-13	1.587917e-13	...	6.211839e-13	8.285990e-13	chcon_s_100	0
1	1.270260e-12	8.164477e-13	5.903729e-13	3.056763e-13	...	1.079865e-12	2.844522e-12	chcon_s_100	0
2	7.709517e-13	1.717117e-13	3.039963e-13	1.663584e-13	...	1.407080e-12	4.519942e-12	chcon_s_100	0
3	2.800220e-12	3.611335e-13	9.828645e-13	3.723583e-13	...	6.247119e-13	1.532856e-12	chcon_s_100	0
4	9.507999e-13	4.833313e-13	1.694255e-12	1.434799e-13	...	1.444978e-12	2.585149e-12	chcon_s_100	0
...
10638	2.783871e-12	4.492069e-13	6.107854e-13	2.169733e-13	...	5.649186e-14	1.532088e-13	mcon_str_41	3
10639	1.394714e-12	2.106615e-12	2.131392e-12	1.251115e-12	...	1.013557e-12	8.060742e-13	mcon_str_41	3
10640	3.272053e-12	2.410372e-12	3.519692e-12	3.088923e-13	...	3.932170e-13	6.883666e-13	mcon_str_41	3
10641	1.226257e-11	3.028533e-12	6.798046e-13	3.032247e-13	...	5.681097e-13	6.363103e-12	mcon_str_41	3
10642	8.328733e-11	2.666764e-12	6.353130e-12	8.844715e-13	...	9.100197e-13	1.183650e-12	mcon_str_41	3

10643 rows x 380 columns

Рисунок 3 — Полученный набор данных ЭЭГ

Каждый объект выборки — вектор размерности $63 * 6 + 1 + 1 = 380$, в котором первые $63 * 6 = 378$ столбцов есть амплитуды колебаний электромагнитной активности, которые регистрировались каждым из 63 электродов (каждый электрод в свою очередь характеризуется шестью частотными диапазонами спектра). Также есть столбец "name", содержащий идентификатор участника эксперимента, и столбец "task" с соответствующим типом задачи, которую решал участник во время снятия ЭЭГ данных. Каждому из 4 типов задачи (лёгкая, средняя, повышенной сложности и трудная) соответствует число от 0 до 3.

Всего 101 участником эксперимента было решено от 22 до 32 задач каждого из 4 типов сложности. Суммарное количество объектов в выборке составляет 10641.

Рассмотрим объект выборки более подробно (см. рис. 4):

- **Красным цветом** выделены амплитуды колебаний электромагнитной активности электродов (каждый из 63 характеризуется шестью частотными диапазонами). В каждой ячейке значением является число;
- **Зелёным цветом** — идентификатор участника;
- **Синим цветом** — тип задачи (число от 0 до 3).

	1_1	1_2	1_3	1_4	...	63_5	63_6	name	task
0	2.148821e-12	1.762882e-12	1.459278e-13	1.587917e-13	...	6.211839e-13	8.285990e-13	chcon_s_100	0

Рисунок 4 — Детальный вид объекта выборки

4 Постановка задачи

4.1 Задача обучения классификатора

4.2 Цели и задачи

4.3 Выбор метода

Список литературы

- [1] *H, Berger. Ueber das Elektroenkephalogramm des Menschen / Berger H // Arch. Psychiat. Nervenkr. — 1929. — Pp. 527–570.*
- [2] *He Y Eguren D, Azorin J M. Brain-machine interfaces for controlling lower-limb powered robotic systems / Azorin J M He Y, Eguren D // J.Neural. — 2018. — no. 15.*
- [3] *Motamedi-Fakhr S Moshrefi-Torbati M, Hill M Etc. Signal processing techniques applied to human sleep EEG signals – a review / Hill M Etc. Motamedi-Fakhr S, Moshrefi-Torbati M // Biomed. Signal Process. Control. — 2014. — Pp. 21–33.*
- [4] *G, Chen. Automatic EEG seizure detection using dual-tree complex wavelet-Fourier features / Chen G // Expert Syst. Appl. — 2014.*
- [5] *Hanjie Liu Jinren Zhang, Qingshan Liu Etc. Minimum spanning tree based graph neural network for emotion classification using EEG / Qingshan Liu Etc. Hanjie Liu, Jinren Zhang // Neural Networks. — 2022. — Vol. 145. — Pp. 308–318.*
- [6] *Alexander Craik, Yongtian He Etc. Deep learning for electroencephalogram (EEG) classification tasks: a review / Yongtian He Etc. Alexander Craik // Journal of Neural Engineering. — 2019. — Vol. 16, no. 031001.*
- [7] *Sternberg, Saul. High-Speed Scanning in Human Memory / Saul Sternberg // Science. — 1966. — Vol. 153, no. 3736. — Pp. 652–654.*
- [8] *Julian Klabes Sebastian Babilon, Babak Zandi Etc. The Sternberg Paradigm: Correcting Encoding Latencies in Visual and Auditory Test Designs / Babak Zandi Etc. Julian Klabes, Sebastian Babilon // Vision. — 2021. — Vol. 5(2), no. 5020021.*
- [9] *Etc., Jensen. Oscillations in the alpha band 9-12 Hz increase with memory load during retention in a short-term memory task / Jensen Etc. // Cereb Cortex. — 2002. — no. 12(8).*
- [10] *Wolach. The mode of short-term memory encoding as indicated by event-related potentials in a memory scanning task with distractions / Wolach, Pratt // Clinical neurophysiology. — 2001. — no. 112(1).*

- [11] *Etc., Bart Rypma*. Dissociating Age-related Changes in Cognitive Strategy and Neural Efficiency Using Event-related fMRI / Bart Rypma Etc. // *Cortex*. — 2005. — Vol. 41. — Pp. 582–594.
- [12] *Corbin*. Is Sternberg’s Memory Scanning Task Really a Short-Term Memory Task? / Corbin, Marquer // *Swiss Journal of Psychology*. — 2013. — Vol. 72(4), no. 181.
- [13] *Slade, Blumhardt*. Working memory dysfunction in major depression: An event-related potential study / Blumhardt Slade, Sharma // *Clinical neurophysiology*. — 2000. — Vol. 111(9), no. 1531-43.
- [14] *Catherine J Archibald 1 Xingchang Wei, James N Scott Etc*. Posterior fossa lesion volume and slowed information processing in multiple sclerosis / James N Scott Etc. Catherine J Archibald 1, Xingchang Wei // *Comparative Study*. — 2004. — Vol. 127(7), no. 1526-34.
- [15] *Verster J C Volkerts E R, Verbaten M N*. Effects of alprazolam on driving ability, memory functioning and psychomotor performance: a randomized, placebo-controlled study. / Verbaten M N Verster J C, Volkerts E R // *Neuropsychopharmacology*. — 2002. — Vol. 27. — Pp. 260–269.
- [16] *Mainak Jas Denis A. Engemann, Yousra Bekhti Etc*. Autoreject: Automated artifact rejection for MEG and EEG data / Yousra Bekhti Etc. Mainak Jas, Denis A. Engemann // *Neuroimage*. — 2017. — no. 159. — Pp. 417–429.
- [17] *Морз, Р.* Бесстолкновительный PIC-метод / Р. Морз // *Вычислительные методы в физике плазмы* / Ed. by Б. Олдера, С. Фернбаха, М. Ротенберга. — М.: Мир, 1974.
- [18] *Киселёв, А. А.* Численное моделирование захвата ионов бесстолкновительной плазмы электрическим полем поглощающей сферы / А. А. Киселёв, Долгонос М. С., Красовский В. Л. // *Девятая ежегодная конференция «Физика плазмы в Солнечной системе»*. — 2014.