

# Поиск зависимостей в биомеханических системах

*Зыков Т.А.<sup>1</sup>, Дорин Д.Д.<sup>2</sup>, Стрижов В.В.<sup>3</sup>*

zykov.ta@phystech.edu

<sup>1</sup>Chair of Data analysis; <sup>2</sup>, <sup>3</sup>Intelligent systems;

Исследуется проблема восстановления зависимости между показаниями датчиков фМРТ и восприятием внешнего мира человеком. Проводится анализ зависимости между последовательностью снимков фМРТ и звуковым рядом. Требуется предложить метод прогнозирования показаний фМРТ по прослушиваемому звуковому ряду. При прогнозировании сложноорганизованных временных рядов, зависящих от экзогенных факторов и имеющих множественную периодичность, связи между рядами устанавливаются с помощью теста Гренджера.

## 1 Введение

Функциональная магнитно-резонансная томография (фМРТ) [9] — метод нейровизуализации, который измеряет активность мозга путем выявления изменений, связанных с кровотоком. Метод нейровизуализации фМРТ находит применение для анализа корреляции активности мозга с такими заболеваниями как аутизм и болезнь Альцгеймера, а также для прогнозирования таких заболеваний как черепно-мозговая травма и имеет потенциал для терапии [2]. Он дает представление о том, какие области мозга участвуют в конкретных психических процессах или задачах, измеряя уровень BOLD (Blood-oxygen-level-dependent). BOLD — это показатель, используемый в функциональной магнитно-резонансной томографии, который отражает изменения кровотока и уровня насыщения кислородом в мозге [4].

Работа посвящена восстановлению зависимости между снимками фМРТ и звуковым рядом. Предполагается, что такая зависимость существует. Кроме того, предполагается, что между снимком и звуковым рядом есть постоянная задержка во времени [7]. Время задержки выступает в качестве гиперпараметра модели. Предлагается метод аппроксимации показаний фМРТ по прослушиваемому звуковому ряду. Проводится первичный анализ влияния типа звукового ряда (музыка или простой диалог) на изменения уровня BOLD и анализ зависимости снимков фМРТ от звукового ряда.

В данной работе рассмотрим одно из ограничений BOLD [7]. Временное разрешение — это наименьший период времени нейронной активности который с высокой точностью можно определить с помощью фМРТ. Время измерений фМРТ происходит с задержкой, что затрудняет регистрацию быстрых нейронных событий. Кроме того, при выборе модели важно учитывать структуру временных рядов: у некоторых вокселей прослеживается тренд на протяжении всей звуковой дорожки или неоднородность шума, при которых нужно использовать оценки Уайта [10].

Для извлечения признаков из звуковой дорожки предлагается использовать разные подходы [5]. В этой работе рассматриваются мел-кепстральные коэффициенты (MFCC). Это характеристики аудиосигнала которые получаются из спектрограммы и используются для анализа звука и обработки речи. Преимуществом такого представления является небольшая размерность.

В качестве выборки будем использовать данные [1], собранные у большой группы людей во время просмотра короткометражного аудиовизуального фильма. Он включает записи фМРТ 30 участников в возрасте от 7 до 47 лет. Данные были получены с использо-

ванием богатого аудиовизуального стимула. Звуковая дорожка состоит из диалогов персонажей и музыкальных вставок.

## 2 Постановка задачи

Требуется предложить метод прогнозирования показаний фМРТ по прослушиваемому звуковому ряду. Обозначим частоту снимков фМРТ  $\mu \in \mathbb{R}$ . Задана последовательность снимков

$$\mathbf{S} = [\mathbf{s}_1, \dots, \mathbf{s}_{\mu t}], \quad \mathbf{s}_\ell \in \mathbb{R}^{X \times Y \times Z}, \quad (1)$$

где  $X, Y$  и  $Z$  — размерности воксельного изображения.

Задана частота дискретизации  $\nu \in \mathbb{R}$ , количество каналов  $k \in \mathbb{N}$  и продолжительность  $t \in \mathbb{R}$  аудиоряда. Задан непрерывный по времени сигнал

$$\mathbf{P} = [\mathbf{p}_1, \dots, \mathbf{p}_{\nu t}], \quad \mathbf{p}_\ell = \begin{pmatrix} p_\ell^1 \\ p_\ell^2 \\ \vdots \\ p_\ell^k \end{pmatrix}, \quad p_\ell^k \in \mathbb{R}, \quad (2)$$

Задача состоит в построении отображения, которое бы учитывало задержку  $\Delta t$  между снимком фМРТ и аудиорядом, а также предыдущие томографические показания. Формально, необходимо найти такое отображение  $\mathbf{f}$ , что

$$\mathbf{f}(\mathbf{p}_1, \dots, \mathbf{p}_{k_\ell - \nu \Delta t}; \mathbf{s}_1, \dots, \mathbf{s}_{\ell-1}) = \mathbf{s}_\ell, \quad \ell = 1, \dots, \mu t, \quad (3)$$

где для  $\ell$ -го снимка фМРТ номер соответствующего сигнала  $k_\ell$  определяется по формуле

$$k_\ell = t\nu = \frac{\ell}{\mu}. \quad (4)$$

Эмбедингами аудиоряда будут мел-кепстральные коэффициенты [8]. То есть для каждого экземпляра сигнала имеем вектор размерности  $\mathbf{d}$ :

$$\mathbf{x}_\ell = [x_1^\ell, \dots, x_d^\ell]^\top \in \mathbb{R}^d, \quad \ell = 1, \dots, \frac{\nu t}{h}. \quad (5)$$

Будем восстанавливать функцию  $\mathbf{f}$ , в предположении марковского свойства.

$$\mathbf{f}(\mathbf{x}_{k_\ell - \nu \Delta t - g}, \dots, \mathbf{x}_{k_\ell - \nu \Delta t}) = \mathbf{s}_\ell - \mathbf{s}_{\ell-1} = \boldsymbol{\delta}_\ell \quad \ell = 2, \dots, \mu t, \quad (6)$$

где  $\boldsymbol{\delta}_\ell = [s_{ijk}^\ell - s_{ijk}^{\ell-1}] = [\delta_{ijk}^\ell] \in \mathbb{R}^{X \times Y \times Z}$  — разность между двумя последовательными снимками.

Учитывая (4), суммарное число пар (сигнал, снимок) равно  $N = \mu(t - \Delta t)$ . Таким образом, для каждого вокселя задана выборка

$$\mathfrak{D}_{ijk} = \{(\mathbf{x}_\ell, \delta_{ijk}^\ell) \mid \ell = 2, \dots, N\}.$$

Поставлена элементарная задача восстановления регрессии

$$y_{ijk} : \mathbb{R}^d \rightarrow \mathbb{R}. \quad (7)$$

Рассмотрим каждый воксель независимо  $\mathbf{Y}_{ijk} \in \mathbb{R}^N$  — воксели,  $\mathbf{X} \in \mathbb{R}^{N \times d}$ . Предполагаемая зависимость

$$\mathbf{Y}_{ijk} = \mathbf{X}\theta + \varepsilon, \quad (8)$$

где  $\theta \in \mathbb{R}^d$  - коэффициенты модели,  $\varepsilon \sim N(0, \Sigma)$  - шум.

Требуется найти параметры  $\hat{\theta}$ , доставляющие максимум функции правдоподобия при заданных гиперпараметрах  $\Delta t$  и  $d$ , где  $d$  – размерность MFCC:

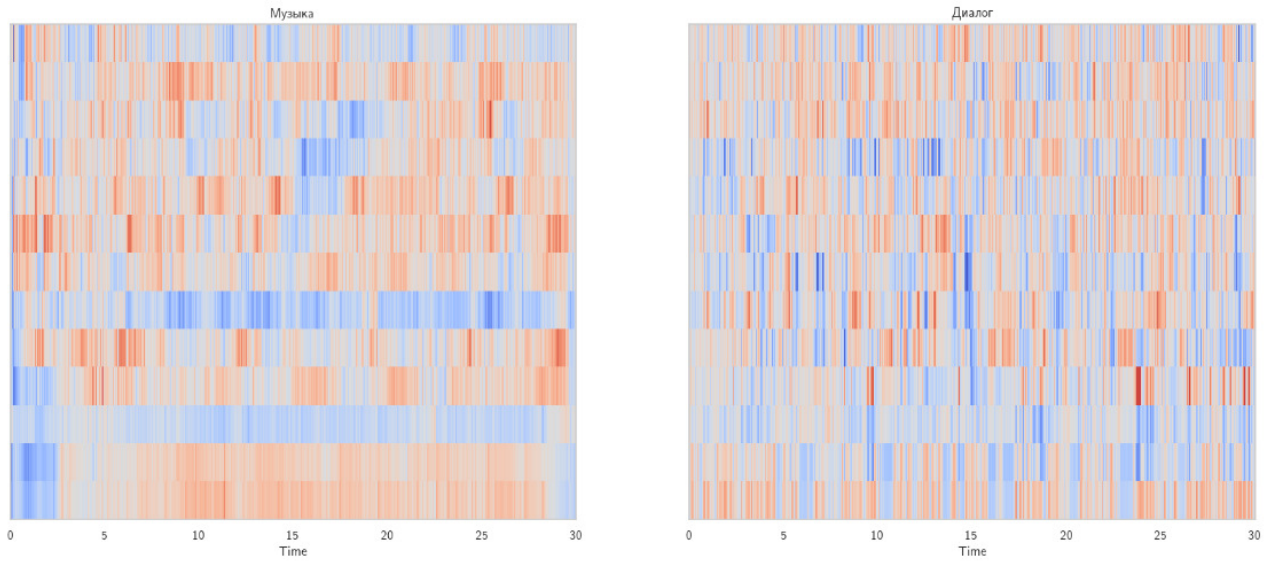
$$L_X(\theta) = \prod_{v=1}^N p_{\theta}(Y_{ijk}^v) \longrightarrow \max_{\theta} \quad (9)$$

### 3 Предварительные сведения

Снимки фМРТ были отнормированы на отрезок  $[0, 1]$  с чисткой от некоторых выбросов. В качестве размерности MFCC взят  $d = 15$ . Общие характеристики обработанного датасета представлены в таблице 1.

В снимках фМРТ присутствуют воксели с неоднородным шумом и/или с линейным трендом.

Для большей информативности входа для предсказания следующего вокселя, вместо одного вектора MFCC подавалось усреднение по окну размера 50. Также добавлена константа (intercept) к данным.



**Рис. 1** Коэффициенты MFCC

**Таблица 1** Описание датасета

Характеристика	Значение
длина короткометражного фильма	389 сек.
частота снимков фМРТ	$641/389 \approx 1.65$ Гц
частота звукового ряда	$\mu$
звуковой ряд	$[-1, 1]^{2 \times 390 \times 44100}$
снимок фМРТ	$[0, 1]^{40 \times 64 \times 64}$

## 4 Вычислительный эксперимент

Ниже приведен пример работы алгоритма на снимках седьмого испытуемого. Все срезы по 20-й координате.

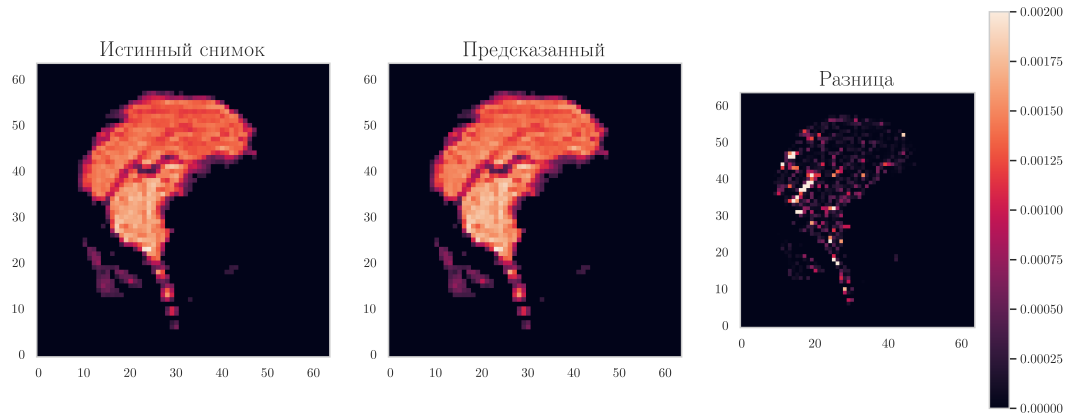


Рис. 2 Пример работы алгоритма

Оценивать качество модели на предсказанных следующих по порядку снимках некорректно. Снимки отличаются на очень малую величину. Обучим вторую модель, одинаковую по архитектуре, на неинформативном входе (случайный тензор) и сравним его в рекурсивной стратегии предсказания.

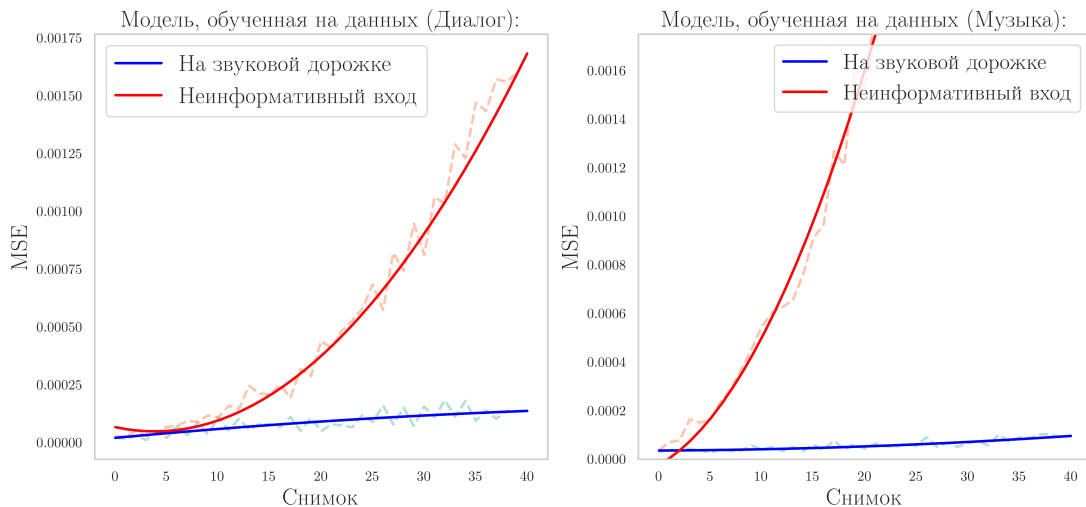


Рис. 3 Сравнение MSE моделей

Заметим, что наша модель работает намного лучше и меньше накапливает ошибку. Следовательно, звуковая дорожка имеет свой эффект на изменение вокселей фМРТ.

Распределение весов модели, которое представлено в рис. 4, невырождено. Это происходит из-за того, что ряды вокселей похожи на стационарные.

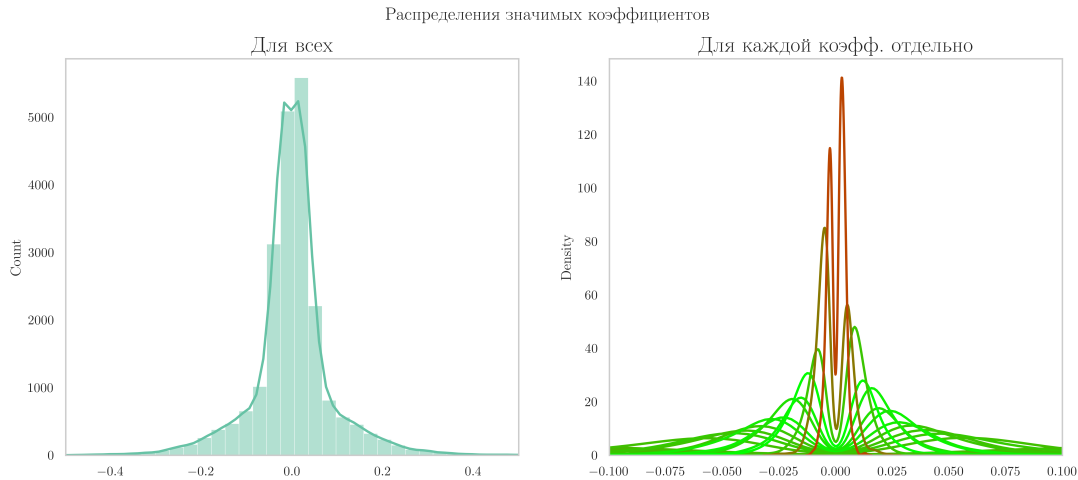


Рис. 4 Распределение весов

## 5 Анализ решения

Очевидно, звук оказывает влияние не на все области мозга. Выделим воксели, на которые звуковой ряд влияет по Грэнджеру [6]. Важно помнить, что причинность по Грэнджеру является необходимым, но не достаточным условием причинно-следственной связи.

Зафиксируем воксель  $i$ . Пусть  $X$  – временной ряд звуковых сигналов,  $Y$  – ряд значений вокселя  $i$  на снимках фМРТ. Строится модель регрессии:

$$Y_t = a_0 + a_1 Y_{t-1} + \dots + a_p Y_{t-p} + b_1 X_{t-1} + \dots + b_p X_{t-p} + \varepsilon_t \quad (10)$$

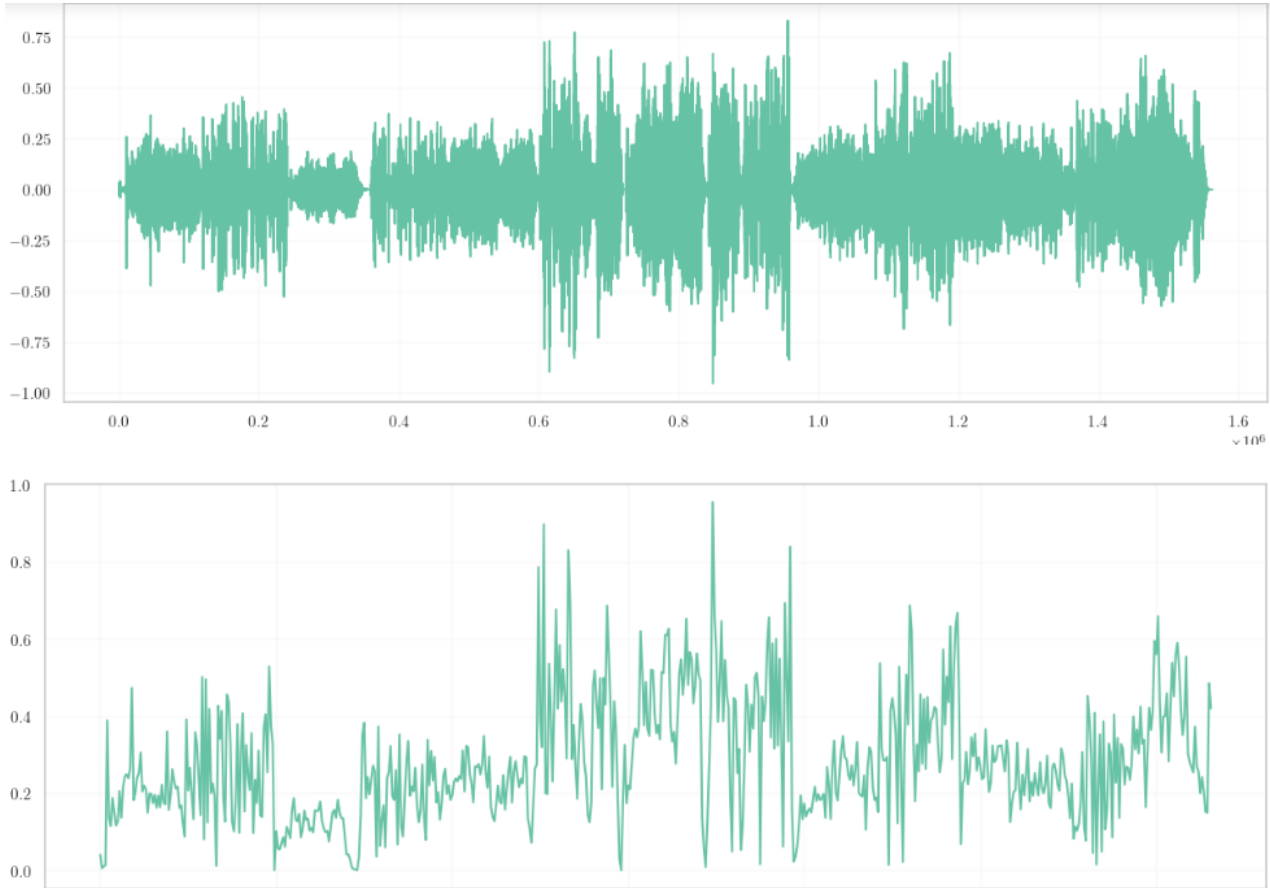
Нулевая гипотеза заключается в том, что коэффициенты при лагах второй переменной одновременно равны нулю. Данную гипотезу предлагается проверить простым критерием Фишера [10].

Рассмотрим пересечение вокселей, у которых  $r$ value меньше 0.3, и височных долей мозга, согласно исследованиям [3]. Пример вокселя, у которого околонулевое  $r$ value на рис. 5.

Оценим ошибку на этих вокселях, так как остальная часть считается шумовой и, предположительно, не является следствием по Грэнджеру. Наименьшая скорректированная MSE достигается при задержке 4-5 секунд, что согласуется с данными, указанными в датасете.

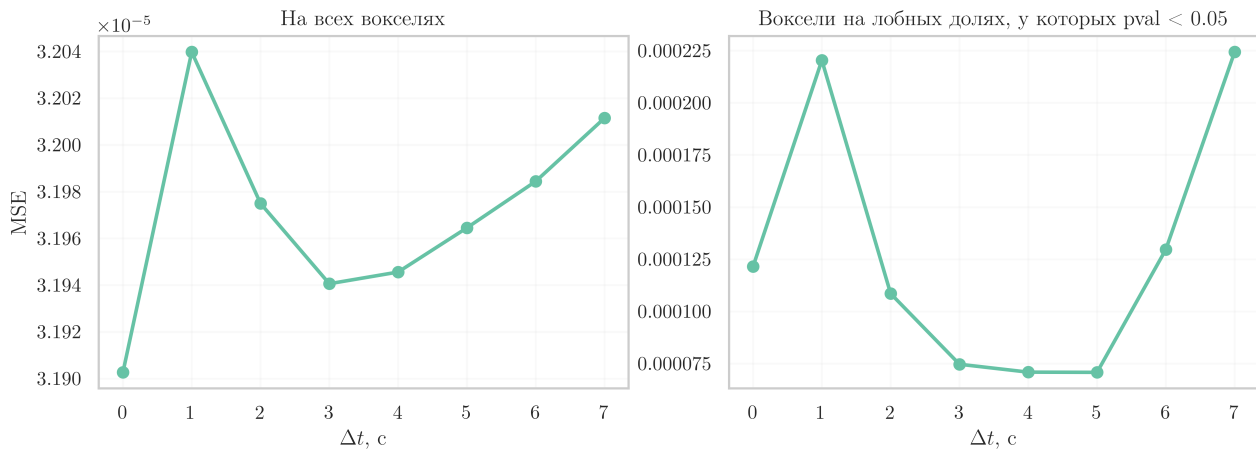
## 6 Заключение

В данной работе рассматривается задача восстановления зависимости между показаниями фМРТ и звуковым рядом. Предложен метод аппроксимации последовательности фМРТ снимков с помощью звукового ряда. Метод учитывает временное разрешение снимков. Линейная модель строится независимо для каждого вокселя изображения фМРТ. Каждая линейная модель строится в предположении, что последовательность изображений фМРТ является марковской. Предложенная модель работает лучше чем неинформативная модель. Следовательно, звуковая дорожка влияет на изменения вокселей. Распределения весов модели невырожденные. Корректировка на воксели, которые предположительно являются следствием по Грэнджеру, уменьшают шум и MSE согласуется с временным разрешением в 4-5 секунд для испытуемого.



**Рис. 5** Воксель, который предположительно зависит от звукового ряда

Зависимость MSE от времени задержки



**Рис. 6** Скорректированный MSE

## 7 Литература

- [1] Julia Berezutskaya, Mariska J. Vansteensel, Erik J. Aarnoutse, Zachary V. Freudenburg, Giovanni Piantoni, Mariana P. Branco, and Nick F. Ramsey. Open multimodal iEEG-fMRI dataset from naturalistic stimulation with a short audiovisual film. *Scientific Data*,

9(1), March 2022.

- [2] Rose Dawn Bharath. Functional mri: Genesis, state of the art and the sequel. *doi: 10.4103/0971-3026.130684*.
- [3] Ioan-Alexandru Florian Corneliu Toader, Calin Petru Tataru. Cognitive crescendo: How music shapes the brain's structure and function. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10605363/>.
- [4] Marcus E. Raichle. The brain's dark energy. *Scientific American* 302 (3): 44–49, 2010.
- [5] Sridhar Krishnan Garima Sharma, Kartikeyan Umapathy. Trends in audio signal feature extraction methods. <https://www.sciencedirect.com/science/article/pii/S0003682X19308795>, 2020.
- [6] C.W.J. Granger. Testing for causality: A personal viewpoint. *Journal of Economic Dynamics and Control*, pages 329–352, 1980.
- [7] Ravi S Menon and Seong-Gi Kim. Spatial and temporal limits in cognitive neuroimaging with fmri. *Trends in cognitive sciences*, 3(6):207–216, 1999.
- [8] S. S. Stevens; J. Volkman; E. B. Newman. A scale for the measurement of the psychological magnitude pitch.
- [9] Ju V Puras and EV Grigorieva. The neurovisualization methods in diagnostics of head injury. part 1. computer tomography and magnetic resonance imaging. *Russian journal of neurosurgery*, (2):7–16, 2014.
- [10] H. White. A heteroscedasticity-consistent covariance matrix estimator and a direct test for heteroscedasticity.