

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ

ГОСУДАРСТВЕННОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ  
ВЫСШЕГО ПРОФЕССИОНАЛЬНОГО ОБРАЗОВАНИЯ  
"МОСКОВСКИЙ ФИЗИКО-ТЕХНИЧЕСКИЙ ИНСТИТУТ  
(ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ)"

ФАКУЛЬТЕТ ИННОВАЦИЙ И ВЫСОКИХ ТЕХНОЛОГИЙ  
КАФЕДРА АНАЛИЗА ДАННЫХ

---

Выпускная квалификационная работа по направлению  
01.03.02 «Прикладная математика и информатика»  
НА ТЕМУ:

**НЕЙРОСЕТЕВЫЕ МОДЕЛИ В ЗАДАЧАХ  
ИНТЕРПРЕТАЦИИ ДАННЫХ ФМРТ ПРИ  
АУДИАЛЬНОЙ СТИМУЛЯЦИИ**

Студентка

Медведева А.Е.

Научный руководитель к.ф-м.н

Артемов А.В.

Зам. зав. кафедрой д.ф-м.н, проф.

Бунина Е.И.

МОСКВА, 2017

# Оглавление

<b>1 Введение</b>	<b>2</b>
1.1 История исследования мозга . . . . .	2
1.2 ФМРТ в измерении активности мозга . . . . .	3
1.3 Функции зон мозга . . . . .	5
<b>2 Обзор литературы</b>	<b>9</b>
2.1 Семантический атлас мозга . . . . .	9
2.2 Применение нейросетевых методов . . . . .	11
2.3 Мотивация исследования . . . . .	13
<b>3 Цель работы</b>	<b>14</b>
3.1 Постановка задачи . . . . .	14
3.2 Цель и задачи исследования . . . . .	15
<b>4 Методы исследования</b>	<b>16</b>
<b>5 Эксперимент</b>	<b>19</b>
5.1 Описание данных . . . . .	19
5.2 Препроцессинг данных . . . . .	19
5.3 Линейные методы . . . . .	20
5.3.1 Анализ весов гребневой регрессии . . . . .	23
5.4 Нейросетевые методы . . . . .	26
5.5 Избирательность в нейросетевых подходах . . . . .	30
<b>6 Сравнение моделей</b>	<b>31</b>
<b>7 Заключение</b>	<b>32</b>

# Часть 1

## Введение

### 1.1 История исследования мозга

Проблема исследования мозга человека, соотношения мозга и психики - одна из самых захватывающих задач, которые когда-либо возникали в науке. Ведь все, что до сих пор исследовалось - и атом, и галактика, и мозг животного - было проще, чем мозг человека. С философской точки зрения неизвестно, возможно ли в принципе решение этой задачи. Ведь, кроме приборов и методов, главным средством познания мозга остается опять-таки наш человеческий мозг. Обычно прибор, который изучает какое-то явление или объект, сложнее этого объекта, в этом же случае мы пытаемся действовать на равных - мозг против мозга.

В прошлом веке были обнаружены зоны мозга, отвечающие за речь, - по имени открывателей их называют области Брока и Вернике (см. 1.5). Были сделаны великие открытия, но возможности методик того времени для изучения человеческих функций весьма ограничены: психологические тесты, клинические наблюдения и начиная с тридцатых годов электроэнцефалограмма.

Прорыв в понимании правого полушария произошел в 1960-х годах, когда Роджер Сперри обнаружил, что правое полушарие мозга просто обрабатывало информацию по-разному: где левая сторона обрабатывала информацию последовательно, правая часть обрабатывала ее более сложноустроенным способом; где левая сторона распознает имя человека, правая сторона распознает их лицо; где левая сторона хороша в логике и анализе, правая сторона имеет логичный смысл, распознает узоры и интерпретирует эмоции и невербальные выражения.

Также в 1960-х годах невролог Пол Маклин изучал эволюцию человеческого мозга. Он пришел к выводу, что в нашем черепе не один мозг, а три, каждый из которых представляет собой отдельный эволюционный слой, который сформировался на более старом слое перед ним.

В последнее время технология сканирования ФМРТ позволила нейробиологам понять, что, хотя в мозге без сомнений есть области, специализированные для обработки определенных типов информации, несколько частей мозга вовлечены практически во все, что мы делаем. Этот прорыв позволил нам понять, что наши мысли и характер определяются меньше самими мозговыми областями, но больше относительной силой связи нейронных сетей, которые их соединяют.

## 1.2 ФМРТ в измерении активности мозга

Функциональная магнитно-резонансная томография, или ФМРТ, - это метод измерения активности головного мозга. Он работает путем обнаружения изменений насыщения крови кислородом, а также изменение кровотока, возникающего в ответ на нейронную активность. Когда зона мозга более активна, она потребляет больше кислорода, и чтобы удовлетворить этот растущий спрос, увеличивается приток крови к рабочей области. ФМРТ может быть использована для создания карт активности, показывающих, какие части мозга вовлечены в конкретный психический процесс.

Каждое ФМРТ измерение – последовательность трехмерных сканов мозга, где каждый скан есть набор значений сигнала на каждом маленьком участке мозга, называемым вокселием и представляющим собой параллелепипед с сторонами порядка нескольких миллиметров (см. 1.1).

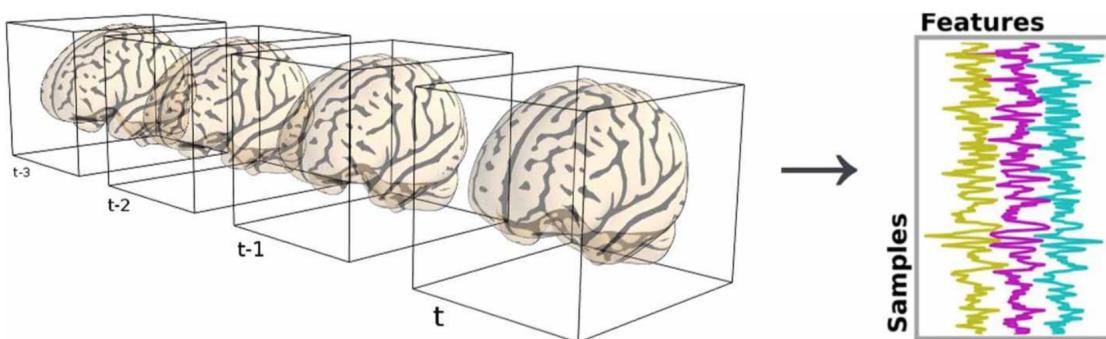


Рис. 1.1: ФМРТ измерение.

ФМРТ имеет несколько преимуществ перед другими методами измерения мозговой активности:

1. Этот метод является неинвазивным и не использует рентгеновское излучение, что делает его безопасным для человека.
2. Имеет хорошую пространственно-временное разрешение. На рисунке 1.2 показано сравнение ФМРТ с другими методами измерения активности мозга, такими как энцефалография(МЭГ, ЭЭГ) и позитронно-эмиссионная томография (ПЭТ).

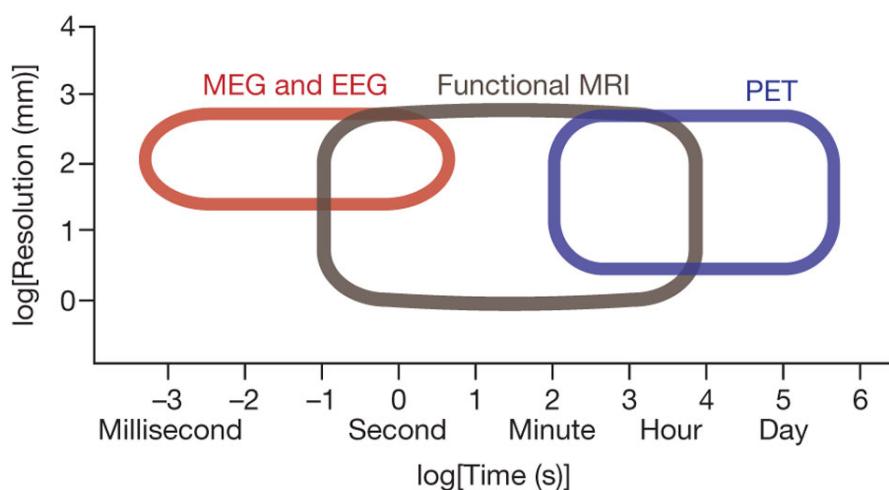


Рис. 1.2: Сравнение ФМРТ, ПЭТ и МЭГ.

3. Его легко использовать для постановки экспериментов и далее анализировать (см. [1]).

Из недостатков можно выделить следующие:

1. Для точного измерения необходимо, чтобы человек оставался неподвижным. Движения головы, вызванные биением сердца и дыханием, являются одним из основных источников шума.
2. Наличие шума, вызванного неоднородностью магнитного поля, используемого для измерения.
3. ФМРТ измеряет уровень активности не напрямую, а косвенно.
4. Уровень насыщения крови кислородом, который измеряет ФМРТ, изменяется очень медленно. Его изменение происходит с некоторой задержкой (обычно

порядка пары секунд), так как необходимо некоторое время для реакции сосудистой системы в ответ на потребность мозга в глюкозе. После чего уровень сигнала обычно поднимается до пика примерно через 5 секунд после стимула. Если нейроны остаются активными, пик распространяется на плоское плато. После остановки деятельности, сигнал падает ниже исходного уровня – явление, называемое недолетом. С течением времени сигнал восстанавливается до базовой линии (см. 1.3).

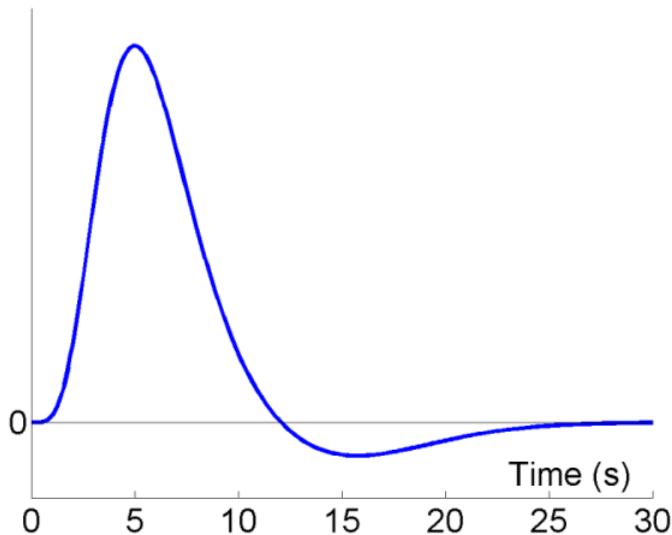


Рис. 1.3: Уровень насыщения крови в зависимости от времени.

### 1.3 Функции зон мозга

В головном мозге человека ученые выделяют три основные части: задний мозг, средний мозг и передний мозг. Исторически более древними считаются задний и средний мозг. Они отвечают за жизненно важные внутренние функции организма: поддержание тока крови, дыхание. За человеческие формы коммуникации с внешним миром (мышление, память, речь) отвечает передний мозг, он и будет рассмотрен далее.

В коре головного мозга можно выделить четыре доли: затылочную, теменную, височную и лобную. В соответствии с принципом латерализации – разделения по полушариям – каждая доля имеет свою пару (см. 1.4).

Лобные доли можно условно назвать командным пунктом головного мозга. Здесь находятся центры, не столько отвечающие за отдельное действие, сколько обес-



Рис. 1.4: Зоны мозга человека.

печивающие такие качества, как самостоятельность и инициативность человека, его способность к критической самооценке. Поражение лобных долей вызывает появление беззаботности, бессмысленных устремлений, переменчивости и склонности к неуместным шуткам. С утратой мотивации при атрофии лобных долей человек становится пассивным, теряет интерес к происходящему, часами остается в постели. Нередко окружающие принимают такое поведение за леность, не подозревая, что изменения в поведении есть прямое следствие гибели нервных клеток этой зоны коры головного мозга.

Функции, выполняемые теменными долями, отличаются для доминирующей и недоминирующей сторон. Доминирующая сторона (обычно левая) отвечает за способность понимать устройство целого через соотнесение его частей (их порядок, структуру) и за наше умение складывать части в целое. Это относится к самым разным вещам. Например, для чтения необходимо уметь складывать буквы в слова и слова во фразы. То же с цифрами и числами. Недоминантная сторона (обычно правая) – это центр, который, комбинируя информацию, поступающую из затылочных долей, обеспечивает трехмерное восприятие окружающего мира. Нарушение этой области коры приводит к зрительной агнозии – неспособности распознавать предметы, лица, окружающий пейзаж.

Затылочные доли отвечают за переработку зрительной информации. По сути, все, что мы видим, мы видим не глазами, которые лишь фиксируют раздражение воздействующего на них света и переводят его в электрические импульсы. Мы «видим» затылочными долями, которые интерпретируют поступающие от глаз сигналы.

Зная об этом, необходимо отличать у пожилого человека ослабление остроты зрения от проблем, связанных с его способностью воспринимать предметы. Острота зрения (способность видеть мелкие объекты) зависит от работы глаз, восприятие – продукт работы затылочной и теменной долей мозга. Информация о цвете, форме, движении обрабатывается отдельно в затылочной доле коры, прежде чем будет принята в теменной доле для превращения в трехмерное представление.

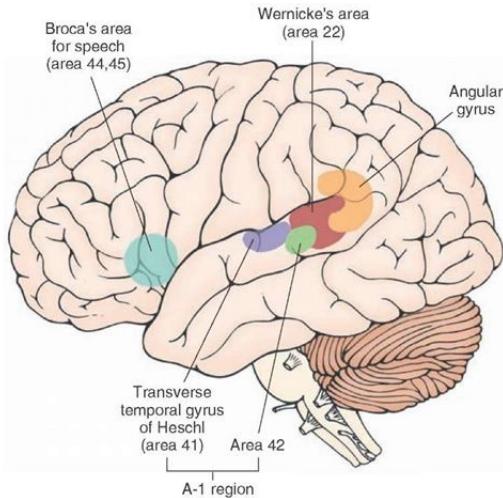


Рис. 1.5: Области Брока, Вернике и Хеши, важные для обработки устной речи.

Височные доли в верхних отделах обрабатывают слуховые ощущения, превращая их в звуковые образы. Поскольку слух – это канал, по которому человеку передаются звуки речи, височные доли (особенно доминантная левая) играют важнейшую роль в обеспечении речевой коммуникации. Именно в этой части мозга производится распознавание и наполнение смыслом обращенных к человеку слов, а также подбор единиц языка для выражения собственных смыслов. Недоминантная доля (правая у правшей) участвует в распознавании интонационного рисунка и выражения лица. Небольшой участок на внутренней поверхности височных долей, имеющий форму морского конька (гиппокамп), контролирует долговременную память человека. Именно височные доли хранят наши воспоминания. Доминантная (обычно левая) височная доля имеет дело с вербальной памятью и названиями объектов, недоминантная используется для зрительной памяти.

Также в височной зоне недавно был найден участок (см.[6]), называемый извилином Hoschi, общий у людей и обезьян, в котором происходят процессы, участвующие в обработке речи и предсказании слова, которое вот-вот прозвучит (см. 1.5). По сло-

вам исследователей, фактически обнаруженные механизмы речи в мозге человека работают как предиктивный ввод текста в мобильном телефоне, предвосхищая то, что вы сейчас услышите.

## Часть 2

### Обзор литературы

#### 2.1 Семантический атлас мозга

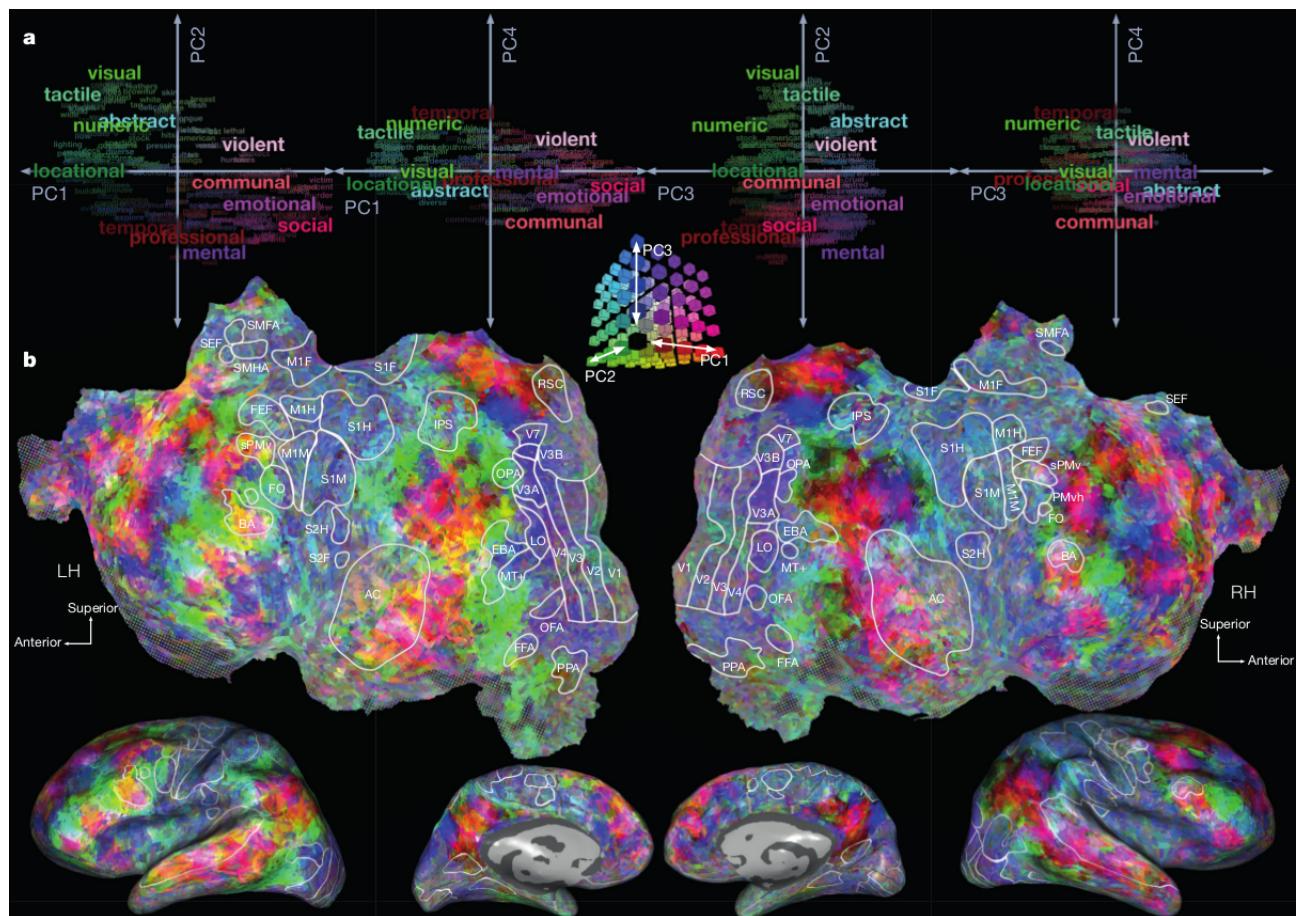


Рис. 2.1: Результаты из работы [5].

Работа [5] группы американских нейробиологов посвящена созданию семантического атласа мозга на основе данных, полученных с помощью ФМРТ. Ученым

удалось выделить области на фронтальной коре, ответственные за конкретные семантические группы, такие как социальные явления, абстрактные понятия, количественные слова.

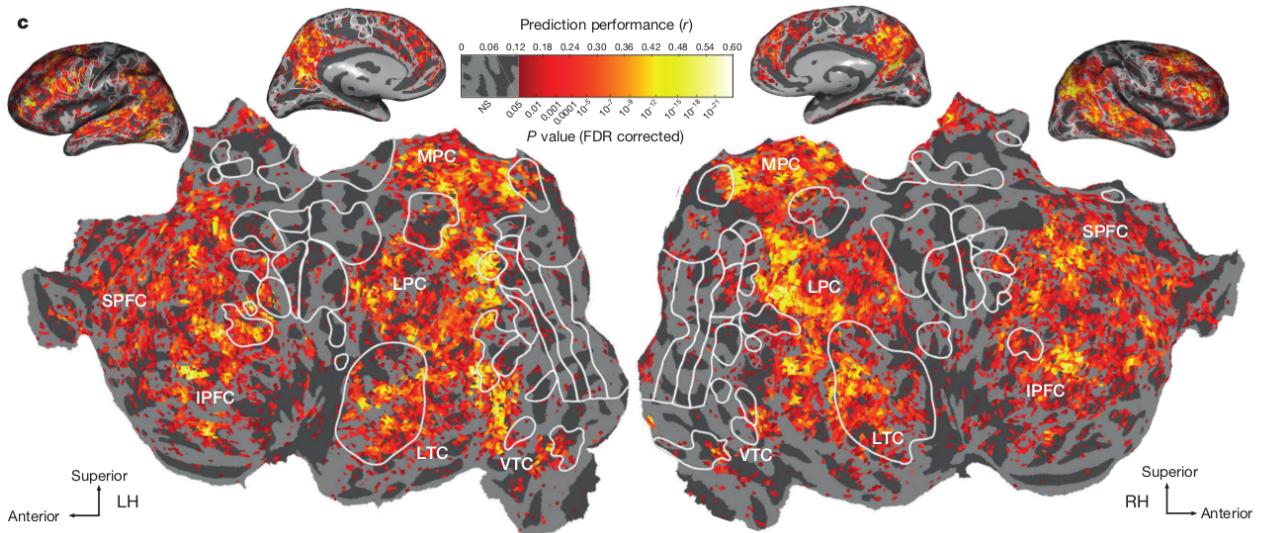


Рис. 2.2: Результаты из работы [5].

Эксперимент был построен следующим образом: 8 испытуемых слушали 2 часа рассказы из радиопередачи. После чего авторы подсчитали на большом корпусе матрицу совместной встречаемости каждого слова из текста аудио и 985 самых часто употребляемых английских слов. Каждая строка этой матрицы соответствовала представлению слова из аудио в 985-мерном пространстве. Для установления зависимости между признаками аудиостимула (представления слов из аудио) и измерениями ФМРТ авторы использовали гребневую регрессию. Далее изучались вектора весов регрессии, соответствующей каждому воксели, для чего они были спроектированы в четырехмерное пространство с помощью метода главных компонент. В это же пространство были спроектированы и слова из аудиостимула. С помощью алгоритма k средних спроектированные слова были сгруппированы в 12 категорий. После чего авторы раскрасили пространство из первых трех главных векторов и визуализировали результат (см. 2.1).

Интересна проекция на первую компоненту: положительные знак имеют слова, относящиеся к человеку и человеческим взаимодействиям, отрицательную – описания восприятия, местоположения, количества. На рис. 2.2 показана средняя корреляция Пирсона в проекции на поверхность долей. Видно, что зоны LTC, VTC, LPC, MPC, SPFC и IPFC предсказываются достаточно хорошо (корреляция значимо боль-

ше нуля), ранее эти зоны считались семантической системой мозга.

Также интересно, что височные зоны, отвечающие за анализ слуховой информации, более всего реагируют на слова из социальной группы (см. 2.3).

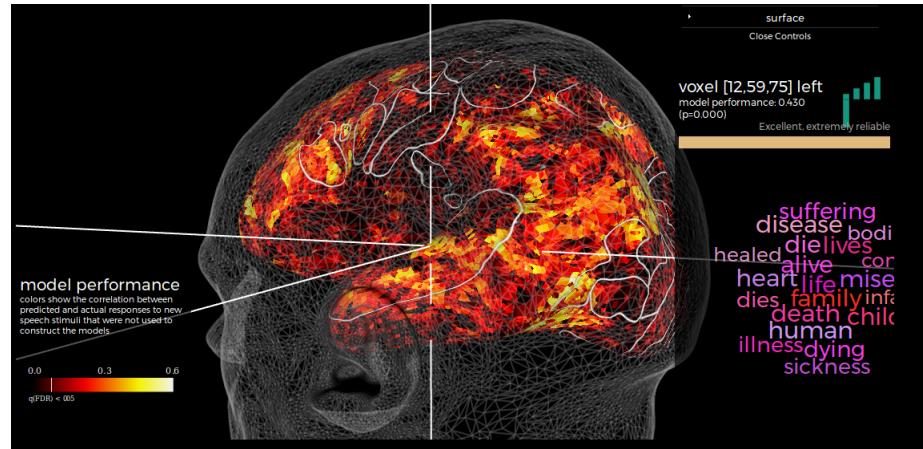


Рис. 2.3: Результаты из [5].

## 2.2 Применение нейросетевых методов

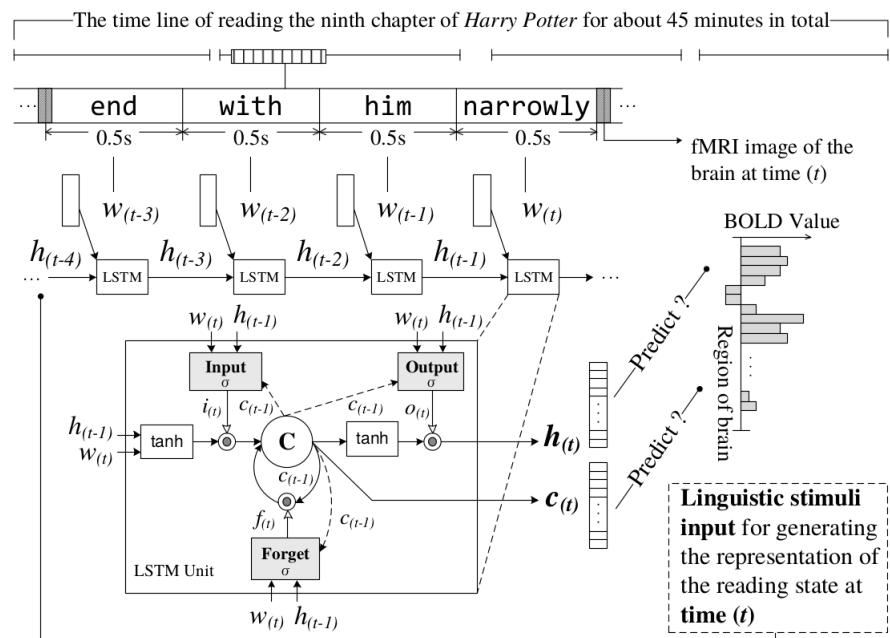


Рис. 2.4: Схема эксперимента из [7].

В статье [7] изучалась предсказательная способность LSTM юнитов в зависимости от архитектуры (наличия гейтов). Эксперимент был устроен следующим

образом: в течение 45 человек равномерно (со скоростью появление слов - 1 слово каждый полсекунды) читали главу из книги "Гарри Поттер".

Далее авторы обучили на оставшихся главах нейронную сеть, состоящую из одного скрытого слоя размерности 100, предсказывать следующее слово в тексте. Затем с помощью метода наименьших квадратов устанавливалась зависимость между вектором внутреннего состояния нейросети  $a_t$  в момент предсказывания следующего слова и вектором активации всего мозга перед прочтением этого слова (см 2.4):

$$y_t = Ma_t, a_t \in \{c_t, h_t\}.$$

В качестве вектора внутреннего состояния сети  $a_t$  может служить как вектор памяти lstm юнитов  $c_t$ , так в "внешнее" внутреннее состояние  $h_t$  - те значения, которые рекуррентный слой передает полно связному. Мерой качества модели авторы использовали косинусное расстояние между предсказанием и действительной активацией мозга, нормализованным на отрезок [0,1]. Несмотря на то, что статья была посвящена изучению LSTM юнитов, авторы получили неплохую предсказательную способность (см. 2.5), максимальное качество 0.86 достигалось при максимально возможном окне просмотра на тесте и использовании вектора внутренней памяти сети.

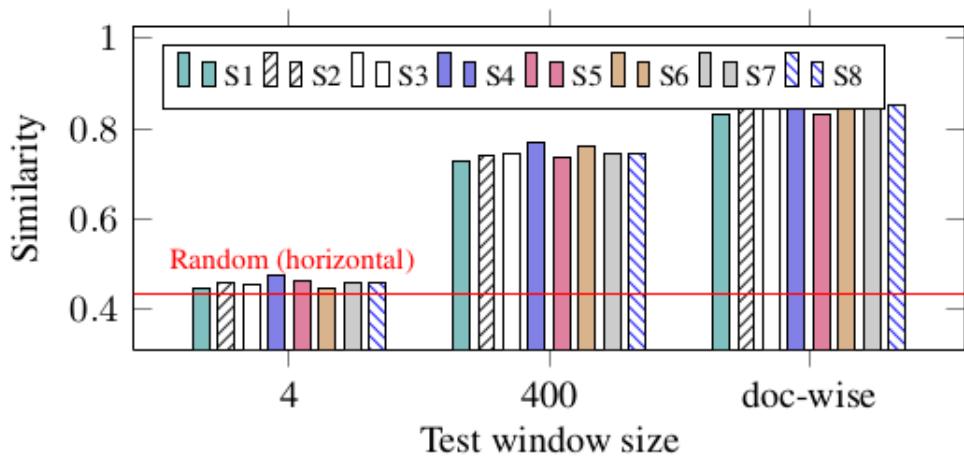


Рис. 2.5: Результаты из работы [7].

## 2.3 Мотивация исследования

В статье [4] отсутствуют результаты для альтернативных семантических представлений, например word 2 vec. Также интерес представляет возможность ли воспроизвести результаты для другого языка и других параметрах эксперимента. В статье [7] целью являлось исследование LSTM сетей, а не точность предсказания мозгового аудиостимула, также стимул в эксперименте был визуальный и слова появлялись на экране равномерно.

# Часть 3

## Цель работы

### 3.1 Постановка задачи

В типичном эксперименте с помощью ФМРТ по внешнему стимулу строится матрица признаков  $\mathbf{X}$ . Например, при аудиальной стимуляции строки матрицы  $\mathbf{X}$  – векторы, соответствующие словам из текста аудио. Измерения ФМРТ – матрица ответов  $\mathbf{Y}$ , столбцы которой соответствуют вокселям, строки – вектор активации мозга в конкретный момент времени. Далее задача сводится к построению модели  $f$ , максимизирующую корреляцию Пирсона на тестовой выборке:

$$Pearson(\mathbf{Y}_{test}, f(\mathbf{X}_{test})) \rightarrow max_f.$$

В случае, если для каждого вокселя своя модель из класса  $\mathcal{F}$ , то максимизируется средняя корреляция по вокселям. Обозначим число вокселей за  $N$ ,  $i$ -ый столбец матрицы  $\mathbf{Y}$  за  $\mathbf{Y}^{:,i}$ . Тогда функционал для максимизации выглядит следующим образом:

$$\frac{1}{N} \sum_{i=1}^N Pearson(\mathbf{Y}_{test}^{:,i}, f_i(\mathbf{X}_{test})) \rightarrow max_{\mathcal{F}}.$$

Нетрудно убедится в том, что при нормализованном отклике данная задача максимизации эквивалентна минимизации средней квадратичной ошибки между предсказанием и действительным откликом. Обозначим  $\hat{Y} = f(\mathbf{X}_{test})$ ,  $Y = \hat{Y} + \epsilon$ , где  $\hat{Y}$  независим от  $\epsilon$ . Учитывая, что  $\sigma_Y^2 = \sigma_{\hat{Y}}^2 + \sigma_{\epsilon}^2$ , получаем:

$$\begin{aligned}
Pearson(Y, \hat{Y})^2 &= \frac{cov^2(Y, \hat{Y})}{\sigma_Y^2 \sigma_{\hat{Y}}^2} = \frac{cov(\hat{Y} + \epsilon, \hat{Y}) cov(\hat{Y} + \epsilon, \hat{Y})}{\sigma_Y^2 \sigma_{\hat{Y}}^2} = \frac{\sigma_{\hat{Y}}^4}{\sigma_Y^2 \sigma_{\hat{Y}}^2} = \\
&= \frac{\sigma_{\hat{Y}}^2}{\sigma_Y^2} = \frac{\sigma_Y^2 - \sigma_{\epsilon}^2}{\sigma_Y^2} = 1 - \frac{MSE(Y, \hat{Y})}{\sigma_Y^2}.
\end{aligned}$$

## 3.2 Цель и задачи исследования

Целью исследование является разработка и исследование математической модели, алгоритмов и комплексов программ для интерпретации данных ФМРТ при аудиальной стимуляции. Более конкретно, задачи исследования являются:

1. Независимое воспроизведение линейной повоксельной модели ФМРТ-отклика на аудиальный стимул, в том числе:
  - Создание экспериментальной среды, допускающей легкую воспроизведимость исследования.
  - Попытка использования предобученного word 2 vec в предсказании ФМРТ отклика.
  - Сравнение семантических групп с результатами, полученными для английского языка.
2. Разработка нелинейной нейросетевой повоксельной модели ФМРТ-отклика на аудиальный стимул, в том числе:
  - Попытка применения рекуррентных сетей для предсказания ФМРТ отклика.
  - Использование семантических предсказаний на основе рекуррентных сетей.
  - Сравнение линейных и нелинейных моделей.

## Часть 4

# Методы исследования

Пусть  $\mathbf{X}$  – матрица признаков,  $\vec{y}$  – вектор ответов. Опишем модели, используемые в работе.

1. Гребневая регрессия. Предполагаем, что  $\vec{y} = \mathbf{X}\vec{w} + \vec{\epsilon}$ , где  $\vec{\epsilon} \sim N(0, \sigma^2 I)$ ,  $\vec{w} \sim N(0, \frac{\sigma^2}{\alpha} I)$ . Оптимизационная задача тогда выглядит следующим образом:

$$\vec{w}^* = \operatorname{argmin}_{\vec{w}} (\|\vec{y} - \mathbf{X}\vec{w}\|^2 + \alpha \|\vec{w}\|^2),$$

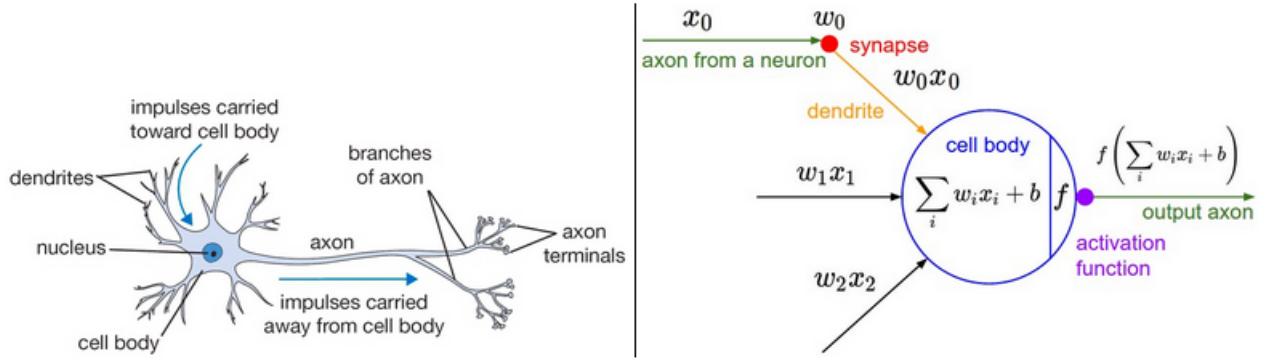
ее решение:

$$\hat{\vec{w}} = (\mathbf{X}^T \mathbf{X} + \alpha I)^{-1} \mathbf{X}^T \vec{y}.$$

2. Полносвязные нейронные сети. Главной частью нейронных сетей является искусственный нейрон ([8]). Он был создан по подобию человеческого нейрона (см 4.1), который принимает сигналы по дендритам, обрабатывает их и передает сигнал по аксону другим нейронам. Архитектура искусственного нейрона весьма проста: принимая на входе вектор  $\vec{x}$ , он возвращает  $f(\vec{x}) = \sigma(\sum_{i=1}^n w_i x_i)$ , где  $\sigma$  – некоторая нелинейная функция активации, например, сигмоида:  $\sigma(x) = \frac{1}{1+e^{-x}}$ .

3. Рекуррентные нейронные сети.

- LSTM-модуль это рекуррентный модуль сети ([3]), способный запоминать значения как на короткие, так и на длинные промежутки времени. Ключом к данной возможности является то, что LSTM-модуль не использует функцию активации внутри своих рекуррентных компонентов. Таким образом,



A cartoon drawing of a biological neuron (left) and its mathematical model (right).

Рис. 4.1: Архитектура искусственного нейрона.

хранимое значение не размывается во времени, и градиент или штраф не исчезает при использовании метода обратного распространения ошибки во времени (Backpropagation through time) при тренировке сети (см. 4.2).

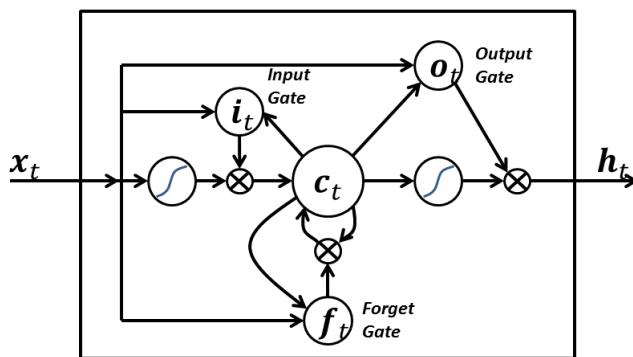


Рис. 4.2: Архитектура LSTM юнита.

- GRU - механизм в рекуррентных нейронных сетях, также использующий ворота (gates), чтобы управлять потоком информации внутри себя, однако имеет меньше параметров, чем LSTM, ввиду того что не имеет отдельную клетку памяти (memory cell), (см.[2]).

4. Метод главных компонент - один из основных способов уменьшить размерность данных, потеряв наименьшее количество информации. Пусть матрица  $X$  имеет размерность  $n, k$ . Для этого решаем задачу нахождения таких чисел  $Z = (z_{ij}), i = 1, \dots, n, j = 1, \dots, m$  и  $U = (u_{ij}), i = 1 \dots k, j = 1 \dots m, m < n$  чтобы:

$$\hat{X}_i = \sum_{l=1}^m z_l u_{jl}$$

был как можно ближе к  $X_i$ . Получаем задачу оптимизации:

$$\sum_{i=1}^k \|X_j - \hat{X}_j\|^2 = \|ZU^T - X\|^2 \longrightarrow \min_{Z,U}$$

Ее решение: матрица  $U$ , такая что ее столбцы являются собственными векторами матрицы  $X^T X$ , соответствующие первым  $m$  наибольшим собственным значениям, а  $Z = XU$ .

# Часть 5

## Эксперимент

### 5.1 Описание данных

Данные были взяты с открытой базы данных ФМРТ openfmri.org (см. [4]). В ходе эксперимента 20 носителей немецкого языка в возрасте от 21 до 38 лет 2 часа (8 сегментов по 15 минут) слушали немецкий аудиофильм "Форрест Гамп версия для слабовидящих. Данная версия отличается от обычной тем, что в моменты, когда нет диалогов, голос за кадром описывает, что происходит на экране. При этом ФМРТ сканер измерял каждые две секунды мозговую активность с пространственным разрешением  $2.75\ mm^3$ .

### 5.2 Препроцессинг данных

К данным ФМРТ была применена коррекция движений а также подавление шума с помощью библиотеки FLIRT, затем ФМРТ сигнал был посегментно нормализан, чтобы иметь среднее ноль и дисперсию 1.

Для дальнейшего работы необходимо знать точное соответствие каждого слова из транскрипции аудио и момента его произношения (выравнивание транскрипции). Так как транскрипция отсутствовала, она создавалась самостоятельно. Сначала на основе описания сцен и субтитров создавалась фразовая транскрипция аудио. Далее каждая фраза и соответствующий фрагмент аудио подавалась на вход библиотеке aeneas, которая выдавала для каждого слова из фразы соответствующий ему момент в аудио. Также был испробован инструмент Google Api Speech Recognition

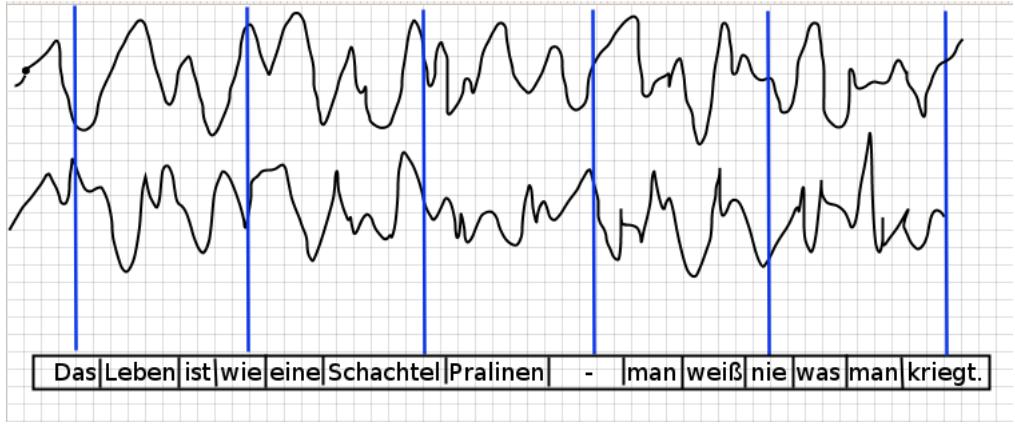


Рис. 5.1: Визуализация эксперимента.

для того чтобы создать более точную версию субтитров, однако данный подход оказался несколько хуже ввиду ошибок распознавания.

Далее с помощью предобученной модели word to vec слова из транскрипции были переведены в трехсотмерные вектора из  $\mathbb{R}^{300}$ . Затем полученные представления слов интерполировались с ядром Ланшица, чтобы согласовать моменты произношения слов и моменты измерения ФМРТ (см. 5.1). Ввиду того что отклик конкретного вокселя происходит с некоторой задержкой, а также ввиду сложной формы сигнала отклика (1.3), для каждого момента измерения стимул, соответствующий этому моменту, был сконкатенирован с тремя предыдущими (см. 5.2). Таким образом получилось 1200 семантических признаков. Далее из 3600 пар стимул - измерение первые 3100 использовались для обучения и последние 500 для валидации.

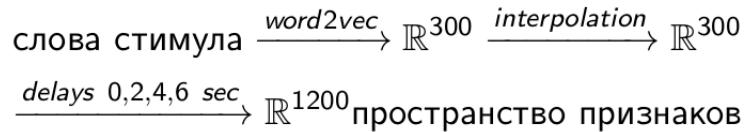


Рис. 5.2: Схема построения пространства признаков.

### 5.3 Линейные методы

Повторяя эксперимент авторов статьи [5] а также избегая переобучения ввиду большого числа признаков и относительно малого числа данных, мы использовали повоксельно гребнеую регрессию для предсказания ФМРТ отклика.

На рисунке 5.3 показаны результаты корреляции на тесте в зависимости от

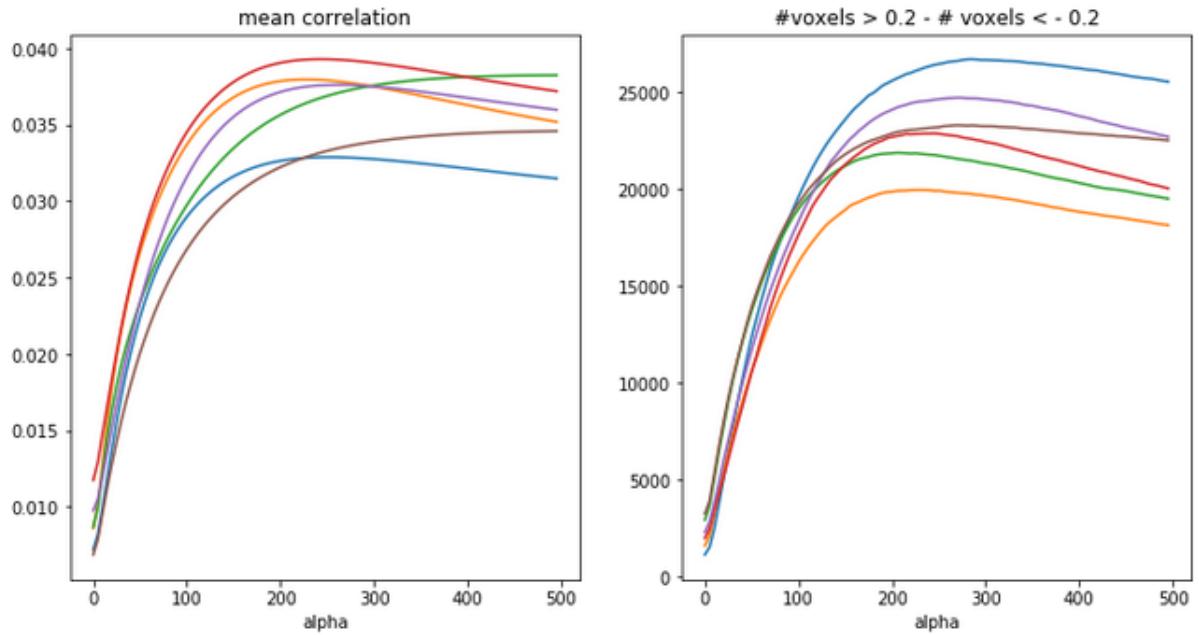


Рис. 5.3: Корреляция на тесте в зависимости от параметра  $\alpha$

параметра регуляризации альфа гребневой регрессии для 6 человек. На первом графике показана средняя повоксельная корреляция, на втором – разница числа вокселей, корреляция которых больше 0.2 и числа вокселей, корреляция которых меньше -0.2. По графику видно, что оптимальные значения параметра альфа примерно одного порядка и средняя корреляция стабильно больше нуля. Также оптимальное значение параметра альфа в статье [5] равняется 185, что примерно соответствует графикам. Далее для каждого субъекта выбиралась свое оптимальное значение параметра альфа, при этом общее для всех вокселей данного человека.

На рисунке 5.4 показана гистограмма корреляций вокселей для третьего испытуемого. Видно, что положительных значений больше и среднее значимо отличается от нуля. Критерий согласия Колмогорова для проверки на нормальность показывает *p-value* равным 0, что означает что данное распределение нельзя считать нормальным. Также критерий Вилкенсона для проверки на симметричность относительно нуля также дает *p-value* равным 0.

На рисунке 5.5 показан пример типичного сигнал ФМРТ на тесте и его предсказания, нормализованные для наглядности. В целом график выглядит оптимально, и убеждает в том, что обладает хорошей предсказательной способностью.

Посчитав для каждого вокселя результат корреляции на тесте и записав полученные результаты в трехмерный тензор, можно визуализировать карты результатов

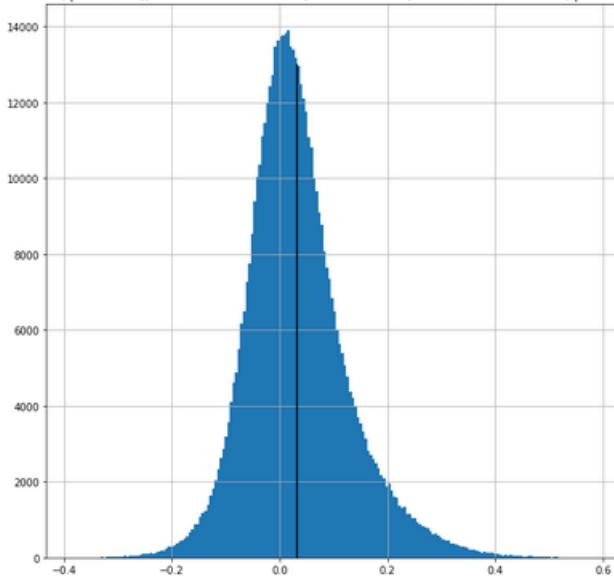


Рис. 5.4: Гистограмма корреляций при  $\alpha = 245$ .

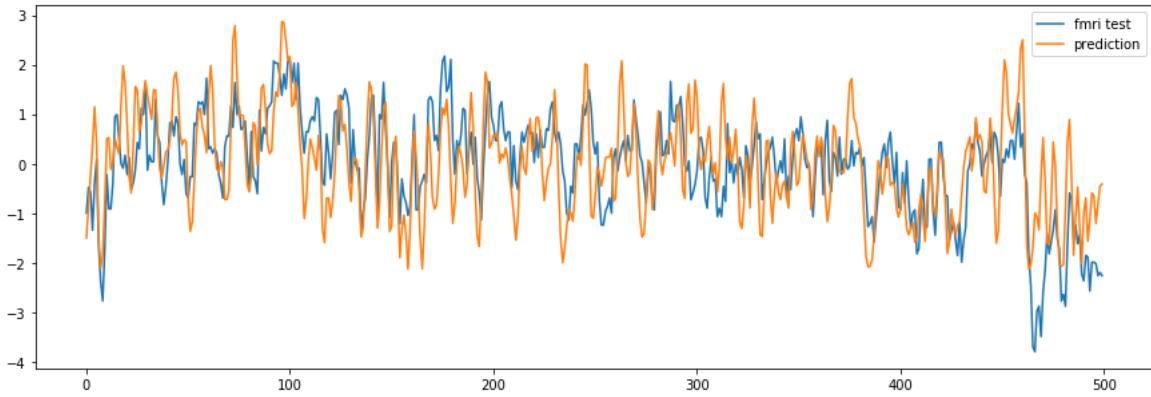


Рис. 5.5: Отклик вокселя и его предсказание.

модели. На рисунке 5.6 показаны последовательные срезы такой карты по оси z для одного и того же человека. Примечательно, что лучше всего для него предсказывается доминантная височная зона (для праворукого - левая), которая отвечает за наполнение смыслом обращенных к человеку слов.

На рисунке 5.7, показаны срезы на одном и том же уровне по оси z для шести различных субъектов, на 5.8 их 3d модели. В целом карты приблизительно повторяют друг друга: хорошо предсказываются височные доли (корреляция достигает 0.4 для всех субъектов), на рисунке 2.3 эти зоны больше всего реагируют на слова из социальной группы.

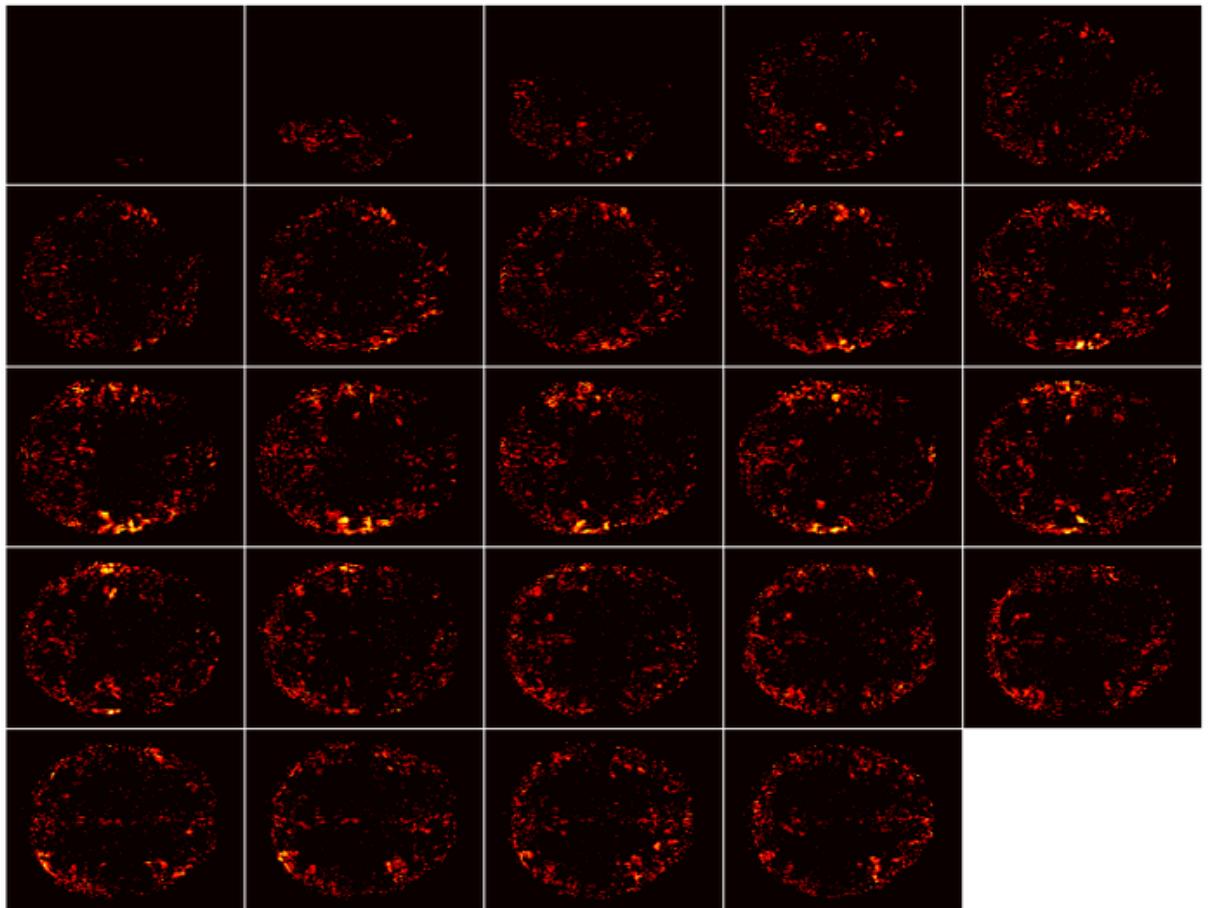


Рис. 5.6: Значимые результаты предсказания (корреляция  $> 0.1$ ) для 12-ого субъекта.

### 5.3.1 Анализ весов гребневой регрессии

После применения гребневой регрессии для каждого voxelя получилось 1200 весов  $\vec{v}_{old} \in \mathbb{R}^{1200}$ . Эти веса были обратно усреднены по представлениям таким образом, чтобы получить 300 мерные вектора:

$$\vec{w} = \frac{1}{4} [(v_1, \dots, v_{300})^T + (v_{301}, \dots, v_{600})^T + (v_{601}, \dots, v_{900})^T + (v_{901}, \dots, v_{1200})^T].$$

Далее для хорошо предсказываемых voxelей искались слова из аудиостимула, косинусное расстояние между представлением которых максимально с весами данного voxelя, что эквивалентно максимизации значения отклика в семантическом пространстве аудиостимула.

Интересно, например, что если брать два voxelя с одинаковыми координатами у двух разных человек и сравнивать топ из 20 слов, то у многих зон есть ненулевое пересечение. Например, для зоны на рисунке (5.9) следующее пересечение:

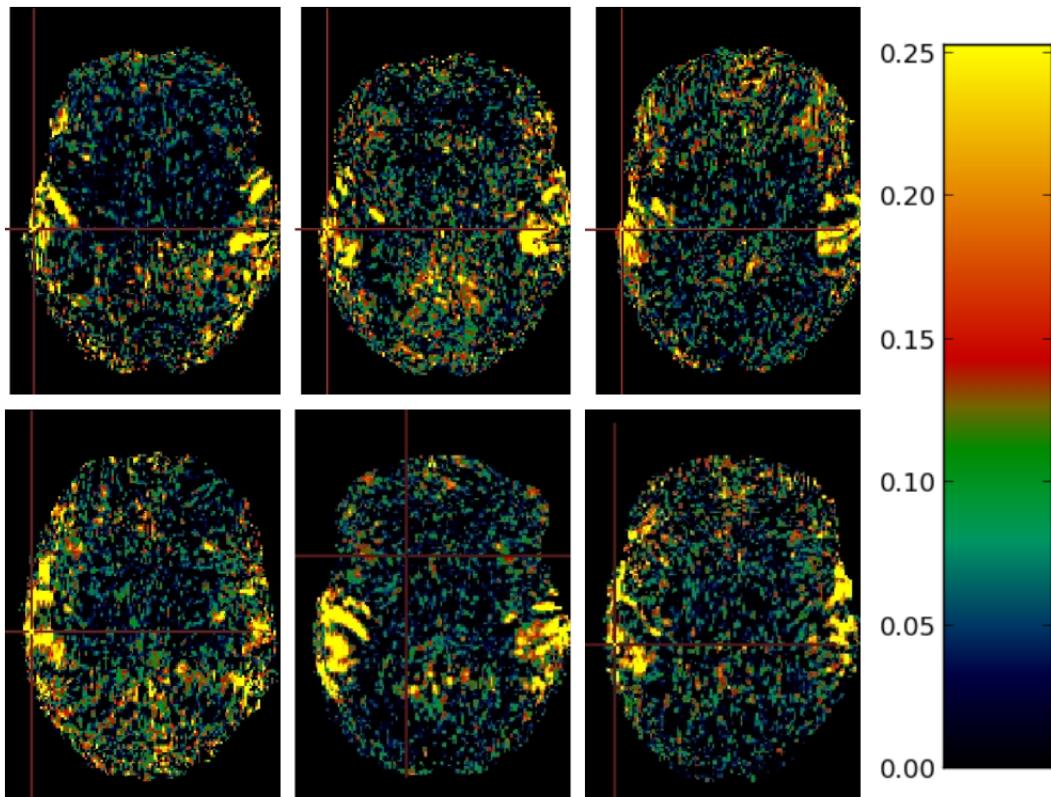


Рис. 5.7: Срезы карты корреляции ФМРТ отклика и предсказания.

- *wirklich* - собственно, в сущности, на самом деле
- *eigentlich* - действительно, в самом деле, право
- *ja* - да, конечно, обязательно
- *jawohl* - да, конечно, совершенно верно
- *aber* - однако, только, же
- *solltest* - должен, быть долженным, быть обязанным, следовать, надлежать
- *tust* - делать, творить, учинить
- *wirst* - становиться, стать, делаться, превращаться

Можно выделить слова, со смыслом "действительно а также глаголы. Продолжая анализировать поучившиеся веса, мы попробовали кластеризовать слова в топе наиболее хорошо предсказываемых вокселей. Были найдены воксели, для которых кластеризуется достаточно хорошо, например: см 5.10, 5.11. На рисунках в нулевом кластере достаточно простые слова: ball, name, miss, high, hot, strong, в первом

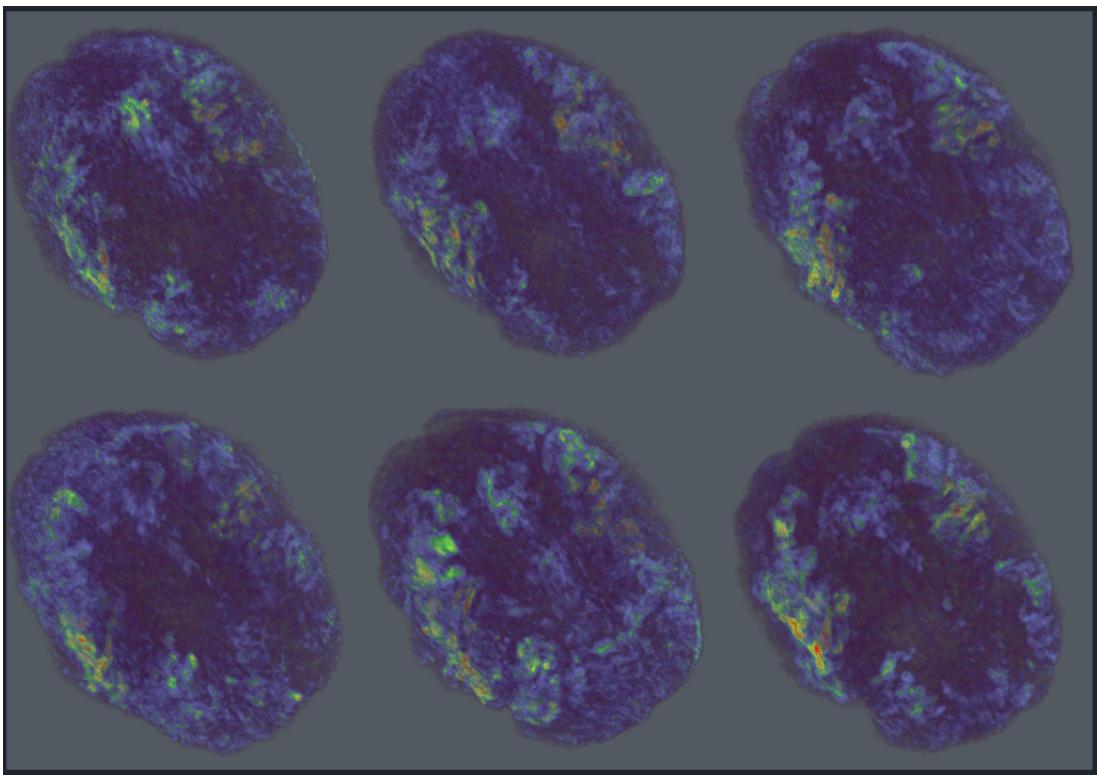


Рис. 5.8: Карта корреляции ФМРТ отклика и предсказания.

же много собственных имен: Elvis, Tommy, Danny, Roger. Также точками показана проекция на плоскость с помощью принципа главных компонент, две компоненты объясняют более трети дисперсии, что достаточно много для семантического трехсотмерного представления.

Мы создали демо, в котором можно выбрать человека и воксель, и будет показано на какие слова данный воксель имеет наибольший отклик (см. 5.12).

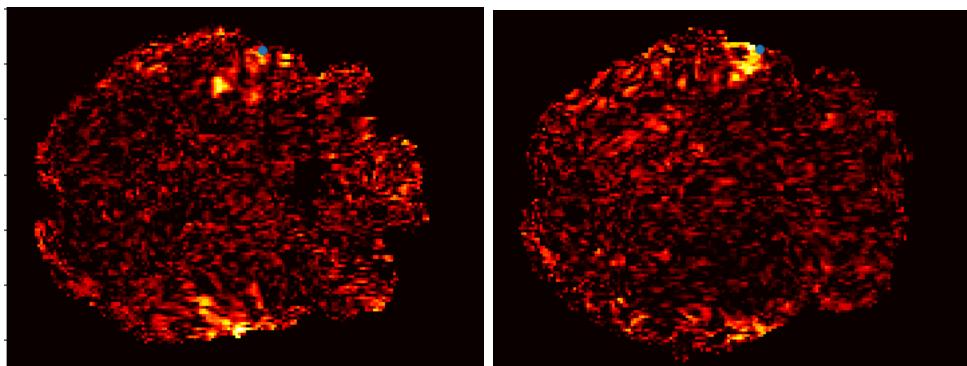


Рис. 5.9: Анализ вокселей для двух человек в одной и той же зоне.

```

Voxel index: (115, 84, 23), pearson correlation: 0.42884961477
Silhouette score: 0.254543602467, class : number_words per cluster: Counter({1: 81, 0: 19})
CLUSTER 0
truck      shit     blue    happens   baby    talent   ball    name    stock   high
apartment   bad      king    baby     little   ball    miss    stock   times   album
hot        strong
CLUSTER 1
john       gary      taylor   buford   lauf    roger
eric       charlie   henry   tex      hinterher wright
bert       bubba    gumbo   schnell billy   robert
dummer    maehler  farmer  alan    will    forrest
harris    sprecher fotos   runners wesley danny  sergeant
footballstar fotos   dann    dumm    schliesslich
wichser   thomas   shrimp  millionaer danach  verdammer schneller riesensprung elmar
haufen    shrimp   reinfallen millionaer
als       reinfallen stoppen bloeder  nachbar  verrueckter
band      jetzt    stoppen donner   trottel  langsamer
feuerball  kriegen  bernd   donner   johnson wieder
namens   wenn    verpruegeln bernd   heran    crockett
meinhardt  verpruegeln
sagten   mieser  coleman
Explained variance ratios per component: 1: 0.236832410098, 2: 0.0969822254737

```

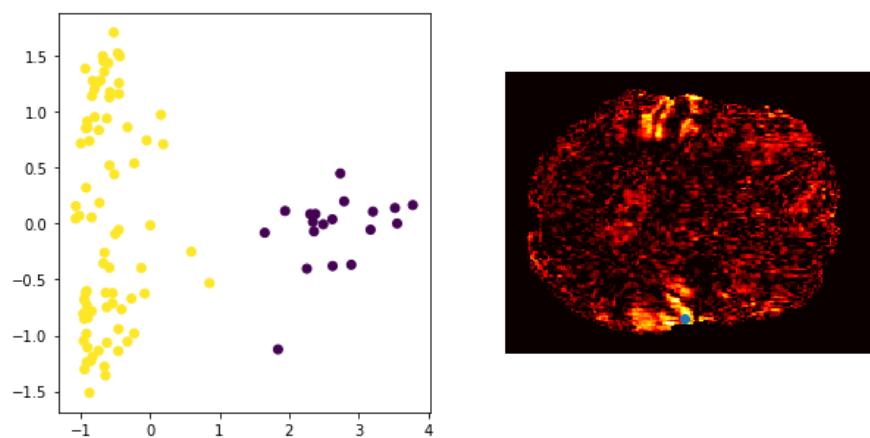


Рис. 5.10: Кластеризация топ 100 слов

```

Voxel index: (120, 70, 29), pearson correlation: 0.473467105688
Silhouette score: 0.269313693047, class : number_words per cluster: Counter({1: 71, 0: 29})
CLUSTER 0
truck      restaurant      happens      partner      name      hotel
blue       apartment      taxi        little       bins       walk
album      chance        moment      hope        coach      film
shit       american      stock       club        talent      baby
field     block        ball       rest
CLUSTER 1
OTOS      JOHN      HENRY      CHARLIE      ROGER      ERIC
elmar    TAYLOR    GARY      GUMBO      SHRIMP      ROBERT
bert     DANNY      RUNNERS   RIESENSprung THOMAS      EIN
lauf     FOOTBALLSTAR  BUBBA      THEATERBESITZER
billy    ALAN      BUFORD    FLACHMANN  DAVY        GEORGE
als      PHOENIX   DOSCH     NAMENS     EINZIGER   WRIGHT
maeher   WILL      TOMMY    VERDAMMLTER ZWEITER    GLASER
harris   COLEMAN   JOHNSON   SCHLIESSLICH ELVIS
bloeder  KYLE      HAARSCHNITT TEX        PINGPONG  VOGEL
sir      SUESSER   FORREST  HAARSCHNITT GESELLSCHAFT BENJAMIN
jawohl   DUMMIES   FUSSEL   VERDAMMTE  ERWIES    ROBIN
CHARLES  DUMMIES  ROCKY    ZIEL

```

Explained variance ratios per component: 1: 0.306257502234, 2: 0.076462258289

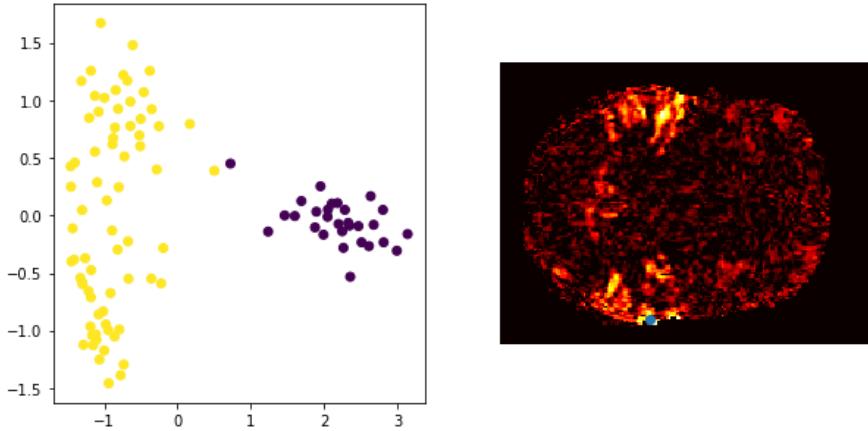


Рис. 5.11: Кластеризация топ 100 слов.

## 5.4 Нейросетевые методы

В первом приближении мы испробовали предсказывать отклик напрямую на основе синтерполированного представления аудиостимула. На графике 5.13 показана средняя корреляция Пирсона на тесте в зависимости от эпохи обучения для 1000 наиболее хорошо предсказываемых по результатам гребневой регрессии вокселей. Горизонтальной линией показан результат гребневой регрессии. Видно, что все архитектуры ведут себя не хуже регрессии, при этом лучший результат у рекуррентной сети (1 скрытый слой с 10 gru/lstm). Это можно объяснить тем, что рекуррентные сети лучше улавливают предыдущий контекст, нежели обычные полно связанные сети.

Так как данных не очень много для того чтобы обучить более мощную модель, мы испробовали несколько другой подход. Мы обучили нейронную сеть (см.

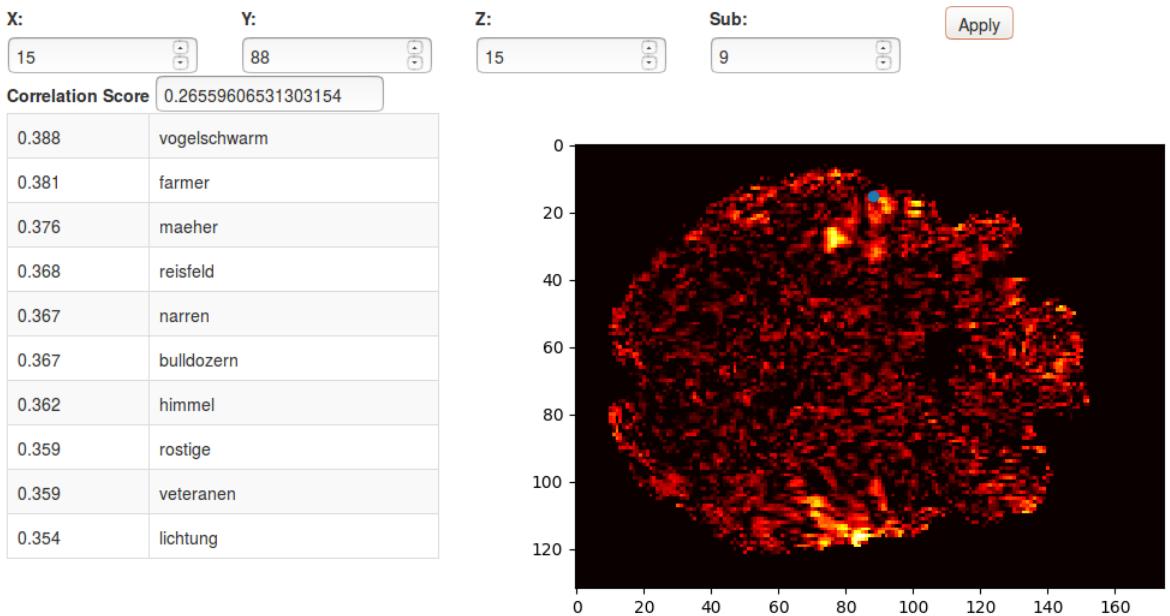


Рис. 5.12: Демо для изучения воксельной избирательности.

5.14) с одним скрытым LSTM слоем размерности 50 и трехсторонним представлением, проинициализированным предобученным word2vecом, использовавшимся в гребневой регрессии, предсказывать следующее слово на корпусе, состоящем из 13 немецких книг и 200000 статей немецкой Википедии. Статьи Википедии выбирались таким образом, чтобы хотя бы одна треть слов из статьи присутствовала в аудиостимуле. Генерация батчей для обучения происходила следующим образом: сначала с вероятностью, пропорциональной длине статьи или книги выбирался текст из обучающего корпуса, далее случайно выбиралось начало последовательности. Ответом для этой последовательности служила она же, сдвинутая на одно слова вперед.

Далее через эту сеть был пропущен аудиостимул и каждому слову сопоставлялось состояние lstm слоя в момент предсказания следующего слова. Далее новые полученные представления также интерполировались и конкатенировались и на основе их предсказывался отклик 1000 наиболее хорошо предсказываемых вокселей по результатам гребневой регрессии. Оказалось, что данное представление дает результат не хуже гребневой регрессии, при этом использует стомерное семантическое пространство, а не трехстороннее.

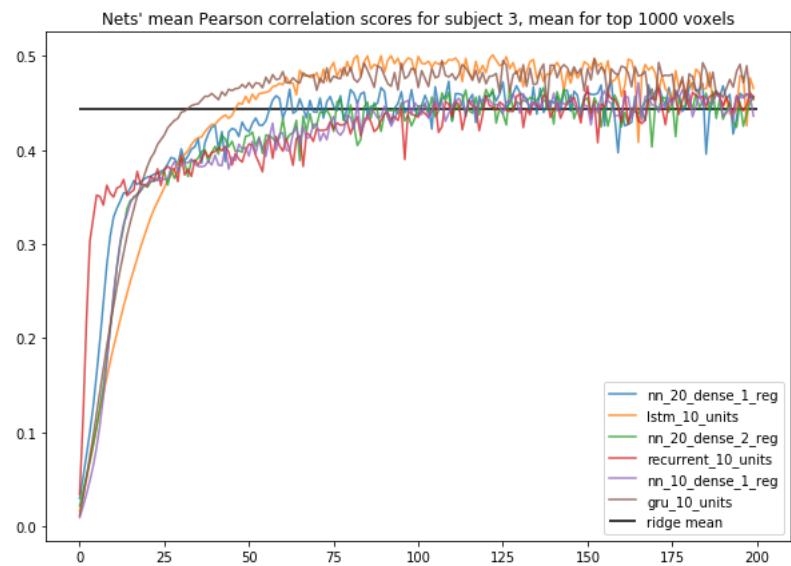


Рис. 5.13: Результат на тесте в зависимости от архитектуры сети и эпохи обучения.

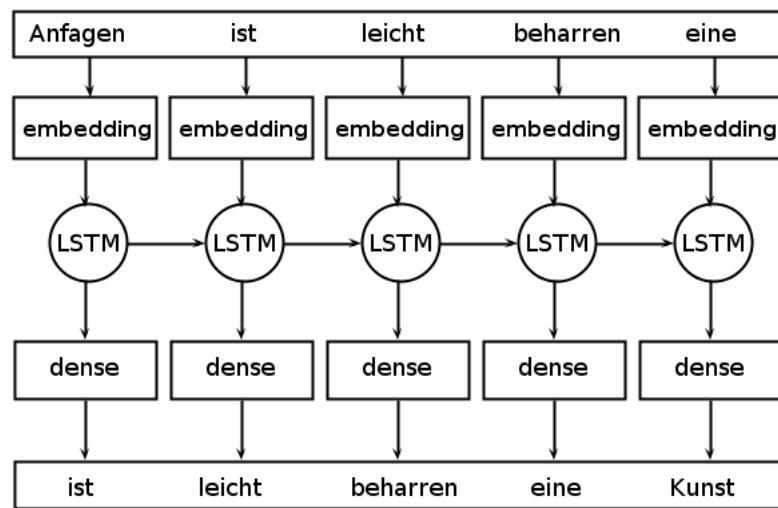


Рис. 5.14: Работа сети на примере фразы: Anfagen ist leicht, beharren eine Kunst.

## 5.5 Избирательность в нейросетевых подходах

Далее мы попробовали проинтерпретировать веса гребневой регрессии. Для этого мы рассмотрели тот lstm unit, вес которого по модулю наибольший и посмотрели его активации на слова из аудиостимула. Оказалось, что наибольшая активация достигается на собственный именах (см 5.15).

---

eine **computeranimation** auf einen schroffen berg mit schneebedeckter flanke fliegt eine reihe sterne zu sie bilden einen kranz um den gipfel **paramount** vor himmel schwebt eine grauweisse feder durch die luft in einer stadt segelt sie ueber die baumkronen eines **parks** ein **tom hanks** als **forrest gump** die feder fliegt ein paar meter ueber den wipfeln **robin** wright als **jenny gary** als lieutenant dan **williamson** als bubba und sally field als **gump** musik alan silvestri drehbuch **eric roth** nach dem gleichnamigen roman von **winston groom** regie robert in der naehe eines kirchturms sinkt die feder vor einem weissen gebaeude langsam tiefer die wurde produziert vom bayerischen rundfunk im auftrag des schweizer fernsehens text koop michalski und beckmann ev redaktion elmar dosch und bernd benecke tonmischung herbert glaser sprecher **thomas meinhardt** die feder

Рис. 5.15: Активации рекуррентного нейрона на слова аудиостимула. Наибольшие активации выделены жирным шрифтом.

# Часть 6

## Сравнение моделей

В следующей таблице представлено сравнение моделей по некоторым субъектам. Метод, использовавший скрытое состояние LSTM сети, показывает результаты не хуже гребневой регрессии, при этом обучаемых параметров у него меньше. Нейросети, предсказывающие отклик напрямую, в целом показывают результаты чуть лучше гребневой регрессии, несмотря на большое число обучаемых параметров (использовалась сильная регуляризация и дропаут).

Субъекты / Модели	Гребневая регрессия	LSTM(wiki)	GRU(stimulus)	FC(stimulus)
Субъект 1	0.48	0.5	0.52	0.53
Субъект 2	0.44	0.41	0.48	0.46
Субъект 3	0.43	0.45	0.47	0.46
Среднее значение	0.45	0.45	<b>0.48</b>	0.46
Число параметров	1200	<b>400</b>	36000	12010

# Часть 7

## Заключение

В заключение хотелось бы отметить несколько выводов, сделанных при работе:

- Подход, описанный в статье [4] воспроизводим на независимых данных и других параметрах эксперимента, при этом предобученный word2vec также показывает неплохие результаты.
- Зоны, лучше всего предсказываемые моделями, более всего реагируют на собственные имена.
- Рекуррентные сети показывают результаты чуть лучше гребневой регрессии засчет того, что улавливают контекст.
- Метод, использующий представление, полученное рекуррентной сети показывает результат не хуже гребневой регрессии, при этом у него в 4 раза меньше обучаемых параметров.

На защиту выносятся:

- Подготовлена экспериментальная среда, позволяющая развивать исследования с данными ([5]).
- Разработана и исследована линейная регрессионная модель BOLD-отклика.
- Разработан демонстрационный пользовательский интерфейс для исследования семантической избирательности модели.
- Разработаны и исследованы нелинейные нейросетевые модели BOLD-отклика, в том числе модели на основе полносвязной и рекуррентной нейронных сетях.

# Литература

- [1] Alexandre Abraham, Fabian Pedregosa, Michael Eickenberg, Philippe Gervais, Andreas Muller, Jean Kossaifi, Alexandre Gramfort, Bertrand Thirion, and G  el Varoquaux. Machine learning for neuroimaging with scikit-learn. *arXiv preprint arXiv:1412.3919*, 2014.
- [2] Junyoung Chung, Caglar Gulcehre, KyungHyun Cho, and Yoshua Bengio. Empirical evaluation of gated recurrent neural networks on sequence modeling. *arXiv preprint arXiv:1412.3555*, 2014.
- [3] Alex Graves and J  rgen Schmidhuber. Framewise phoneme classification with bidirectional lstm and other neural network architectures. *Neural Networks*, 18(5):602–610, 2005.
- [4] Michael Hanke, Florian J Baumgartner, Pierre Ibe, Falko R Kaule, Stefan Pollmann, Oliver Speck, Wolf Zinke, and J  rg Stadler. A high-resolution 7-tesla fmri dataset from complex natural stimulation with an audio movie. *Scientific data*, 1, 2014.
- [5] Alexander G Huth, Wendy A de Heer, Thomas L Griffiths, Fr  d  ric E Theunissen, and Jack L Gallant. Natural speech reveals the semantic maps that tile human cerebral cortex. *Nature*, 532(7600):453–458, 2016.
- [6] Yukiko Kikuchi, Adam Attaheri, Benjamin Wilson, Ariane E Rhone, Kirill V Nourski, Phillip E Gander, Christopher K Kovach, Hiroto Kawasaki, Timothy D Griffiths, Matthew A Howard III, et al. Sequence learning modulates neural responses and oscillatory coupling in human and monkey auditory cortex. *PLoS biology*, 15(4):e2000219, 2017.
- [7] Peng Qian, Xipeng Qiu, and Xuanjing Huang. Bridging lstm architecture and the neural dynamics during reading. *arXiv preprint arXiv:1604.06635*, 2016.

- [8] Frank Rosenblatt. The perceptron: A probabilistic model for information storage and organization in the brain. *Psychological review*, 65(6):386, 1958.