

УДК 004.89, 004.048, 004.67

МОДЕЛИ МАШИННОГО ОБУЧЕНИЯ ДЛЯ АВТОМАТИЧЕСКОГО ВЫЯВЛЕНИЯ НОВЫХ ФУНКЦИОНАЛЬНЫХ ПАТТЕРНОВ В МНОГОМЕРНЫХ ВРЕМЕННЫХ РЯДАХ

© 2024 г. А. И. Майсурадзе^a, *, Л. С. Сидоров^a, **

^aМГУ им. М.В. Ломоносова, Москва, Россия

*e-mail: maysuradze@cs.msu.ru

**e-mail: leon.sidorov@gmail.com

Поступила в редакцию 16.05.2024 г.

После доработки 09.06.2024 г.

Принята к публикации 16.10.2024 г.

Предлагается метод и конкретная архитектура модели машинного обучения, которая помогает исследователям в различных областях автоматически выявлять функциональные закономерности в многомерных временных рядах из серии экспериментов. Исходная задача была формализована в терминах машинного обучения, поэтому у исследователей нет необходимости быть профессионалами в рассматриваемой предметной области. Эффективность метода была продемонстрирована в области нейрофизиологии для данных, где уже известно о существовании приведенного паттерна P300. Для дальнейших исследований представляет интерес обобщить предложенный метод на другие области, например данные с датчиков на производственных линиях или банковские транзакции.

Ключевые слова: многомерные временные ряды, функциональные паттерны, электроэнцефалограмма, волна P300

DOI: 10.31857/S0002338824060106, EDN: SUBTIL

MACHINE LEARNING MODELS TO AUTOMATICALLY DISCOVER NOVEL FUNCTIONAL PATTERNS IN MULTIVARIATE TIME SERIES

А. И. Maysuradze^a, *, Л. С. Sidorov^a, **

^aLomonosov Moscow State University, Moscow, Russia

*e-mail: maysuradze@cs.msu.ru

**e-mail: leon.sidorov@gmail.com

In this paper, we propose a method and a specific architecture for a machine learning model that assists researchers across various fields in automatically identifying functional patterns in multivariate time series from a series of experiments. The initial problem was formalized in terms of machine learning, eliminating the need for researchers to be experts in the specific subject matter under examination. The effectiveness of the method has been demonstrated in the field of neurophysiology with data where the existence of the P300 pattern is already known. For further research, it would be beneficial to generalize the proposed method to other areas, such as sensor data from production lines or banking transactions.

Keywords: multivariate time series, functional patterns, electroencephalogram, P300 wave

Введение. Во многих предметных областях исследуемые данные имеют вид многомерных временных рядов. Это могут быть показания датчиков на производственных линиях, акции на фондовом рынке или истории денежных транзакций. Иногда такие временные ряды возникают как результат дополнительной обработки, например видео может преобразовываться в пу-

чок траекторий специфических точек. При этом на многомерных временных рядах решаются как традиционные задачи классификации, так и специфические задачи поиска аномалий или точки изменения тренда. Соответственно, для всех этих прикладных областей и задач разрабатывают различные модели машинного обучения, которые принимают на вход многомерные временные ряды.

В наши дни исследователи ждут от моделей машинного обучения не просто решения прикладных задач, но и интерпретируемости полученных результатов. Модели должны явно демонстрировать пользователям закономерности, которые они нашли. При этом анализ многомерных временных рядов во многих случаях сводится к выявлению так называемых функциональных паттернов [1], т.е. особенностей поведения временного ряда, соответствующих некоторым интересующим исследователей состояниям системы. На протяжении многих лет эксперты из различных предметных областей уже выявили некоторое количество функциональных паттернов. Основная задача данной работы – разработка методики автоматического поиска подобных паттернов в многомерных временных рядах без обладания какими-либо априорными знаниями о предметной области. Мы хотим предложить модель, которая как частный случай найдет уже известный паттерн, что подтвердит выдвинутую гипотезу. Развивая подобный подход, в будущем исследователи смогут использовать модели машинного обучения, чтобы быстрее находить новые виды паттернов и закономерностей в исходных данных многомерных временных рядов.

В работе в качестве предметной области взята нейрофизиология. Соответственно, работа методики будет продемонстрирована для выявления одного из самых известных функциональных паттернов в анализе электроэнцефалограммы (ЭЭГ), а именно волны Р300. Функциональные паттерны, на которые сделан упор в исследовании, служат обобщением данного понятия для случая произвольного многомерного временного ряда.

Паттерн ЭЭГ – это целостная картина биоэлектрической активности по всем исследуемым областям мозга, отражающая особенности распределения различных компонентов ЭЭГ. Он является качественной характеристикой ЭЭГ, отражающей функциональное состояние головного мозга человека.

Существуют нормальные паттерны, которые сопровождают нас в состоянии покоя или во сне. Они имеют тенденцию к мономорфности: состоят из повторяющихся волн со схожей формой и длиной волны. Всякое внешнее раздражение, примененное впервые, вызывает десинхронизацию и обнаружение частых колебаний.

В работе осуществляется анализ непостоянных паттернов, которые возникают в качестве реакции мозга на какие-либо внешние раздражители. В отличие от нормальных паттернов зачастую они состоят из одной волны и поэтому почти не подходят для визуального анализа. Однако подобные паттерны активно используются для взаимодействия человека с компьютером через BCI (*brain computer interface*).

Одна волна – это одиночное колебание потенциала любой амплитуды и формы. Амплитуда волны λ – величина колебания потенциала от пика до пика, измеряется в микровольтах (мкВ) и милливольтах (мВ). Период T – длительность интервала между началом и концом одиночной волны или комплекса волн. Эти величины связаны соотношением

$$T = \frac{\lambda}{c},$$

где c – это скорость распространения волны.

1. Формальная постановка задачи. У паттерна волны Р300 есть два определения: положительный импульс ($\lambda > 0$), возникающий в интервале от 250 до 500 мс, или же третий положительный пик после возникновения стимула (отсюда альтернативное название Р3).

Оба эти определения говорят о том, что данный паттерн характеризуется количеством и расположением пиков с положительной амплитудой на ЭЭГ, поэтому для дальнейшего анализа ключевое значение имеют операции локализации и измерения максимума (\max и $\arg\max$).

В данной работе эксперименты проводятся на наборе данных, где целевая переменная характеризует стимул, который должен вызывать волну Р300. Подробнее набор данных описан в разд. 1. Решается задача бинарной классификации, где 1 означает наличие воздействия, а 0 – его отсутствие. Эксперимент по постановке не гарантирует, что во время воздействия возникнет волна Р300, однако шанс этого события крайне велик.

В таком случае модель машинного обучения не получает информацию о наличие волны Р300 напрямую, однако может выделить ее в качестве метапризнака. Основная задача данной работы заключается в проверке этой гипотезы. Объектом выборки является многомерный временной ряд фиксированной длины. На каналах задано отношение соседства по цепи, поэтому

каждое наблюдение $X_i \in \mathbb{R}^{N \times T}$ можно задать матрицей размера $N \times T$, где N – количество каналов (электродов), а T – количество промежутков времени, в которые были осуществлены снятия показаний с прибора.

2. Анализ литературы. Отметим, что в нейрофизиологии базовым методом определения паттерна выступает усреднение временных рядов по каждому классу. Даже для ярко выраженных паттернов такой подход требует привлечения десятков респондентов. В наших данных присутствует только два респондента, соответственно, базовый метод не использовался ни в работе [1], ни в нашем исследовании.

Рассмотрим основные подходы интерпретации моделей машинного обучения в смежной области – медицине, где подобное свойство играет ключевое значение. Как и в разработанном нами подходе, возможность интерпретации результатов в этих сетях была заложена на этапе проектирования.

В работе [2] авторы использовали декодер сети U-Net, чтобы кодировать внутренние органы человека в разных масштабах. Последующее применение механизма внимания позволяет модели фокусироваться на ключевых участках исходного изображения и в явном виде визуализировать их вклад в итоговый прогноз. Эта техника напоминает анализ функциональных паттернов в многомерных временных рядах. Аналогичный подход, но для задачи сегментации изображений, был применен в работе [3].

Алгоритм глубокого обучения из работы [4] учитывает текущие и предыдущие снимки компьютерной томографии пациента для прогнозирования риска развития рака легких. Авторы использовали метод интегрированного градиента [5] для получения карт выраженности и пригласили опытных радиологов для проверки точности этих карт. Эксперты полностью согласились с тем, что модель действительно фокусировалась на опухолях.

В работе [6] предложили глубокую сверточную нейронную сеть для прогнозирования смертности, основанную на измерении 22 различных показателей в отделении интенсивной терапии, таких как индекс натрия, диурез и т.д. В данной работе интерпретация производилась посредством визуализации сверточных фильтров нейронной сети на разных масштабах.

Авторы [7] представили сеть, которая объединяет извлеченные из ResNet признаки и предварительные клинические знания для прогнозирования риска смертности пациентов на основе компьютерной томографии. Чтобы дополнительно подтвердить эффективность предложенной модели, они использовали метод t-SNE [8] для уменьшения размерности карт признаков злокачественных и доброкачественных новообразований и обнаружили, что злокачественные и доброкачественные наблюдения были хорошо разделены.

Как мы видим из приведенных примеров, в области медицины уже предпринимались многочисленные попытки интерпретации полученных моделью машинного обучения прогнозов, однако в основном они касались моделей для анализа компьютерных снимков и других задач компьютерного зрения. Наш же подход предполагает интерпретацию многомерных временных рядов и поиск новых типов закономерностей в анализируемых сигналах.

3. Предложенная архитектура. Основной подход в изучении многомерных временных рядов заключается в использовании архитектуры сверточной нейронной сети, которая применяет отдельные фильтры для анализа временных и пространственных зависимостей. Пространственные сверточные слои используются для извлечения признаков из данных ЭЭГ, отражающих активность различных электродов, в то время как временной сверточный слой извлекает признаки временных рядов. Затем полно связанный слой агрегирует всю полученную информацию для решения поставленной задачи. Этот подход был применен на рассматриваемом наборе данных в работе [9].

Подобное семейство архитектур нейронных сетей весьма богато и обладает хорошей интерпретируемостью. Модели такого типа состоят из двух блоков обработки, которые представлены пространственными и временными свертками соответственно. Анализируя каждый из них, можно выявлять пространственные и временные характеристики функциональных паттернов. Эти блоки могут быть применены в любом порядке, однако основной подход заключается в том, чтобы сначала обработать пространственную информацию, а только потом перейти к анализу временных зависимостей. Наша архитектура также будет придерживаться данного принципа.

Рассматриваемый набор данных состоит из относительно небольшого количества наблюдений с ярко выраженным дисбалансом классов, а специфика многомерных временных рядов не позволяет нам использовать большие предобученные модели. Поэтому цель этой работы заключается в разработке модели с относительно небольшим количеством параметров, кото-

рая потенциально была бы применима к многомерным временными рядам произвольной природы. По этой причине пространственный слой представляет собой свертку 1×1 по аналогии с [9], где выход слоя $z^s(j)$ формируется следующим образом:

$$z^s(j) = b^s + \sum_{i=0}^{N-1} w_i^s X_m(i, j).$$

Здесь $i = \overline{0, N-1}$, $j = \overline{0, T-1}$, N – количество электродов, а T – количество временных отсчетов, b^s – сдвиг пространственного слоя, w_i^s – вес i -го электрода в пространственном слое, а $X_m(i, j)$ – элемент в i -й строке и j -м столбце исходной матрицы наблюдения X_m .

В качестве слоя для обработки времени нами был выбран механизм внимания из [10] (*self attention*). Голова внимания (*attention head*) была взята лишь одна, чтобы добиться наименьшего количества параметров в итоговой модели.

Выход временного слоя $z^t(X) \in \mathbb{R}^{T \times 1}$ считается по следующей формуле:

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{D}}\right)V,$$

$$z^t(X) = \text{Attention}(W^Q X, W^K X, W^V X) W^{OT},$$

где $\text{Attention}(Q, K, V) \in \mathbb{R}^{T \times D}$ – оператор внимания, зависящий от матриц $Q, K, V \in \mathbb{R}^{T \times D}$, $X \in \mathbb{R}^{T \times 1}$ – вход слоя, а $W^Q, W^K, W^V, W^O \in \mathbb{R}^{1 \times D}$ – его матрицы весов, T – количество временных отсчетов, а D – внутренняя размерность механизма внимания, которая подбирается экспериментально. Если $Y \in \mathbb{R}^{A \times B}$, то обозначим

$$\text{softmax}(Y)_{ij} = e^{Y_{ij}} / \sum_{k=1}^B e^{Y_{ik}}, \quad i = \overline{1, A}, \quad j = \overline{1, B}.$$

Механизм внимания способен учитывать контекст каждого наблюдения, т.е. он также способен воспроизвести операции *max* и *argmax*. Согласно разд. 1, этого достаточно, чтобы определять наличие волны Р300.

Предложенная нами модель машинного обучения способна определять амплитуду волны λ и оценивать ее форму посредством поиска минимума и максимума, что в контексте анализа волны Р300 имеет ключевое значение. Нейронная сеть поочередно применяет фильтры к каналам и временными промежуткам на исходных данных, выявляя тем самым важность тех или иных электродов на голове человека или моментов времени во время записи.

Итоговая модель (рис. 1) состоит из свертки 1×1 для обработки пространственной информации и слоя внимания [10] для поиска функциональных паттернов во времени. На выходе расположен один линейный слой для решения задачи классификации на основе полученных представлений. Кроме того, каждый блок, кроме линейного слоя на выходе, сопровождается нелинейной функцией ReLU (*rectified linear unit*) [11].

Приведенная модель использует идеи визуализации фильтров в сверточных слоях и анализ карт внимания для интерпретации полученных результатов и локализации функционального паттерна. При решении исходной задачи модели предлагается сравнить состояние покоя пациента с реакцией на внешнее воздействие.

4. Рассматриваемый набор данных. Данные были собраны с помощью парадигмы странностей (англ. *oddball paradigm*). *Oddball paradigm* – это дизайн эксперимента, используемый в психологических и нейрофизиологических исследованиях. Испытуемому демонстрируют последовательности повторяющихся стимулов, которые изредка прерываются нестандартным воздействием. Реакция испытуемого на этот “странный” стимул впоследствии записывается и анализируется, подробнее этот процесс описан в работе [12].

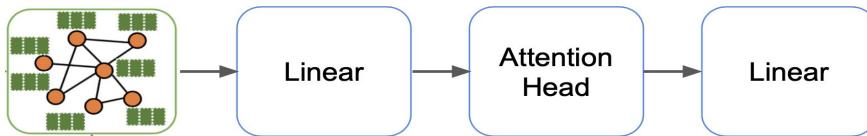


Рис. 1. Предложенная архитектура.

Пользовательские интерфейсы для сбора данных были представлены в работе [13]. Всего в этой работе описывались семь подобных интерфейсов, однако мы решили сфокусироваться на одном функциональном паттерне и рассматривать лишь те парадигмы, что успешно применялись для изучения волны Р300. Такие подходы генерируют большое количество наблюдений, что позволяет успешно обучать искусственные нейронные сети.

4.1. П р о ц е с с с б о р а д а н н ы х. Представленный набор данных был записан с использованием символьной матрицы, с помощью которой человек может писать слова, концентрируя свое внимание на отдельных символах. Этот набор данных является частью соревнования [1] и содержит множество экспериментальных результатов, полученных с применением большого разнообразия моделей. Исходно данные были собраны с целью понять, насколько хорошо алгоритмы могут предсказать целевой символ из данных о мозговой активности. Соответственно, в работе [1] также описаны существующие подходы из классического машинного обучения, которые уже достигают достаточно высокого уровня качества на представленном наборе данных. Но в то время исследователи еще не рассматривали, что именно находят их модели.

Продолжительность одного фрагмента не фиксирована и зависит от постановки задачи. Данные могут быть разделены на фрагменты разной длины. Однако волна Р300 может быть обнаружена в масштабе 1 с, поэтому нет смысла использовать более длительные временные интервалы. Мы работали с фрагментами по 600 мс.

Набор данных содержит ЭЭГ двух здоровых добровольцев (А и В). Графический пользовательский интерфейс для проведения эксперимента показан на рис. 2. Он представляет собой матрицу 6×6 символов. В нейтральном состоянии все символы имеют одинаковую низкую яркость. Стока или столбец могут быть временно подсвечены.



Рис. 2. Интерфейс для сбора данных из [13].

Для очередного целевого символа проводят 15 блоков подсвечиваний матрицы, которые в совокупности называются эпохой. В каждом блоке каждая из строк и столбцов подсвечивается по одному разу, т.е. всего 12 подсвечиваний. В блоке столбцы и строки загорались в случайном порядке, т.е. в разных блоках одной эпохи порядок подсвечиваний разный. Испытуемому предлагалось сосредоточиться на целевом символе и молча считать количество мигающих повторений строки и столбца, содержащих целевой символ. Символ определяется парой из строки и столбца. Таким образом, в каждом блоке целевой символ подсвечивается дважды, всего в эпохе – 30 раз. Для каждого подсвечивания легко определить, участвовал ли в нем целевой символ.

Все сигналы ЭЭГ были собраны с помощью 64-электродного скальпа, отфильтрованы в диапазоне от 0.1 до 60 Гц и оцифрованы с частотой 240 Гц. Более подробно прочитать об этом можно в работе [1].

4.2. П р е о б р а з о в а н и е ц е л е в о г о п р и з н а к а. Согласно исследованиям нейрофизиологов, предполагается, но формально не гарантируется, что подсвечивание целевого символа дает волну Р300. Хотя мы знаем, когда стоит ожидать сигнал Р300, его появление зависит от испытуемого, причем сам испытуемый тоже не контролирует этот процесс. Производство волны Р300 не является феноменом сознания, оно возникает из-за внешних стимулов (мерцание строк и столбцов). Кроме того, волна Р300 интерферирует со многими другими волнами, соответственно, ее выраженность может сильно варьироваться.

Для рассматриваемого набора данных часто задача ставится следующим образом. Проверяется, содержит ли целевой символ в подсвечиваемых строке или столбце. В таком случае задача может рассматриваться как бинарная классификация. На вход подается ЭЭГ, соответствующая подсвечиванию одной строки или столбца, а выходом служит бинарная метка класса (является ли текущее подсвечивание целевым). Как уже говорилось выше, на каждое подсвечивание отводится запись длиною в 600 мс, что больше продолжительности самого подсвечивания, однако последний объект эпохи не обрежется, потому что каждая запись заканчивается паузой в 2.5 с, во время которой матрица гаснет.

Такая постановка задачи позволяет нам сфокусироваться на выявлении функциональных паттернов и интерпретации полученных результатов. Итоговые размеры обучающей и тестовой выборок для каждого из пациентов приведены в табл. 1.

Нейрофизиологи выявили, что во время подсвечивания целевого символа в человеческом мозге возникает волна Р300. Предполагается, что символ соответствует пересечению накопле-

Таблица 1. Размеры выборок для каждого из пациентов

Целевой класс	Обучающая выборка	Тестовая выборка
Есть подсвечивание	2550	3000
Нет подсвечивания	12 750	15 000

ния нескольких волн Р300. Наибольшее накопление волн Р300 для вертикальных мигающих огней определяет столбец желаемого символа. То же справедливо и для горизонтальных мигающих огней и строк.

Хотя из обучающей разметки мы явно знаем, когда загорался целевой символ, эксперимент по постановке не гарантирует, что во время воздействия возникнет именно волна Р300, хотя шанс этого события и крайне велик. Соответственно, модель не получает информацию о наличии волны Р300 напрямую. Мы хотим убедиться, что модель машинного обучения действительно сможет самостоятельно найти функциональный паттерн, который нейрофизиологам уже известен.

5. Эксперименты. Для проведения экспериментов будем использовать предоставленное авторами из работы [1] разделение выборки на обучение и тест. Таким образом, можно сравнить наши алгоритмы с другими подходами, в том числе с теми, которые были разработаны во время оригинального соревнования на используемых данных.

Как отмечалось ранее, обучающая выборка состоит из 85 символов, а тестовая – из 100. При переходе от классификации символов к классификации строк и столбцов получаем, соответственно, обучающую выборку из 15 300 наблюдений и тестовую выборку из 18 000 наблюдений. При обучении важен порядок поступления данных, лучших результатов мы добились при соответствии постановке эксперимента, который проводился при записи исходных данных. Мы установили размер мини-батча $K = 180$, именно столько наблюдений описывает один целевой символ из исходной задачи распознавания целевых символов.

Согласно *oddball paradigm*, на каждые 12 объектов приходится 2 наблюдения положительного класса и 10 отрицательного. Следовательно, в задаче существует значительный дисбаланс классов (1 к 6). Кроме того, каждый батч содержит ровно 30 объектов положительного класса и 150 объектов отрицательного, что гарантируется из постановки эксперимента. Для подобной классической проблемы в качестве функции ошибки L была выбрана бинарная кросс-энтропия:

$$L = - \sum_{n=1}^K \left[y_n \log x_n + (1 - y_n) \log(1 - x_n) \right],$$

где $y_n \in \{0,1\}$ – метки классов, а $x_n \in [0,1]$ – вероятность положительной метки, полученная из модели.

5.1. П р о с т р а н с т в е н н ы й а н а л и з. Теперь, когда процедура обучения определена, можно перейти непосредственно к анализу полученных результатов. Для начала визуализируем веса свертки 1×1 , которая является первым этапом обработки входных данных. Благодаря своей структуре она выучивает некоторый вес в соответствие каждому электроду на голове человека. Таким образом, на выходе получаем карту важности электродов с точки зрения модели (рис. 3), другими словами, локализацию функционального паттерна в пространстве.

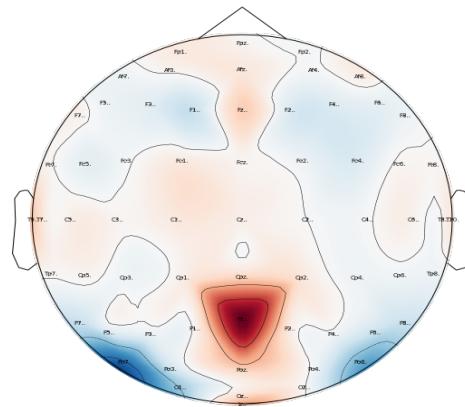


Рис. 3. Карта важности каналов в линейном слое.

Заметим, что модель самостоятельно выявила оптимальную линейную комбинацию исходных электродов. Рс7 и Рс8 были взяты с отрицательным знаком, чтобы учесть негативные значения потенциала, предшествующие Р300, а электрод Рz, наоборот, — с положительным, чтобы уловить пик самой волны. Эту гипотезу подтверждают значения потенциала на выделенных электродах (рис. 4), для построения графика было взято произвольное наблюдение из тестовой выборки, на котором присутствует волна Р300. Для Рz характерен провал в значении потенциала непосредственно перед пиком Р300, а Рс7 и Рс8 вместе образуют самый высокий пик на данных ЭЭГ.

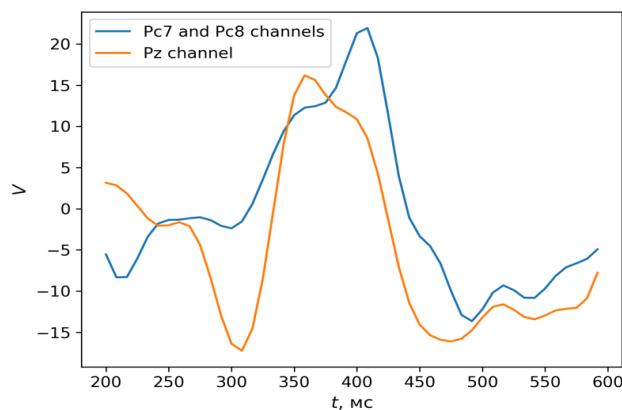


Рис. 4. Потенциал на выделенных электродах.

Также важно отметить, что модель предложила использовать всего три канала ЭЭГ, выделив их как точки концентрации информации о наблюдаемом явлении. Этот результат не противоречит нашим знаниям о человеческом мозге, ведь именно затылочная его доля отвечает за распознавание визуальных образов, которыми и являются целевые символы на экране компьютера.

5.2. Временной анализ. Также немаловажную роль в интерпретации результатов сыграл и слой с механизмом внимания (рис. 5). На графике с картой внимания для одного конкретного наблюдения волны Р300 можно заметить, что модель концентрирует свое внимание на максимальном и минимальном значениях промежуточного представления ЭЭГ. Это говорит о том, что нейронная сеть действительно самостоятельно приходит к преобразованиям вида max и argmax, что служит необходимым признаком для выявления функционального паттерна Р300, согласно разд. 1.

Взаимное расположение положительных и отрицательных пиков играет ключевую роль в обоих определениях волны Р300. Таким образом, предложенная модель может не только локализовать возникновение функционального паттерна в пространстве, но и определять наличие паттерна по его характерным особенностям. Другими словами, алгоритм выявил вол-

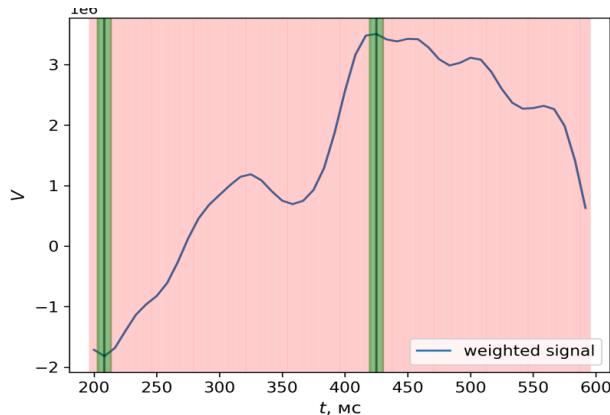


Рис. 5. Карта внимания модели для преобразованной волны P300.

ну P300 из данных, хотя изначальная постановка задачи говорила о его существовании лишь косвенно, через этап сбора данных. При этом данные наших экспериментов (табл. 2) подтверждают, что качество решения исходной задачи бинарной классификации сразу же оказалось на уровне лучших алгоритмов из работы [1] и нейросетевого подхода из [9].

Таблица 2. Сравнение точности бинарной классификации с классическими алгоритмами

Модель	Пациент А, %	Пациент В, %
Логистическая регрессия	65.9	68.4
Градиентный бустинг	67.6	69.0
Сверточная нейросеть	72.4	75.7
Предложенная модель	72.7	75.2

Заключение. Предложена методика, помогающая исследователям по серии экспериментов автоматически выявить функциональный паттерн в многомерных временных рядах. При этом было достаточно формализовать исходную задачу в терминах машинного обучения и не требовалось углубляться в предметную область.

Работоспособность методики продемонстрирована на данных ЭЭГ. Модель выявляет паттерн волны P300 с высокой точностью, а также способна явно выделить его отличительные признаки. Вся эта информация была неявно извлечена из постановки задачи.

Данный подход может быть доработан и расширен для других предметных областей, что и представляет интерес для дальнейших исследований.

СПИСОК ЛИТЕРАТУРЫ

- Blankertz B., Muller K.R., Krusienski D.J., Schalk G., Wolpaw J.R., Schlogl A., et al. The BCI Competition iii: Validating Alternative Approaches to Actual BCI Problems // IEEE Transactions on Neural Systems and Rehabilitation Engineering. 2006. V. 14 (2). P. 153–159.
- Oktay O., Schlemper J., Folgoc L.L., Lee M., Heinrich M., Misawa K. et. al. Attention U-Net: Learning Where to Look for the Pancreas // arXiv preprint arXiv:1804.03999, 2018.
- Lee H., Yune S., Mansouri M., Kim M., Tajmir S.H., Guerrier C.E. et. al. An Explainable Deep-learning Algorithm for the Detection of Acute Intracranial Haemorrhage from Small Datasets // Nature Biomedical Engineering. 2019. V. 3 (3). P. 173–182.
- Ardila D., Kiraly A.P., Bharadwaj S., Choi B., Reicher J.J., Peng L. et. al. End-to-end Lung Cancer Screening with Three-dimensional Deep Learning on Low-dose Chest Computed Tomography // Nature Medicine. 2019. V. 25 (6). P. 954–961.
- Sundararajan M., Taly A., Yan Q. Axiomatic Attribution for Deep Networks // Intern. Conf. on Machine Learning. PMLR. Sydney, Australia, 2017. P. 3319–3328.
- Caicedo-Torres W., Gutierrez J. ISeeU: Visually Interpretable Deep Learning for Mortality Prediction Inside the ICU // J. Biomedical Informatics. 2019. V. 98. P. 103269.

7. *Guo H., Kruger U., Wang G., Kalra M.K., Yan P.* Knowledge-based Analysis for Mortality Prediction from CT Images // IEEE Biomedical and Health Informatics. 2019. V. 24 (2). P. 457–464.
8. *Van der Maaten L., Hinton G.* Visualizing Data Using t-SNE // J. Machine Learning Research. 2008. V. 9(11).
9. *Cecotti H., Graser A.* Convolutional Neural Networks for P300 Detection with Application to Brain-computer Interfaces // IEEE Transactions on Pattern Analysis and Machine Intelligence. 2010. V. 33 (3). P. 433–445.
10. *Vaswani A., Shazeer N., Parmar N., Uszkoreit J., Jones L. et al.* Attention is All you Need // Advances in Neural Information Processing Systems. 2017. V. 30.
11. *Krizhevsky A., Sutskever I., Hinton G. E.* Imagenet Classification with Deep Convolutional Neural Networks // Advances in Neural Information Processing Systems. 2012. V. 25.
12. *Squires N.K., Squires K.C., Hillyard S.A.* Two Varieties of Long-latency Positive Waves Evoked by Unpredictable Auditory Stimuli in Man // Electroencephalography and Clinical Neurophysiology. 1975. V. 38 (4). P. 387–401.
13. *Schalk G., McFarland D.J., Hinterberger T., Birbaumer N., Wolpaw J.R.* BCI2000: A General-purpose Brain-computer Interface (BCI) System // IEEE Transactions on Biomedical Engineering. 2004. V. 51 (6). P. 1034–1043.