

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ
МОСКОВСКИЙ ФИЗИКО-ТЕХНИЧЕСКИЙ ИНСТИТУТ (государственный университет)
ФАКУЛЬТЕТ УПРАВЛЕНИЯ И ПРИКЛАДНОЙ МАТЕМАТИКИ
ВЫЧИСЛИТЕЛЬНЫЙ ЦЕНТР ИМ. А. А. ДОРОДНИЦЫНА РАН
КАФЕДРА «ИНТЕЛЛЕКТУАЛЬНЫЕ СИСТЕМЫ»

Жуков Андрей Александрович

Классификация движений в однонаправленном нейрокомпьютерном интерфейсе

010900 — Прикладные математика и физика

ВЫПУСКНАЯ КВАЛИФИКАЦИОННАЯ РАБОТА БАКАЛАВРА

Научный руководитель:
д. ф.-м. н. Воронцов Константин
Вячеславович

Москва

2015

Содержание

1	Введение	4
2	Постановка задачи	5
3	Методы формирования признаков	6
3.1	Метод главных компонент — PCA	6
3.2	Метод независимых компонент — ICA	6
3.3	Использование коэффициентов ковариации сигнала	7
3.4	Метод общих пространственных структур — CSP	8
4	Вычислительный эксперимент	9
4.1	Описание исходных данных и общего алгоритма исследования	9
4.2	Предварительная фильтрация	9
4.3	Алгоритмы классификации	10
4.4	Обозначения моделей	12
4.5	Определение числа компонент в используемых моделях	12
4.6	Оценка точности полученных моделей по другим экспериментам	13
5	Заключение	14

Аннотация

Данная работа посвящена классификации движений конечностей человека по данным электроэнцефалографии для последующего применения в нейрокомпьютерных интерфейсах. Исследуется эффективность различных способов формирования пространства признаков и методов машинного обучения для решения данной задачи. В качестве признаков используются показания электроэнцефалографа пациентов, выполняющих моторную активность. Предлагается использовать метод адаптивного бустинга для повышения точности классификации.

Ключевые слова: *нейрокомпьютерный интерфейс, метод главных компонент, метод общих пространственных структур, машина опорных векторов, адаптивный бустинг.*

1 Введение

Нейрокомпьютерные интерфейсы являются одним из перспективных прикладных направлений современной нейрофизиологии. Данная технология позволяет увеличить эффективность терапии, направленной на восстановление моторных функций после инсультов. В развитых странах инсульты являются одной из основных причин инвалидности и серьезно влияют на качество жизни переживших. Для переживших инсульт восстановление моторной активности необходимо для возможности выполнения ежедневных задач, но это восстановление, как правило, проходит по-разному для разных больных и, как правило, не полно [1]. Восстановлению моторной активности способствует активное использование участков мозга, ответственных за выполнение движений, и нейрокомпьютерный интерфейс позволяет создать положительную обратную связь для пациентов, для которых получение её иначе было бы затруднено, что приводит к повышению эффективности восстановительной терапии [2]. Нейрокомпьютерный интерфейс также способен существенно улучшить качество жизни больных боковым амиотрофическим склерозом, расширяя их возможности по самостоятельному взаимодействию с окружающим миром.

Нейрокомпьютерный интерфейс может считывать информацию о мозговой активности с помощью нескольких различных технологий, каждая из которых имеет свои достоинства и недостатки [3]. В данной работе используются данные электроэнцефалографии (далее — ЭЭГ). Использование ЭЭГ в нейрокомпьютерном интерфейсе удобно тем, что устройство для считывания данных ЭЭГ компактно и мобильно, а также обеспечивает высокое временное разрешение сигнала и малое время отклика, что необходимо для удобства взаимодействия с пользователем [3].

Нейрокомпьютерный интерфейс должен с высокой точностью распознавать намерения пользователя, разделяя активность его активность на несколько определенных типов движения. Данная задача является типичной задачей классификации. Несмотря на большое количество исследований на эту тему, задача не потеряла актуальность, так как не был предложен алгоритм, который позволял бы распознавать движения произвольного пользователя с высокой точностью [4, 5].

2 Постановка задачи

Дана выборка — временной ряд $\mathfrak{D} = \{(\mathbf{x}_t, y_t)\}, t \in \mathcal{T} = \{1 \dots T\}$, состоящая из наблюдений в T различных моментов времени. Каждое наблюдение описывается показаниями n различных сенсоров $\mathbf{x}_t \in \mathbb{R}^n$ и принадлежит одному из 3 классов $y_t \in \mathbf{Y} = \{1, 2, 3\}$, характеризующих, соответственно, состояния покоя и движения левой и правой рукой. Исходное пространство объекты—признаки включает нестационарные колебательные процессы, происходящие на различных частотах, которые не поддаются интерпретации классификаторами при прямом использовании временного ряда как обучающей выборки, поэтому необходимо осуществить преобразование пространства признаков. Для этого предлагается разбить исходный временной ряд на отрезки определенной длины, осуществить на них фильтрацию сигнала в различных диапазонах частот, а затем выделить для каждого из отрезков признаки. Для выделения признаков используются различные методы, и мы рассмотрим наиболее популярные из них, такие, как:

- использование коэффициентов ковариации сигнала
- выделение общих пространственных структур (Common Spatial Patterns, CSP)
- метод главных компонент (Principle Component Analysis, PCA)
- анализ независимых компонент (Independent Component Analysis, ICA)

Результатом применения этих методов является модифицированная выборка $\mathfrak{D}' = \{(\mathbf{x}'_{t'}, y'_{t'})\}, t' \in \mathcal{T}' = \{1 \dots T'\}$ меньшей размерности, состоящая из меньшего количества объектов. Полученная выборка используется для обучения классификатора.

Таким образом, исходная задача разделяется на две подзадачи:

1. Определение оптимального пространства признаков \mathbf{x}' .
2. Выбор оптимального классификатора \mathbf{a} .

В качестве критерия качества используется оценка точности классификации с помощью перекрестной проверки. С ее помощью задачу можно сформулировать в терминах оптимизации:

$$CV(\mathbf{x}', \mathbf{a}, \mathfrak{D}') \xrightarrow{\mathbf{a}, \mathbf{x}'} \max$$

3 Методы формирования признаков

Ниже приведен обзор используемых в данной работе методов обработки сигналов.

3.1 Метод главных компонент — PCA

Метод главных компонент широко применяется при преобразованиях признаков для уменьшения размерности [6], в том числе и в задачах анализа мозговой активности [5, 7, 8]. Метод основан на последовательном подборе ортогональных компонент линейного преобразования, каждая из которых максимизирует дисперсию в проекции матрицы наблюдений на данную компоненту.

Пусть \mathbf{X} — матрица наблюдений. Найдем первую такую компоненту:

$$\mathbf{w}_{(1)} = \underset{\mathbf{w}}{\operatorname{argmax}} \frac{\|\mathbf{X}\mathbf{w}\|^2}{\|\mathbf{w}\|^2}.$$

В общем случае (на k -ом шаге) введем остаточную матрицу

$$\hat{\mathbf{X}}_k = \mathbf{X} - \sum_{s=1}^{k-1} \mathbf{X}\mathbf{w}_{(s)}\mathbf{w}_{(s)}^T,$$

и будем находить k -ую компоненту в виде

$$\mathbf{w}_{(k)} = \underset{\|\mathbf{w}\|=1}{\operatorname{argmax}} \left\{ \|\hat{\mathbf{X}}_k\mathbf{w}\|^2 \right\}.$$

Векторы $\mathbf{w} \in \mathbf{W}$ называются главными компонентами (principal components). Преобразованную матрицу наблюдений можно выразить через матрицу главных компонент и исходную матрицу наблюдений:

$$\mathbf{X}' = \mathbf{X}\mathbf{W}.$$

При разбиении матрицы наблюдений на отрезки наблюдений $\mathbf{X}_{t'}$ преобразованную матрицу наблюдений можно выразить аналогичным образом: $\mathbf{X}'_{t'} = \mathbf{X}_{t'}\mathbf{W}$. Данная матрица используется для признакового описания выбранного отрезка наблюдений.

3.2 Метод независимых компонент — ICA

Метод независимых компонент (Independent Component Analysis) является другим способом обработки сигналов и используется в задачах определения независимых

источников [9]. В частности, данный метод использовался в задачах классификации моторной активности по данным ЭЭГ [7, 10]. В рамках модели независимых компонент предполагается, что каждое наблюдение является линейной комбинацией независимых случайных величин.

Предположим, что каждое отдельное наблюдение \mathbf{x}_t является линейной комбинацией n независимых и ненормальных случайных величин \mathbf{s}_t :

$$\mathbf{x}_t = \mathbf{A}\mathbf{s}_t.$$

Тогда исходную матрицу наблюдений \mathbf{X} можно записать в виде

$$\mathbf{X} = \mathbf{A}\mathbf{S},$$

Матрицу наблюдений случайных величин можно выразить через матрицу наблюдений исходных:

$$\mathbf{S} = \mathbf{X}\mathbf{A}^{-1}.$$

Таким образом, для разбиения исходных наблюдений на временные отрезки $\mathbf{X}_{t'}$ имеем преобразованную матрицу наблюдений $\mathbf{X}'_{t'} = \mathbf{X}_{t'}\mathbf{A}^{-1}$, которая используется для признакового описания данного отрезка наблюдений. Матрица \mathbf{A} оценивается итерационным методом, который подробно описан в работе [9].

3.3 Использование коэффициентов ковариации сигнала

В данном методе предполагается, что наблюдения различных классов являются случайными величинами из различных семейств нормальных распределений $\mathbf{x}_t \sim N(0, \Sigma_y)$. Искомые распределения оцениваются по обучающей выборке: $\mathbf{x}_t \sim N(0, \hat{\Sigma}_y)$. Для определения класса каждого отдельного наблюдения можно воспользоваться методом максимального правдоподобия:

$$\mathbf{a}(\mathbf{x}_t) = \arg \max_{y \in \mathbf{Y}} p(y)p(\mathbf{x}_t|y)$$

Такой способ классификации наблюдений не дает достаточно высокой точности классификации и является неустойчивым к шуму. Предварительные эксперименты показали, что для повышения надежности целесообразно объединять отдельные наблюдения в последовательные интервалы: $\mathbf{x}_{t_j} \rightarrow \mathbf{X}'_i$, где $t_j \in t'_i$, все наблюдения принадлежат одному классу и t'_i покрывает интервал порядка одной секунды. Для измененной

выборки можно записать метод максимального правдоподобия в следующем виде:

$$\mathbf{a}(\mathbf{X}_{t'}) = \arg \max_{y \in \mathbf{Y}} \prod_{t \in t'} p(y)p(\mathbf{x}_t|y)$$

Предлагается использовать оценки ковариаций на каждом отдельном интервале t'_i в качестве признаков описания данных интервалов: $\mathbf{X}'_j = \hat{\Sigma}_j$. Такой подход даст схожую точность классификации, позволяет использовать широкий диапазон алгоритмов классификации и позволяет с легкостью добавлять в описания объектов дополнительные признаки.

3.4 Метод общих пространственных структур — CSP

Метод общих пространственных структур (Common Spatial Patterns) является наиболее широко используемым методом при формировании признаков для классификации движений [4, 11–16]. В данном методе находится линейное преобразование исходной матрицы сигналов, максимизирующее отношение дисперсий преобразованных матриц для двух различных классов.

Пусть $\mathbf{X}_1, \mathbf{X}_2$ — матрицы наблюдений для двух классов. Тогда найдем матрицу фильтров (общих пространственных структур) с m компонентами:

$$\mathbf{W}^* = \underset{\mathbf{W}}{\operatorname{argmax}} \frac{\|\mathbf{X}_1 \mathbf{W}\|^2}{\|\mathbf{X}_2 \mathbf{W}\|^2}.$$

Если записать матрицы ковариации для наблюдений двух классов $\Sigma_i = \mathbf{X}_i \mathbf{X}_i^T$, $i \in \{1, 2\}$, задачу поиска компонент \mathbf{w}_i матрицы \mathbf{W} можно записать [4] в виде задачи оптимизации

$$\min \sum_{i=1}^m \mathbf{w}_i \Sigma_2 \mathbf{w}_i^T + \sum_{i=m+1}^{2m} \mathbf{w}_i \Sigma_1 \mathbf{w}_i^T$$

с ограничениями

$$\mathbf{w}_i (\Sigma_1 + \Sigma_2) \mathbf{w}_i^T = 1, \quad i \in \{1, \dots, 2m\},$$

$$\mathbf{w}_i (\Sigma_1 + \Sigma_2) \mathbf{w}_j^T = 0, \quad i \neq j, \quad i, j \in \{1, \dots, 2m\}.$$

Векторы $\mathbf{w} \in \mathbf{W}$ называются векторами пространственных фильтров (структур). Для отрезка наблюдения \mathbf{X}_i преобразованной матрицей наблюдений будет $\mathbf{Z}_i = \mathbf{X}_i \mathbf{W}$.

4 Вычислительный эксперимент

4.1 Описание исходных данных и общего алгоритма исследования

В вычислительной части работы используются данные 6 экспериментов ЭЭГ, проведенных с одним пациентом, выполняющего 3 вида активности: отдых (класс 1) и движения левой и правой рук (классы 2 и 3). Мозговая активность измеряется с частотой 500Гц с помощью 40 сенсоров, расположенных на поверхности головы пациента. Сенсоры ЭЭГ измеряют потенциалы электрического поля в микровольтах.

Исходный сигнал разбивается на отрезки времени в одну секунду, после чего осуществляется предварительная фильтрация сигнала. К фильтрованным сигналам применяются описанные выше преобразования признаков. Преобразованные признаки используются для оценки классификации активности. Точность классификации оценивается помощью стократной перекрестной проверки.

4.2 Предварительная фильтрация

Сигнал, получаемый при использовании сенсоров, имеет характеристики, затрудняющие классификацию. Во-первых, на целевой сигнал наложены колебания сети переменного тока 50Гц, обеспечивающей питание прибора. Амплитуда колебаний 50Гц на 4 порядка превышает амплитуду низкочастотных колебаний. Во-вторых, показания приборов являются смещенными, причем величина смещения со временем меняется. Предполагается, что смещение связано с характеристикой измеряющих приборов. Для решения таких проблем применялись различные методы:

- Применение частотных фильтров с бесконечной импульсной характеристикой с различными частотными диапазонами
- Применение сезонного дифференцирования для сезонности 50Гц: $\mathbf{x}_t = \mathbf{x}_t - \mathbf{x}_{t-10}$
- Выравнивание тренда путем вычитания среднего по отрезку: $\mathbf{x}_t = \mathbf{x}_t - \text{mean}_{t \in t'}(\mathbf{x}_t)$

В ходе экспериментов было получено, что при работе с коэффициентами ковариации сигнала наибольшая точность достигается при сезонном дифференцировании сигнала. При использовании методов CSP, PCA, ICA лучшие результаты получаются при вычитании среднего.

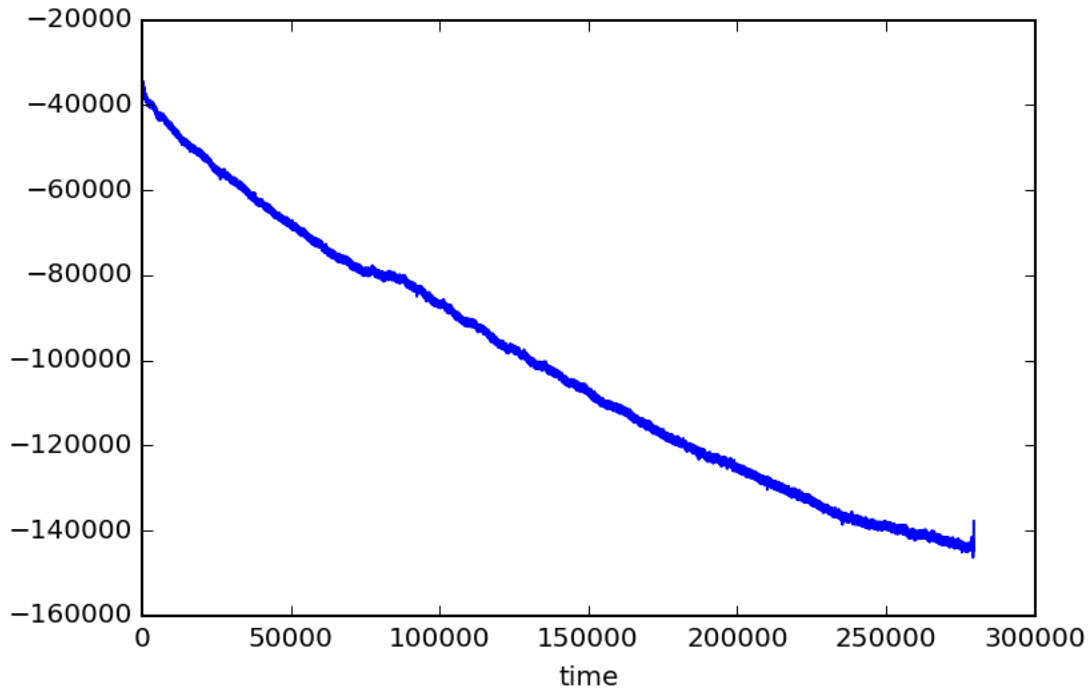


Рис. 1: Смещение сигнала с течением времени

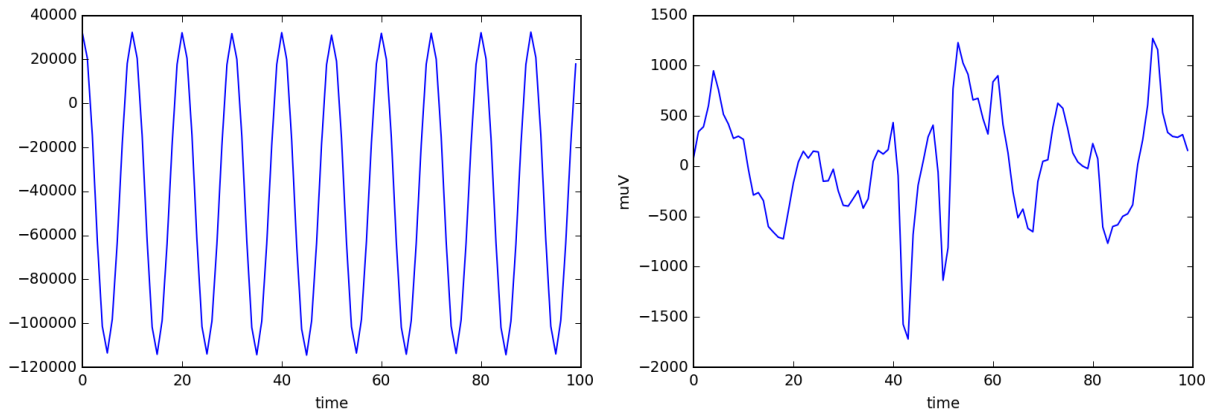


Рис. 2: Вид сигнала до и после сезонного дифференцирования

4.3 Алгоритмы классификации

Для классификации были использованы алгоритмы машины опорных векторов (SVM) и адаптивного бустинга (AdaBoost), реализованные в библиотеке scikit-learn [17]. Использование SVM обусловлено тем, что это наиболее часто используемый метод для классификации моторной активности. Использование AdaBoost обусловлено гипотезой о существенно линейной неразделимости классов в пространстве признаков (характерный вид признакового пространства при обработке сигналов с помощью CSP рис. 3), формируемом методом CSP, и вместе с тем возможностью применения решающих деревьев для классификации. формой пространства признаков.

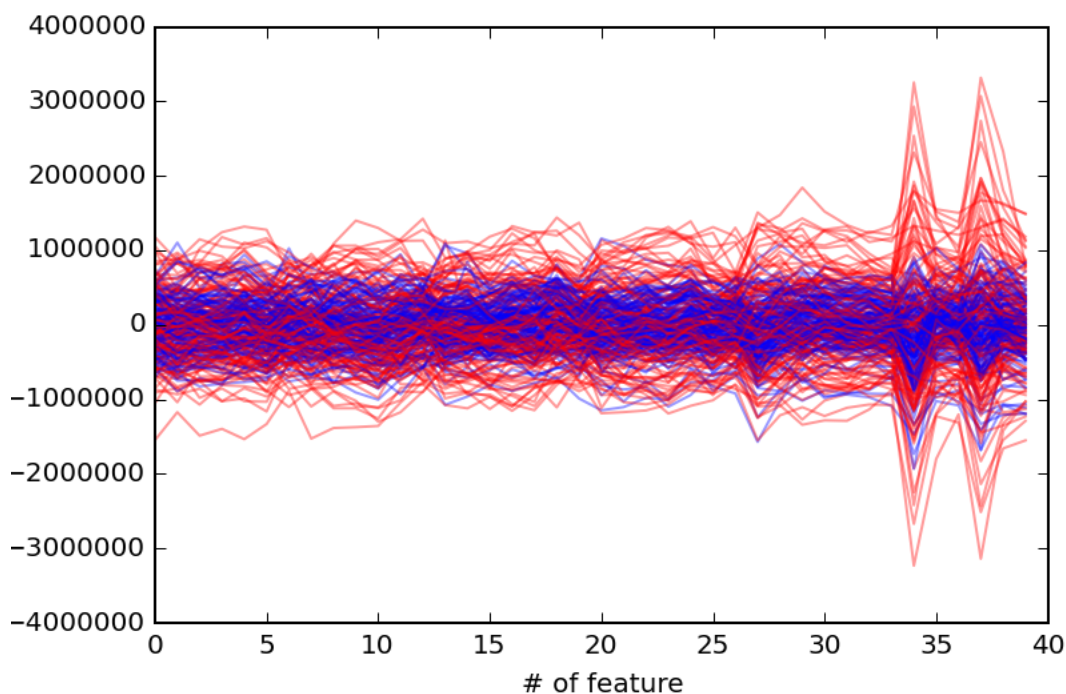


Рис. 3: Вектор компонент сигнала после фильтрации CSP для 2х классов

Оптимальным в модели SVM было использование L2 регуляризации с параметром $C = 0.01$. В модели AdaBoost оптимальное количество классификаторов было определено равным $n = 75$.

4.4 Обозначения моделей

- PCA — Использование метода главных компонент с SVM-классификатором
- ICA — Использование метода независимых компонент с SVM-классификатором
- CSP — Использование метода общих пространственных структур с SVM-классификатором
- AdaCSP — Использование метода общих пространственных структур с AdaBoost-классификатором
- Cov — Использование метода коэффициентов матрицы ковариации с SVM-классификатором
- CovICA — Использование метода коэффициентов матрицы ковариации с последующим применением ICA и SVM-классификатором
- CovPCA — Использование метода коэффициентов матрицы ковариации с последующим применением PCA и SVM-классификатором

4.5 Определение числа компонент в используемых моделях

Число компонент для моделей PCA, CSP, ICA, CovPCA, CovICA было по одному из экспериментов с пациентом. Исследовалась классификация движений.

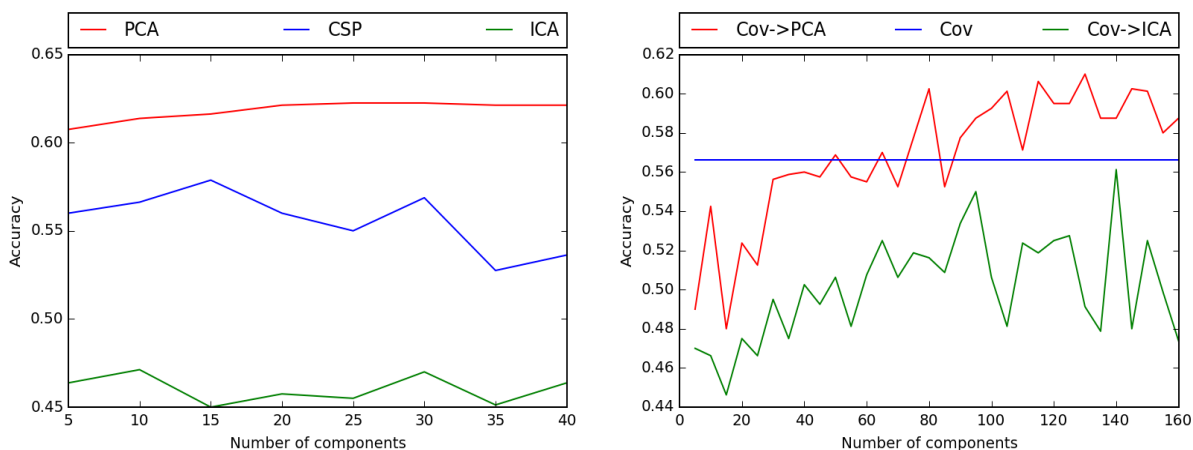


Рис. 4: Оценка параметров моделей для классификации видов движений

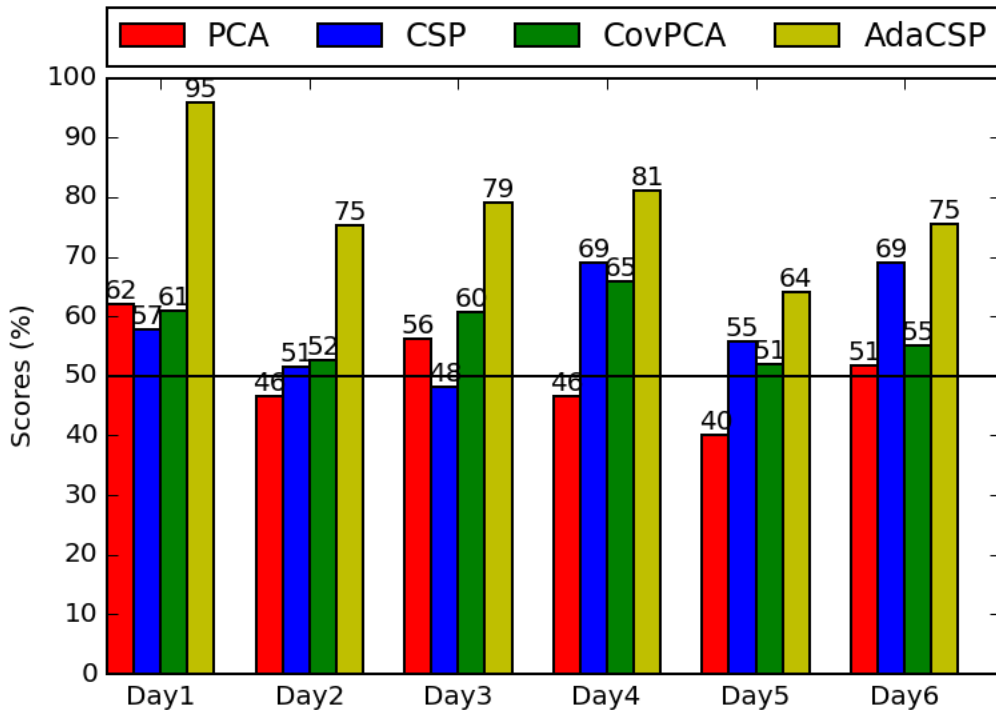


Рис. 5: Точность различных моделей в экспериментах для одного пациента: движения левой и правой рук

В обоих случаях использование метода независимых компонент дало значительно худшие по сравнению с другими методами результаты при любом количестве компонент. Исходя из полученных результатов, для эксперимента с использованием коэффициентов матриц ковариации число компонент было выбрано равным 130, для эксперимента с преобразованием исходных признаков — 15.

4.6 Оценка точности полученных моделей по другим экспериментам

Настроенные по одному из экспериментов модели были затем использованы для классификации движений в других экспериментах пациента. Результаты показаны на рис. 5. Наиболее эффективным оказался метод AdaBoost с использованием метода CSP формирования признаков, предоставляющий в каждом случае наивысшую по сравнению с другими методами точность классификации. При решении задачи определения наличия активности (в данном случае за один

класс берется наличие активности, за другой — отсутствие) наиболее точные результаты показал метод использования коэффициентов матрицы ковариации, как видно на рис. 6.

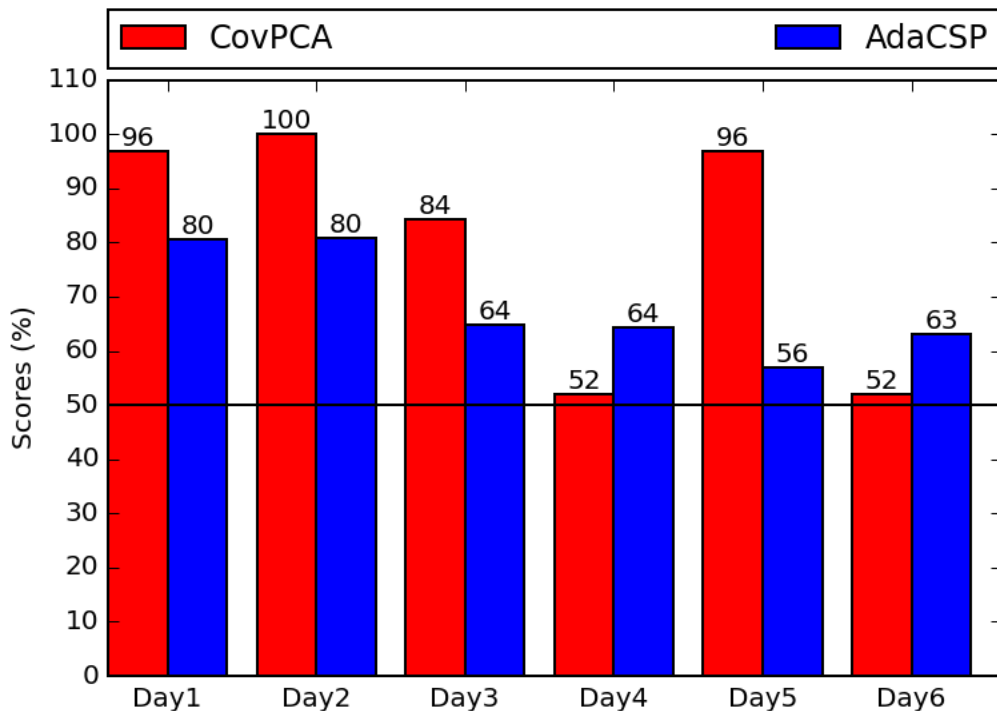


Рис. 6: Точность различных моделей в экспериментах для одного пациента: определение наличия активности

Средние точности оптимальных алгоритмов для классификации различных типов движений и наличия активности составили 78.45% и 80.3% соответственно.

5 Заключение

В ходе работы был проведен обзор существующих методов формирования признаков в задаче классификации движений. Наиболее эффективным методом для классификации различных типов движений в данном эксперименте оказался метод общих пространственных структур с выравниванием данных с помощью вычитания среднего по отрезку и использованием адаптивного бустинга в качестве классификатора. Для определения наличия движения как такового эффективнее всего показал себя метод использования коэффициентов ковариации.

ации с последующей обработкой методом главных компонент. В дальнейшем предлагается использовать комбинацию данных алгоритмов для решения задачи мультиклассовой классификации.

Список литературы

- [1] P. W. Duncan, L. B. Goldstein, D. Matchar, G. W. Divine, and J. Feussner, “Measurement of motor recovery after stroke,” *Stroke*, vol. 23, pp. 1084–1089, 1992.
- [2] K. K. Ang and C. Guan, “Brain-Computer Interface in Stroke Rehabilitation,” *Journal of Computer Science and Engineering*, vol. 7, no. 2, pp. 139–146, 2013.
- [3] S. Baillet, J. C. Mosher, and R. M. Leahy, “Electromagnetic brain mapping,” *IEEE Signal Processing Magazine*, vol. 18, no. November, pp. 14–30, 2001.
- [4] M. Arvaneh, C. Guan, K. K. Ang, and C. Quek, “Optimizing spatial filters by minimizing within-class dissimilarities in electroencephalogram-based brain-computer interface,” *IEEE Transactions on Neural Networks and Learning Systems*, vol. 24, no. 4, pp. 610–619, 2013.
- [5] X. Yu, P. Chum, and K. B. Sim, “Analysis the effect of PCA for feature reduction in non-stationary EEG based motor imagery of BCI system,” *Optik*, vol. 125, no. 3, pp. 1498–1502, 2014.
- [6] J. Law and I. T. Jolliffe, “Principal Component Analysis.,” *The Statistician*, vol. 36, no. 4, p. 432, 1987.
- [7] J. Hill, T. N. Lal, M. Schroeder, T. Hinterberger, G. Widman, C. Elger, B. Schölkopf, and N. Birbaumer, “Classifying Event-Related Desynchronization in EEG, ECoG and MEG signals},” 2006.
- [8] L. Ke and R. Li, “Classification of EEG signals by multi-scale filtering and PCA,” in *Proceedings - 2009 IEEE International Conference on Intelligent Computing and Intelligent Systems, ICIS 2009*, vol. 1, pp. 362–366, 2009.

- [9] a. Hyvärinen, a. Hyvärinen, E. Oja, and E. Oja, “Independent component analysis: algorithms and applications.,” *Neural networks : the official journal of the International Neural Network Society*, vol. 13, no. 4-5, pp. 411–30, 2000.
- [10] N. Xu, X. Gao, B. Hong, X. Miao, S. Gao, and F. Yang, “BCI competition 2003 - Data set IIb: Enhancing P300 wave detection using ICA-based subspace projections for BCI applications,” *IEEE Transactions on Biomedical Engineering*, vol. 51, no. 6, pp. 1067–1072, 2004.
- [11] K. K. Ang, Z. Y. Chin, H. Zhang, and C. Guan, “Filter Bank Common Spatial Pattern (FBCSP) in Brain-Computer Interface,” *2008 IEEE International Joint Conference on Neural Networks (IEEE World Congress on Computational Intelligence)*, pp. 2391–2398, 2008.
- [12] K. K. Ang, K. S. G. Chua, K. S. Phua, C. Wang, Z. Y. Chin, C. W. K. Kuah, W. Low, and C. Guan, “A Randomized Controlled Trial of EEG-Based Motor Imagery Brain-Computer Interface Robotic Rehabilitation for Stroke,” *Clinical EEG and neuroscience*, 2014.
- [13] K. K. Ang, C. Guan, K. S. G. Chua, B. T. Ang, C. W. K. Kuah, C. Wang, K. S. Phua, Z. Y. Chin, and H. Zhang, “A large clinical study on the ability of stroke patients to use an EEG-based motor imagery brain-computer interface,” *Clinical EEG and Neuroscience*, vol. 42, no. 4, pp. 253–8, 2011.
- [14] K. K. Ang, C. Guan, K. S. G. Chua, K. S. Phua, C. Wang, Z. Y. Chin, L. Zhou, K. Y. Tang, G. J. E. Joseph, and C. Kuah, “A clinical study of motor imagery BCI performance in stroke by including calibration data from passive movement,” *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, pp. 6603–6606, 2013.
- [15] H. J. Hwang, S. Kim, S. Choi, and C. H. Im, “EEG-Based Brain-Computer Interfaces: A Thorough Literature Survey,” *International Journal of Human-Computer Interaction*, vol. 29, no. 12, pp. 814–826, 2013.
- [16] K.-R. Müller, M. Krauledat, G. Dornhege, G. Curio, and B. Blankertz, “Machine Learning Techniques for Brain-Computer Interfaces,” *Machine Learning*, vol. 49, pp. 11–22, 2004.

- [17] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, “Scikit-learn: Machine learning in Python,” *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.