

Введение

Электроэнцефалография (ЭЭГ) – медицинский метод визуализации, осуществляющий запись электрического поля с поверхности кожи головы человека. Эти электрические потенциалы генерируются мозговой активностью и фиксируются в виде колебаний напряжения при помощи набора проводящих электродов, расположенных по всей поверхности кожи головы [4-6]. Запись, выполненная с использованием данного метода, называется энцефалограммой. Для науки имеет большое значение распознавание сигналов энцефалограммы головного мозга. Это служит основой для построения нейрокомпьютерных интерфейсов. Исследование, проведенное в данной работе, нацелено на анализ методов регуляризации классификации с целью повышения эффективности выделения полезного сигнала.

В разделе «Анализ предметной области и постановка задачи» приводятся описания электроэнцефалографии, принцип фиксации сигналов мозговой деятельности, а также задачи распознавания полезных сигналов при быстром последовательном предъявлении визуальных раздражителей.

В разделе «Существующие методы решения задачи классификации сигналов энцефалограммы при быстром последовательном предъявлении визуальных раздражителей» приведен один из методов, при помощи которого можно распознать полезный сигнал – линейный дискриминант Фишера, а также результаты классификации сигналов при помощи данного метода.

В разделе «Разработка метода классификации энцефалограмм при быстром последовательном предъявлении визуальных раздражителей на основе метода опорных векторов» представлена байесовская постановка задачи классификации, а также описание процесса разработки процедуры оптимизации регуляризованного критерия на основании метода градиентного спуска.

В разделе «Экспериментальное исследование предложенных подходов» приводятся результаты работы на модельных и экспериментальных данных.

1. Анализ предметной области и постановка задачи

1.1 Энцефалография

ЭЭГ начала применяться с 1929 года [6], широко применялась в медицине для решения клинических и исследовательских задач. Она используется для исследования эпилептической активности и эпилептических припадков, изучения стадий сна, а так же мысленных процессов после травм головы [4,7]. Кроме того, поскольку ЭЭГ обеспечивает визуальное представление мозговых процессов, в то числе и аномалий, связанных с функциями мозга, ЭЭГ широко применяется для изучения и диагностики неврологических и психических расстройств, а так же для мониторинга развития мозга и когнитивных функций[4].

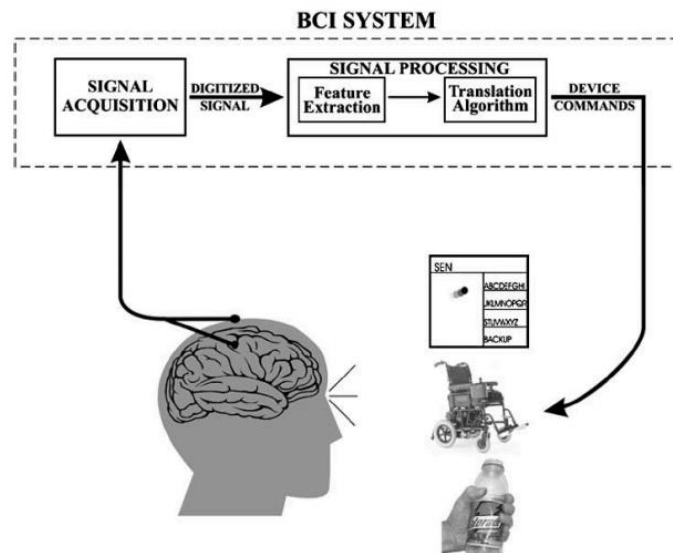


Рисунок 1 – Основные операции и последовательность обработки сигналов в нейрокомпьютерном интерфейсе (Brain-computerinterface - BCI), которые позволяют мысленно управлять внешними устройствами, такими как текстовый процессор, кресло-коляска и нейропротезное устройство[6]

В последние десятилетия, благодаря достижениям в области обработки сигналов и распознаванию образов, а также повышению доступности мощного и недорогого компьютерного оборудования, ЭЭГ послужила основой для создания нейрокомпьютерного интерфейса (НКИ). НКИ

обеспечивает немышечный канал связи между людьми и внешними устройствами[8].

Как показано на рисунке 1, основанный на ЭЭГ НКИ принимает исходные электрофизиологические сигналы (сигналы ЭЭГ), которые отражают намерения пользователя и применяют методы обработки сигналов для их интерпретации в команды, управляющие внешними устройствами, к примеру текстовый процессор или нейропротезное устройство. Данная последовательность обработки сигнала включает в себя алгоритм классификации для категоризации намерений пользователя. Для людей с тяжелыми двигательными расстройствами или инвалидностью НКИ оказывает существенную помощь. Помимо области клинической помощи, НКИ могут принести пользу даже здоровым людям в таких областях, как образование, когнитивное развитие, безопасность, развлечение и т.д.[9]

Сегодня многие НКИ, основанные на ЭЭГ, работают на основе принципа вызывания и обнаружения связанных с событиями потенциалов (ERP) – нейронного ответа, связанных с представлением определенных стимулов или событий[8]. Из многочисленных методов выявления ERP в приложениях НКИ в последние годы парадигма быстрого последовательного предъявления (RSVP) была среди наиболее часто используемых. RSVP включает в себя предъявление субъекту серии визуальных стимулов с очень высокой скоростью, при этом один стимул заменяется другим на небольшой площади экрана [10,11]. Из этих стимулов те, которые вызывают интерес у субъекта (в зависимости от выполняемой задачи) называются целевым стимулом. Когнитивный ответ от субъекта в качестве целевого содержимого приводит к созданию или повышению определенных компонентов ERP в сигнале ЭЭГ, который гораздо менее заметен в реакции на нецелевые объекты. Используя RSVP, этап классификации НКИ обычно сводится к проблеме распознавания двух классов, то есть решает, относится ли стимул к целевому или нецелевому классу, анализируя записанный сигнал ЭЭГ и обнаруживая целевые компоненты ERP. На основании высокой

скорости и эффективности по времени и точности, достигаемые при помощи RSVP, данная парадигма представляет интерес для данного исследования.

1.2 Физиологические основы электроэнцефалографии

При проведении ЭЭГ измеряют суммарные постсинаптические токи. Потенциал действия (ПД, кратковременное изменение потенциала) в пресинаптической мембране аксона (длинного цилиндрического отростка нервной клетки) вызывает высвобождение нейромедиатора в синаптическую щель. Нейромедиатор, или нейротрансмиттер, — химическое вещество, осуществляющее передачу нервных импульсов через синапсы между нейронами. Пройдя через синаптическую щель, нейромедиатор связывается с рецепторами постсинаптической мембраны. Это вызывает ионные токи в постсинаптической мембране. В результате во внеклеточном пространстве возникают компенсаторные токи. Именно эти внеклеточные токи формируют потенциалы ЭЭГ. ЭЭГ нечувствительна к ПД аксонов [2].

Хотя за формирование сигнала ЭЭГ ответственны постсинаптические потенциалы, поверхностная ЭЭГ не способна зафиксировать активность одного дендрита или нейрона. Правильнее сказать, что поверхностная ЭЭГ представляет собой сумму синхронной активности сотен нейронов, имеющих одинаковую ориентацию в пространстве, расположенных радиально к коже головы. Токи, направленные по касательной к коже головы, не регистрируются. Таким образом, во время ЭЭГ регистрируется активность радиально расположенных в коре апикальных дендритов. Поскольку вольтаж поля уменьшается пропорционально расстоянию до его источника в четвертой степени, активность нейронов в глубоких слоях мозга зафиксировать гораздо труднее, нежели токи непосредственно около кожи [2].

Токи, регистрируемые на ЭЭГ, характеризуются различными частотами, пространственным распределением и взаимосвязью с различными состояниями мозга (например, сон или бодрствование). Такие колебания потенциала представляют собой синхронизированную активность целой сети нейронов.

Идентифицированы лишь немногие нейронные сети, ответственные за регистрируемые осцилляции (например, таламокортикальный резонанс, лежащий в основе «сонных веретен» — учащенных альфа-ритмов во время сна), тогда как многие другие (например, система, формирующая затылочный основной ритм) пока не установлены [2].

1.3 Использование электроэнцефалографии в научных целях

Использование ЭЭГ в ходе нейробиологических исследований имеет целый ряд преимуществ перед другими инструментальными методами. Во-первых, ЭЭГ представляет собой неинвазивный способ исследования объекта. Во-вторых, нет такой жесткой необходимости оставаться в неподвижном состоянии, как при проведении функциональной МРТ. В-третьих, в ходе ЭЭГ регистрируется спонтанная активность мозга, поэтому от субъекта не требуется взаимодействия с исследователем (как, например, это требуется в поведенческом тестировании в рамках нейропсихологического исследования). Кроме того, ЭЭГ обладает высоким разрешением во времени по сравнению с такими методами, как функциональная МРТ, и может использоваться для идентификации миллисекундных колебаний электрической активности мозга[5].

В психологии электроэнцефалография широко используется для изучения функционального развития человеческого мозга так же, как и неврологических расстройств, которые влияют на поведение человека. Многие исследования в этой области также изучают характеристики нейронной деятельности при воздействии раздражителей определенного типа, например, зрительных, слуховых или тактильных [10]. Обычно ЭЭГ регистрируются в тот момент, когда субъект осуществляет выполнение задачи, призванной вызвать определенную когнитивную реакцию[10]. Однако, помимо того, ЭЭГ содержит полезный сигнал, оно подвергается воздействию ряда негативных факторов. К таким факторам можно отнести технические (шум от сети или усилителя, помехи от линий электропередач и т.д.) или физиологические (моргание, движения глаз, мышечные сокращения) [9]. Полученное в результате низкое соотношение сигнал/шум означает, что полезный сигнал становится трудно различим. Методы решения данной проблемы обычно включают в себя усреднения значений сигнала при повторных испытаниях по каждому стимулу. Для этого осуществляется предварительная обработка сигнала ЭЭГ. Во-первых, требуется

синхронизировать сигнал ЭЭГ с началом каждого стимула. В цифровых системах это осуществляется путем отправки маркерных кодов в начале каждого стимула, таким образом связывая каждый отдельный сегмент необработанного сигнала ЭЭГ с соответствующим стимулом. На рисунке 2 представлен сигнал ЭЭГ с сегментами, связанными со стимулами при помощи маркерных кодов.

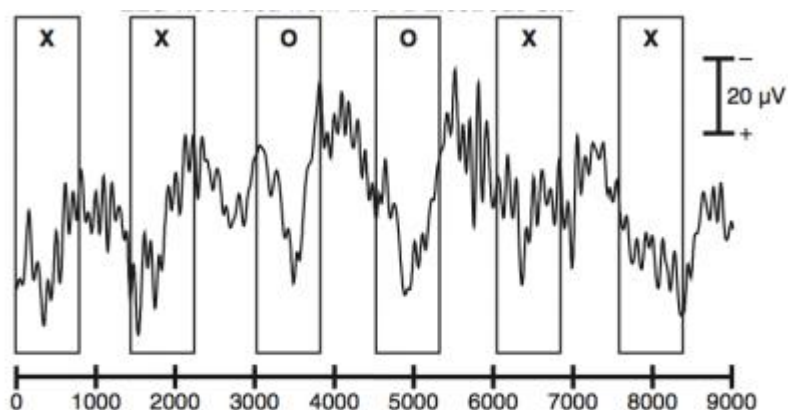


Рисунок 2 - Представлен сигнал ЭЭГ с сегментами, связанными со стимулами при помощи маркерных кодов

Сегменты ЭЭГ с одинаковыми кодовыми обозначениями обозначаются эпохами и затем извлекаются и усредняются. Данный процесс проиллюстрирован на рисунках 3 и 4.

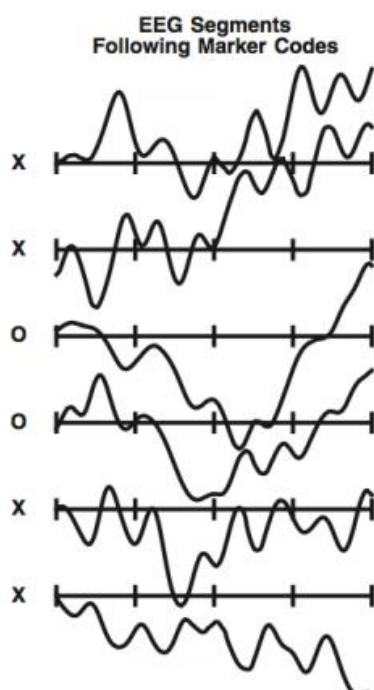


Рисунок 3 – Отдельные сегменты ЭЭГ, соответствующие стимулам X и O

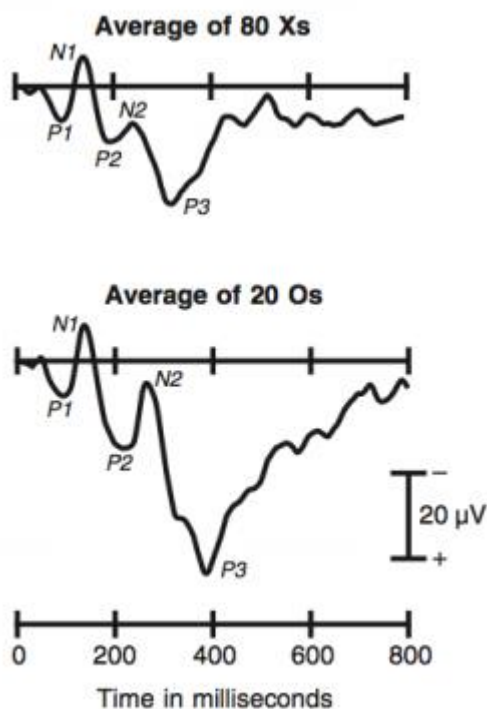


Рисунок 4 – Усредненные формы волн сегментов ЭЭГ

Сигналы, зарегистрированные при наличии когнитивных раздражителей, применяются во многих приложениях, реализующих нейрокомпьютерный интерфейс (НКИ). Одной из основных целей НКИ является попытка точно интерпретировать реакцию пользователя еще до того, как пользователь отреагирует на это воздействие. В НКИ-приложениях реализация такой функции подразумевает анализ и классификацию отдельных сегментов ЭЭГ в реальном времени. При этом необходимо сразу принять решение относительно того, является ли реакция на раздражение целевой. Следовательно, мы не можем избавиться от низкого значения отношения сигнал/шум вследствие отсутствия повторных испытаний и последующего усреднения.

1.4 Потенциал Р300

Потенциал – значение напряжения на электродах. Р300 – позитивный компонент вызванного потенциала. Название состоит из двух частей: Р означает positive, 300 означает момент времени обнаружения данного компонента. То есть Р300 – часть отклика мозга, максимум амплитуды которого возникает примерно через 300 миллисекунд после начала стимула. Она может проявиться при использовании экспериментальной парадигмы (oddball), которая дает возможность сопоставлять реакцию на «значимый — незначимый» стимул. На рисунках 5 и 6 представлена экспериментальная парадигма и форма Р300 в ответ на целевой стимул по сравнению с ответом на нецелевой стимул.

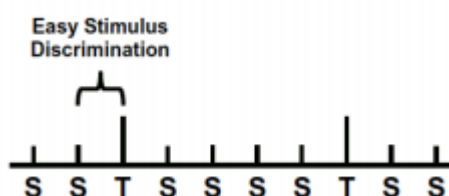


Рисунок 5 – Экспериментальная парадигма, представленная последовательностью нецелевых стимулов (обозначенных «S») и целевых (обозначенных «Т»)

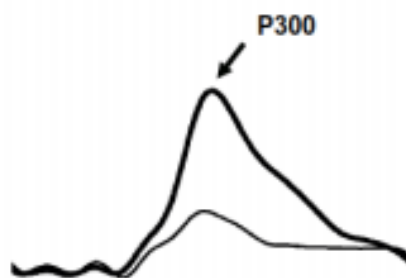


Рисунок 6 – Типичная форма сигнала компонента Р300, соответствующий представлению целевого стимула по сравнению с формой волны, соответствующей нецелевому стимулу

1.5 Методология быстрого последовательного визуального предъявления

Методология RSVP (Rapid serial visual presentation – быстро последовательное визуальное предъявление) реализует принцип интегрального восприятия большого количества визуальных данных, минуя этап осознанного принятия решений. Это обеспечивается передовой вычислительной мощностью визуальной системы человека, которая, как было установлено, способна идентифицировать и классифицировать содержимое изображения как целевое или нецелевое с частотой более 10 Гц [9]. Принимая данный факт во внимание, Пол Сайда (2010) вместе с коллегами предложили два проекта системы «Компьютерного зрения», которая сочетает компьютерное зрение и ЭЭГ с парадигмой RSVP для создания интерфейса для просмотра в режиме реального времени и идентификации пользовательских целевых изображений [10]. Предлагаемые системы могут облегчить общие задачи, связанные с просмотром и визуальным анализом изображений, или могут приносить пользу специалистам в таких областях, как радиолокация, анализ спутниковых изображений и т.д. В другом исследовании, проводимом Джоном Турианом и его коллегами (2011), парадигма RSVP и анализ P300 использовались для определения правильности распознавания лиц в режиме реального времени. Результаты этого исследования показали, что P300 может использоваться не только для идентификации содержимого изображения, соответствующего поставленной задаче, но и для разработки моделей для обучения [9].

Методология RSVP может иметь множество других применений. В недавнем совместном исследовании Департамента психологии и Центра обработки речи, изображений и сигналов (отдел электронной инженерии) в Университете Суррея был изучен НКИ с RSVP-подходом для потенциального использования RSVP в медицинской области в качестве метода скрининга (проведения определенных обследований и тестов, направленное на выявление возможных онкологических заболеваний).

В данной работе информация, полученная в ходе маммографического скрининга, использовалась в качестве стимулов, а испытуемые были специалистами в области маммографии. Набор стимулов состоял из небольшого количества сегментов маммограммы, содержащих изображение молочной железы с злокачественной опухолью (целевые) и большого количества сегментов маммограммы здоровой молочной железы (нецелевые). На рисунке 7 проиллюстрированы целевые и нецелевые стимулы.

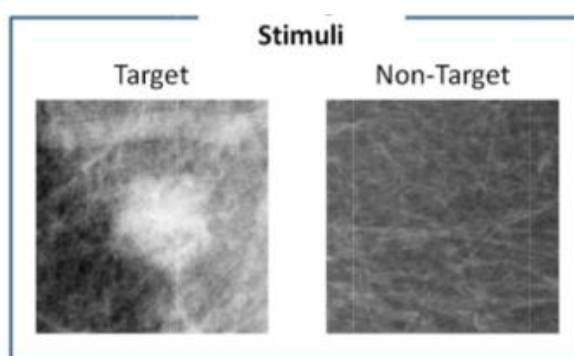


Рисунок 7 – Целевые и нецелевые стимулы

Применяя методологию RSVP, изображения были представлены испытуемым со скоростью 10 изображений в секунду (10 Гц). Для анализа использовался подход НКИ. В экспериментах наблюдалось, что вызванные потенциалы, особенно компоненты P300, соответствующие целевым изображениям, могут быть достоверно выявлены в ходе классификации. Было установлено, что эффективность классификации улучшается при использовании нескольких электродов.

Результаты показали, что данный метод может быть подходящим для использования в медицинской области в качестве метода высокопроизводительного скрининга или в качестве быстрого и экономически эффективного способа подготовки экспертов по скринингу [8].

Для исследования использовались экспериментальные данные ЭЭГ, полученные путем анализа скрининга маммографии на основе методологии RSVP. Проект фокусируется на улучшении обнаружения вызванного

потенциала путем извлечения полезных функций и повышении надежности классификации.

С этой целью была разработана модель, представленная на рисунке 8.

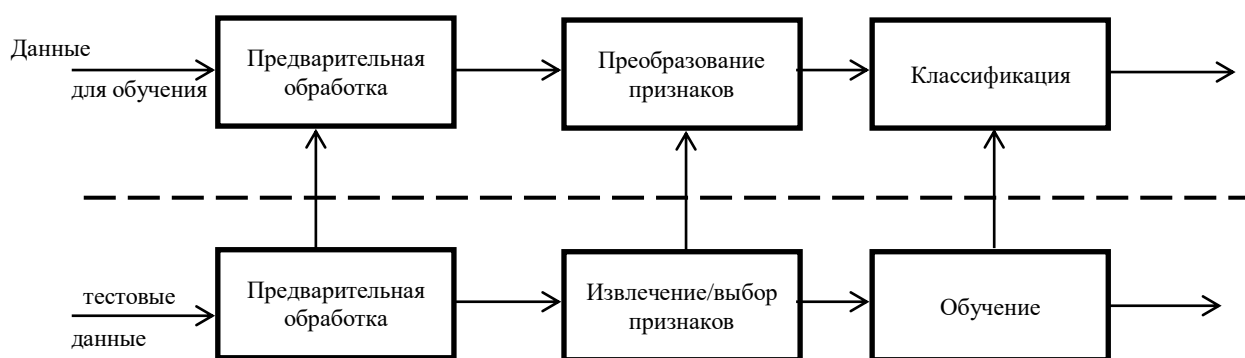


Рисунок 8 – Модель распознавания целевых стимулов

Картинку с русскими буквами

Модель содержит несколько основных этапов и два связанных между собой режима: режим обучения и режим классификации. Данные для обучения (trainingpattern) используются для обучения классификатора, а тестовые данные (testpattern) –используются для проверки корректности работы классификатора. Ниже приводится описание этапов.

- Предварительная обработка (Preprocessing) – этот шаг применяется как к данным для обучения, так и к тестовым данным для представления их в более удобной форме. Например, сегментация сигнала ЭЭГ, удаление шумов. Обычно это осуществляется на оборудовании для записи ЭЭГ, а также путем цифровой обработки данных.

- Извлечение признаков (FeatureExtraction) – данный этап представляет собой уменьшение размерности признакового пространства. Поскольку ЭЭГ представляет собой запись показателей величины напряжения во времени, каждый сегмент ЭЭГ содержит большое количество замеров. Многие сегменты ЭЭГ содержат избыточную информацию, которая не является существенной при классификации. Следовательно, необходимо выделить функцию отбора только значимых характеристик сигнала.

Используя линейное и нелинейное отображение исходных выборок сигнала, каждый сегмент ЭЭГ может быть описан с использованием гораздо меньшего числа переменных [7].

- Отбор признаков (FeatureSelection) – при использовании различных методов извлечения признаков, их количество все еще может быть довольно высоким, и не все они могут быть полезны при классификации. Следовательно, выполняется отбор признаков, тем самым еще больше уменьшая размерность признакового пространства и повышая эффективность классификации.

Извлечение признаков осуществляется при помощи:

1) Временных функций.

Максимальная амплитуда:

$$s_{max} = \max(s[n]),$$

где $s[n]$ - сегмент сигнала энцефалограммы, содержащий n значений амплитуды сигнала;

Сумма положительных значений амплитуды сигнала:

$$A_p = \sum_{n=1}^N 0.5(s[n] + |s[n]|)$$

Среднеквадратический уровень сигнала:

$$V_{RMS} = \sqrt{\frac{1}{N} \sum_{n=1}^N (s[n])^2}$$

Разница между максимальной и минимальной величинами сигнала:

$$s_{pp} = s_{max} - s_{min}$$

где $s_{min} = \min(s[n])$

Индекс максимального значения амплитуды сигнала:

$$n_{s_{max}} = \{n | s[n] = s_{max}\}$$

Индекс минимального значения амплитуды сигнала:

$$n_{s_{min}} = \{n | s[n] = s_{min}\}$$

$$SSA = \sum_{n=2}^{N-1} 0.5 \left| \frac{s[n-1] - s[n]}{|s[n-1] - s[n]|} + \frac{s[n+1] - s[n]}{|s[n+1] - s[n]|} \right|$$

2) Понижение частоты дискретизации – выявления признаков потенциала Р300, используя сегменты энцефалограммы, к которым было применено понижение частоты дискретизации.

3) Взаимнокорреляционная функция – показывает, насколько совпадает шаблон, описывающий целевую характеристику потенциала Р300 с входным сигналом.

4) Характеристики амплитудного спектра – показывают разность амплитуд между целевыми и нецелевыми сигналами.

- Обучение (Learning) – в режиме обучения отобранные признаки из целевых и нецелевых образцов ЭЭГ подаются в качестве входных данных классификатору, который в свою очередь анализирует распределение данных и вычисляет оптимальную границу решения для разделения двух классов.

Обучив классификатор, система может определять целевые и нецелевые образцы ЭЭГ, которые не входили в обучающую выборку. Формальная постановка задачи классификации представлена в разделе 1.5

1.6 Постановка задачи классификации

Пусть \mathbf{X} – множество описаний объектов, \mathbf{Y} – конечное множество меток классов. Существует некая зависимость – отображение $y^*: \mathbf{X} \rightarrow \mathbf{Y}$. Также в нашем распоряжении имеется обучающая выборка:

$$\mathbf{X}^N = \{(\mathbf{x}_1, y_1), \dots, (\mathbf{x}_N, y_N)\}$$

Требуется построить алгоритм $a: \mathbf{X} \rightarrow \mathbf{Y}$, способный классифицировать произвольный объект $\mathbf{x} \in \mathbf{X}$

2. Существующие методы решения задачи классификации сигналов энцефалограммы при быстром последовательном визуальном предъявлении

Одним из алгоритмов классификации, применявшихся ранее, является линейный дискриминант Фишера. Данный метод основан на линейном дискриминантном анализе Фишера – контролируемом алгоритме обучения. После того, как каждый объект обучающей выборки был представлен в виде точки в многомерном пространстве, путем линейного дискриминантного анализа осуществляется поиск направления в пространстве признаков такого, что при проецировании всех точек на это направление, расстояние между двумя классами становится максимальным [12]. Вычисление такого направления основано на критерии Фишера, который определяет оптимальное направление для проецирования данных, которое в свою очередь минимизирует дисперсию прогнозируемых точек, принадлежащих каждому классу [12].

При некоторых предположениях байесовский классификатор сводится к формуле

$$g(\mathbf{x}) = \operatorname{argmax}_{y \in Y} \lambda_y P_y p_y(\mathbf{x}) \quad (3)$$

где Y – множество классов, \mathbf{x} принадлежит множеству объектов X , P_y – априорная вероятность класса y , $p_y(\mathbf{x})$ – функция правдоподобия класса y , λ_y – весовой коэффициент (цена ошибки на объекте класса y)

Если принять гипотезу о том, что классы распределены по нормальному закону, то формула (3) принимает вид:

$$g(\mathbf{x}) = \operatorname{argmax}_{y \in Y} \left(\ln(\lambda_y P_y) - \frac{1}{2} (\mathbf{x} - \mu_y)^T \Sigma_y^{-1} (\mathbf{x} - \mu_y) - \frac{1}{2} \ln(|\Sigma_y^{-1}|) - \frac{n}{2} \ln(2\pi) \right)$$

μ_y – приближения вектора математического ожидания и матрицы ковариации класса y .

$$\mu_y = \frac{1}{l_y} \sum_{i=1}^l x_i \cdot \Sigma_y = \frac{1}{l_y} \sum_{i=1}^l (x_i - \mu_y)(x_i - \mu_y)^T \quad (4)$$

где l - длина обучающей выборки, l_y - количество объектов класса y в обучающей выборке.

Данный алгоритм классификации является квадратичным дискриминантом. Он имеет ряд недостатков, одним из которых является плохая обусловленность (вырожденность) матрицы ковариаций Σ_y при малом количестве обучающих элементов класса y , вследствие чего при обращении данной матрицы Σ_y^{-1} может получиться достаточно искаженный результат, и алгоритм классификации окажется неустойчивым. Линейный дискриминатор Фишера решает данную проблему.

Основная идея данного алгоритма заключается в следующем. Если принять гипотезу о том, что ковариационные матрицы равны между собой, то алгоритм классификации примет следующий вид [3]:

$$g(\mathbf{x}) = \operatorname{argmax}_{y \in Y} \left(\ln \lambda_y P_y - \frac{1}{2} \mu_y^T \Sigma^{-1} \mu_y + \mathbf{x}^T \Sigma^{-1} \mu_y \right)$$

Формулу (3) можно переписать следующим образом:

$$g(\mathbf{x}) = \operatorname{argmax}_{y \in Y} (w_0 + \mathbf{w}^T \mathbf{x}) \quad (5)$$

Простота данного метода классификации является одним из его достоинств, так как в случае бинарной классификации в двумерном пространстве признаков разделяющей поверхностью будет прямая. Если классов больше двух, то разделяющая поверхность будет кусочно-линейной. Но одним из главных преимуществ алгоритма является уменьшение эффекта плохой обусловленности ковариационной матрицы при недостаточных данных [3].

При малых l_y приближения (4) дадут плохой результат, поэтому в задачах, в которых заранее известно, что классы имеют различные формы, иногда бывает выгодно воспользоваться эвристикой дискриминанта Фишера и считать матрицы ковариаций всех классов одинаковыми. Это позволит вычислить «среднюю» матрицу ковариаций, используя всю выборку:

$$\Sigma = \frac{1}{l} \sum_{i=1}^l (x_i - \mu_{y_i})(x_i - \mu_{y_i})^T$$

Используя данную матрицу ковариаций, алгоритм классификации станет более устойчивым [3].

Рассмотрим случай, когда конкретный объект из обучающей выборки описывается N-мерным вектором признаков $\mathbf{x} = [x_1, x_2, \dots, x_N]$, или точкой в N-мерном пространстве признаков. Если оптимальное направление проекции, вычисляемое посредством линейного дискриминантного анализа, описывается вектором $\mathbf{w} = [w_1, w_2, \dots, w_N]$, то проекция точки \mathbf{x} на направление, определяемое вектором \mathbf{w} , определяется следующим образом:

$$w_1x_1 + w_2x_2 + w_3x_3 + \dots + w_Nx_N = \mathbf{w}^T\mathbf{x}(1)$$

Проекция может быть выполнена с использованием линейной комбинации компонент \mathbf{x} , и \mathbf{w} можно рассматривать в качестве вектора весов. Как только проекция будет применена ко всем точкам в обучающей выборке, все проецируемые точки будут лежать в одном направлении. Это означает, что исходный многомерный набор значений будет сведен к новому одномерному набору в направлении, определяемом \mathbf{w} , что значительно упрощает задачу классификации.

На рисунке 9 представлен прогноз, определяемый линейным дискриминатором Фишера в двумерном пространстве.

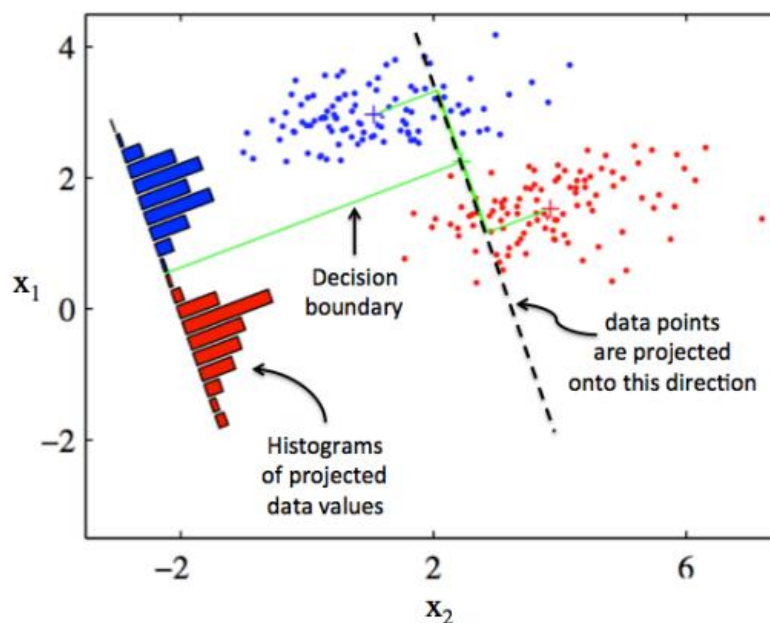


Рисунок 9 - Прогноз, определяемый линейным дискриминатором Фишера в двумерном пространстве

Рисунок 9 иллюстрирует линейный дискриминантный анализ Фишера в двумерном пространстве признаков, где x_1 и x_2 являются осями, которые соответствуют двум различным признакам. Данные разных классов обозначены разными цветами. Пунктирная линия показывает направление проекции, определяемое линейным дискриминантным анализом Фишера, и гистограммы точек, проецируемых на это направление. Видно, что проекция приводит к очень хорошему разделению классов при очень небольшом перекрытии гистограмм. Разделив точки посредством проецирования, в области пространства объектов можно определить границу решения (зеленая линия), перпендикулярной направлению проецирования.

Как только направление максимального разделения двух классов было найдено, линейный дискриминатор Фишера может быть реализован функцией:

$$g(\mathbf{x}) = \mathbf{w}^T \mathbf{x} + w_0(2)$$

где $g(\mathbf{x})$ используется для классификации точки \mathbf{x} , то есть отнесению ее к одному из двух классов. Уравнение $g(\mathbf{x}) = 0$ представляет собой границу решения, перпендикулярную вектору \mathbf{w} [49,50]. Это гиперплоскость, которая разделяет N -мерное пространство на два подпространства для разделения двух классов, имеет размерность $N - 1$. w_0 – смещение, используемое для перемещения гиперплоскости для корректировки порога с целью улучшения точности классификации [12].

Недостаток данного алгоритма - неустойчивость и низкое качество классификации при малом количестве обучающих объектов [3].

Результаты классификации при помощи линейного дискриминатора Фишера в проведенной работе приведены в таблице 1.

Таблица 1 – Результаты классификации при помощи линейного дискриминатора Фишера

Количество электродов	Количество признаков	Точность классификации
6	198	0.75

3. Разработка метода классификации энцефалограмм при быстром последовательном предъявлении визуальных раздражителей на основе метода опорных векторов

3.1 Байесовская постановка задачи классификации с учетом априорной информации о характере сигналов

В настоящее время данный метод считается одним из лучших методов классификации. Рассмотрим задачу классификации на два непересекающихся класса, в которой объекты описываются n -мерными вещественными векторами

$\mathbf{x}(\omega) = \{x_1, \dots, x_n\}^T \in R^n$ – вектор признаков, соответствующий объекту $\omega \in \Omega$, где Ω – некоторое множество объектов. Каждому объекту соответствует метка класса $y \in \{-1, 1\}$. В нашем распоряжении имеется обучающая выборка [4].

$$(X, Y) = \{(x_i, y_i), i = 1, \dots, n\}$$

$$a(\mathbf{x}) = \text{sign}(\mathbf{a}^T \mathbf{x} - b)$$

Пусть в пространстве признаков R^n объективно представлена разделяющая гиперплоскость $\mathbf{a}^T \mathbf{x} + b = 0$. В качестве модели генеральной совокупности будем рассматривать два параметрических семейства плотностей распределения вероятностей $\phi_1(\mathbf{x}|\mathbf{a}, b)$ и $\phi_{-1}(\mathbf{x}|\mathbf{a}, b)$, связанных с двумя классами объектов $y = 1$ и $y = -1$, и сконцентрированные преимущественно по разные стороны гиперплоскости. Совместную плотность распределения конечного множества векторов признаков объектов известных классов в составе обучающей совокупности будем понимать как плотность распределения выборки независимых реализаций этих двух распределений [4].

$$\Phi(x_j|y_j, \mathbf{a}, b, j = 1, \dots, N) = \prod_{j=1}^N \phi_{y_j}(x_j|y_j, \mathbf{a}, b) =$$

$$\left(\prod_{j:y_j=1}^N \phi_1(x_j|y_j, \mathbf{a}, b) \right) \left(\prod_{j:y_j=-1}^N \phi_{-1}(x_j|y_j, \mathbf{a}, b) \right)$$

Пусть, далее, выбрана априорная плотность совместного распределения вероятностей $\psi(\mathbf{a}, b)$ для параметров распределений $\phi_1(\mathbf{x}|\mathbf{a}, b)$ и $\phi_{-1}(\mathbf{x}|\mathbf{a}, b)$. Тогда апостериорная плотность совместного распределения параметров \mathbf{a} и b относительно обучающей совокупности определяется по формуле Байеса[4]:

$$p(\mathbf{a}, b|y_j, j = 1, \dots, N) = \frac{\psi(\mathbf{a}, b) \Phi(x_j|y_j, \mathbf{a}, b, j = 1, \dots, N)}{\int \psi(\mathbf{a}', b') \Phi(x_j|y_j, \mathbf{a}', b', j = 1, \dots, N) d\mathbf{a}' db'}$$

Поскольку знаменатель не зависит от целевых переменных, то достаточно рассматривать только числитель:

$$\begin{aligned} p(\mathbf{a}, b|y_j, j = 1, \dots, N) &\propto \psi(\mathbf{a}, b) \Phi(x_j|y_j, \mathbf{a}, b, j = 1, \dots, N) \\ &= \psi(\mathbf{a}, b) \left(\prod_{j:y_j=1}^N \phi_1(x_j|y_j, \mathbf{a}, b) \right) \left(\prod_{j:y_j=-1}^N \phi_{-1}(x_j|y_j, \mathbf{a}, b) \right) \end{aligned}$$

Принцип максимизации плотности апостериорного распределения в пространстве параметров модели генеральной совокупности приводит к байесовскому правилу обучения [4]:

$$\begin{aligned} (\hat{\mathbf{a}}, \hat{b} | (\mathbf{x}_j, y_j), j = 1, \dots, N) &= \\ &= \arg \max_{\mathbf{a}, b} p(\mathbf{a}, b | y_j, j = 1, \dots, N) = \\ &= \arg \max_{\mathbf{a}, b} \left[\log \psi(\mathbf{a}, b) + \sum_{j:y_j=1} \phi_1(x_j|y_j, \mathbf{a}, b) + \sum_{j:y_j=-1} \phi_{-1}(x_j|y_j, \mathbf{a}, b) \right] \end{aligned}$$

В работе [1] предложена следующая вероятностная модель генеральной совокупности. Рассмотрим несобственные плотности распределения $\phi_1(\mathbf{x}|\mathbf{a}, b)$ и $\phi_{-1}(\mathbf{x}|\mathbf{a}, b)$, определяемые выражениями [4]

$$\phi_1(\mathbf{x}|\mathbf{a}, b) = \begin{cases} 1, & \mathbf{a}^T \mathbf{x} + b \geq 1 \\ \exp[-c(1 - (\mathbf{a}^T \mathbf{x} + b))], & \mathbf{a}^T \mathbf{x} + b < 1 \end{cases}$$

$$\phi_{-1}(\mathbf{x}|\mathbf{a}, b) = \begin{cases} 1, & \mathbf{a}^T \mathbf{x} + b \leq -1 \\ \exp[-c(1 - (\mathbf{a}^T \mathbf{x} + b))], & \mathbf{a}^T \mathbf{x} + b > -1 \end{cases}$$

Наглядное представление этой пары несобственных плотностей распределения приведено на рисунке 11.

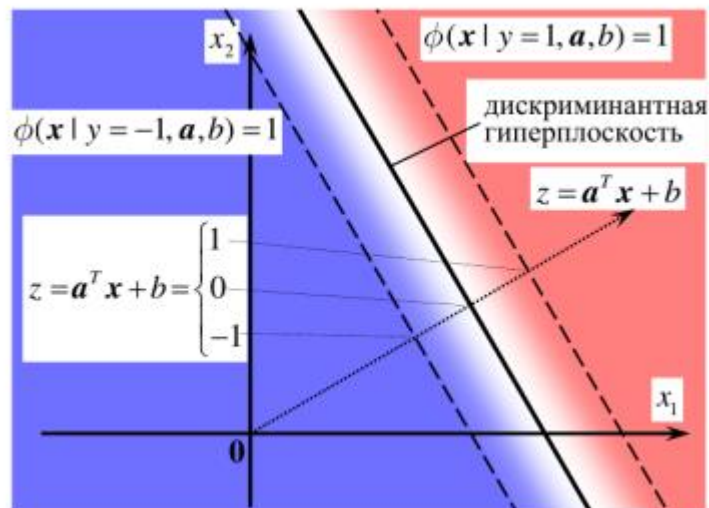


Рисунок 11 – Графическое представление параметрического семейства плотностей распределения двух классов объектов в двумерном пространстве

Направляющий вектор разделяющей гиперплоскости \mathbf{a} будем считать априори нормально распределенным с независимыми компонентами, имеющими одинаковые априорные дисперсии σ^2 . Что же касается параметра положения разделяющей гиперплоскости b , то будем считать, что отсутствуют какие-либо априорные предположения о его значении, что выражается равномерным несобственным распределением, равным единице на всей числовой оси. Тогда совместное априорное распределение параметров разделяющей

гиперплоскости

будет выражаться плотностью[5]

$$\psi(\mathbf{a}, b) \propto \exp\left(-\frac{1}{2\sigma^2} \mathbf{a}^T \mathbf{B} \mathbf{a}\right)$$

Очевидно, что это несобственная плотность, поскольку ее интеграл по переменной b не существует. Для несобственных плотностей распределения векторов признаков объектов двух классов и априорного несобственного распределения параметров разделяющей гиперплоскости байесовский критерий обучения примет вид

$$(\hat{\mathbf{a}}, \hat{b} | (\mathbf{x}_j, y_j), j = 1, \dots, N) = \arg \max_{\mathbf{a}, b} p(\mathbf{a}, b | y_j, j = 1, \dots, N) =$$

$$= \arg \max_{\mathbf{a}, b} \left[-\frac{1}{2\sigma^2} \mathbf{a}^T \mathbf{B} \mathbf{a} + c \sum_{\substack{j: y_j=1 \\ \mathbf{a}^T \mathbf{x} + b < 1}} (1 - (\mathbf{a}^T \mathbf{x} + b)) \right. \\ \left. + c \sum_{\substack{j: y_j=-1 \\ \mathbf{a}^T \mathbf{x} + b > -1}} (1 + (\mathbf{a}^T \mathbf{x} + b)) \right]$$

При выборе коэффициента штрафа на сумму смещений точек обучающей совокупности $C = 2\sigma^2 c$ критерий полностью идентичен классическому критерию обучения методом опорных векторов.

Предположим, имеются объекты в двумерном пространстве действительных значений. На рисунке 12 представлены объекты в двумерном пространстве признаков.

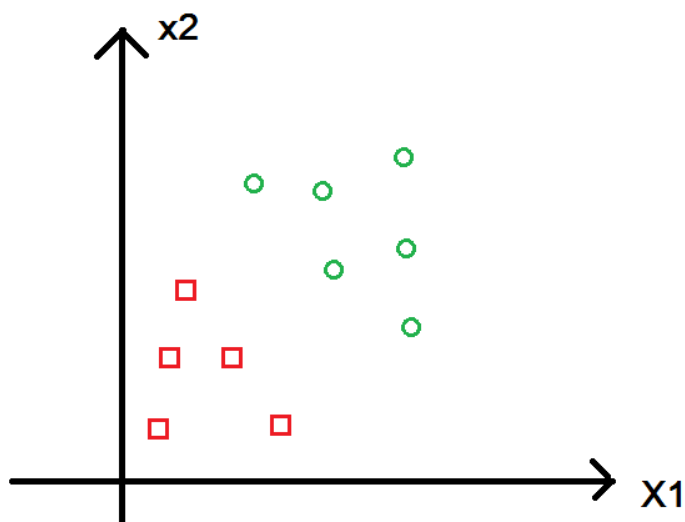


Рисунок 12 – Объекты в двумерном пространстве признаков

На рисунке 12 можно увидеть пример обучающей выборки для случая двух признаков (двух классов).

Будем искать решающее правило в виде гиперплоскости, разделяющей объекты двух классов. Обозначим \mathbf{a} – вектор нормали к разделяющей гиперплоскости, b – кратчайшее расстояние (перпендикуляр) от начала координат до разделяющей гиперплоскости. Тогда разделяющая гиперплоскость делит пространство на два класса. Если: $\mathbf{a}^T \mathbf{x} - b \leq 0$, то $y(\mathbf{x}) = -1$. Если $\mathbf{a}^T \mathbf{x} - b \geq 0$, то $y(\mathbf{x}) = +1$. Данное условие можно переписать следующим образом:

$$y(\mathbf{a}^T \mathbf{x} - b) - 1 \geq 0$$

Объекты, лежащие на границах разделяющей полосы, называются опорными векторами.

Следовательно, задачу определения решающего правила можно свести к задаче минимизации квадратичной функции потерь:

$$J(\mathbf{a}, b) = \mathbf{a}^T \mathbf{a} + C \sum_{i=y_i(\mathbf{a}^T \mathbf{x}_i + b) < 1}^N (1 - y_i(\mathbf{a}^T \mathbf{x}_i + b))^2$$

3.2. Разработка процедуры оптимизации регуляризованного критерия на основании метода градиентного спуска

3.2.1 Метод градиентного спуска

Метод градиентного спуска – метод нахождения локального экстремума функции (минимума или максимума) с помощью движения вдоль градиента. Основная идея метода заключается в том, чтобы идти в направлении наискорейшего спуска, а это направление задается антиградиентом $-\nabla F$. На рисунке 11 представлена иллюстрация последовательных приближений к точке экстремума.

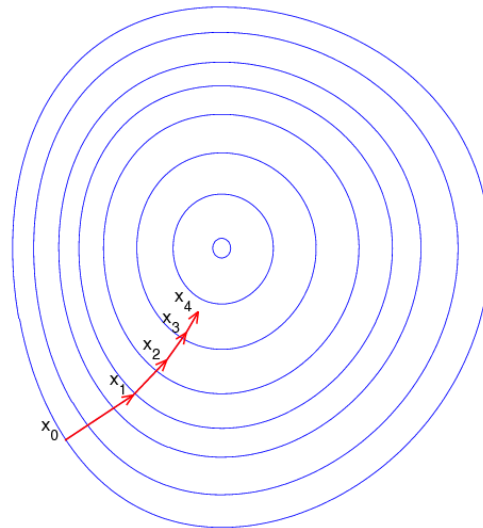


Рисунок 13 – Иллюстрация последовательных приближений к точке экстремума

$$\vec{x}^{[k+1]} = \vec{x}^{[k]} - \lambda^{[k]} \nabla F(\vec{x}^{[k]}) \quad (2)$$

где $\lambda^{[k]}$ выбирается

- Постоянной
- Дробным шагом, то есть длина шага в процессе спуска делится на некое число (например, 2)
- Наискорейшим спуском (привести формулу)

Данный метод применим для поиска минимума квадратичной функции потерь (1).

Следовательно, выражение (2) можно представить в виде:

$$\mathbf{a}_{k+1} = \mathbf{a}_k - \alpha_k \nabla J(\mathbf{a}_k) = \mathbf{a}_k - \alpha_k (\mathbf{a}_k + C * \sum_{i=y_i(\mathbf{a}_k^T \mathbf{x}_i + b) < 1}^N (1 - y_i(\mathbf{a}_k^T \mathbf{x}_i + b)) * (-y_i \mathbf{x}_i)) \quad (3)$$

$$b_{k+1} = b_k - \alpha_k \nabla J(b_k) = b_k - \alpha_k * C * \sum_{i=y_i(\mathbf{a}_k^T \mathbf{x}_i + b_k) < 1}^N (1 - y_i(\mathbf{a}_k^T \mathbf{x}_i + b_k)) * -y_i \quad (4)$$

Метод заключается в следующем:

- 1) Вычисляется разница между значениями функции на (k+1)-ом и k-ом шаге
- 2) Проверяется условие останова: $J(\mathbf{a}_{k+1}, b_{k+1}) - J(\mathbf{a}_k, b_k) < \epsilon$, где ϵ – точность вычисления

Параметры α_k и b_k вычисляются следующим образом:

$$\alpha_{k+1} = \frac{\alpha_k}{2}$$

$$b_{k+1} = \frac{b_k}{2}$$

Остановка вычислений осуществляется при условии если $J(\mathbf{a}_{k+1}, b_{k+1}) - J(\mathbf{a}_k, b_k) > 0$

3.2.2. Оптимизация регуляризованного критерия

Предположим, что отдельные признаки $(x_i, i = 1, \dots, n)$ в составе вектора \mathbf{x} представляют собой результат упорядоченного измерения некоторого свойства объекта вдоль координаты той или иной природы, причем есть основания полагать, что соседние признаки несут почти идентичную информацию о принадлежности объекта к определенному классу [3].

Такое предположение эквивалентно принятию тезиса о существовании априорной информации о значениях коэффициентов $(a_i, i = 1, \dots, n)$ в составе вектора параметров \mathbf{a} , заключающееся в том, что соседние коэффициенты, скорее всего, не слишком сильно отличаются друг от друга, т.е. плавно изменяются при увеличении индекса i .

Для того, чтобы в процессе обучения предпочтение отдавалось решающим правилам с плавным изменением коэффициентов линейной части, можно, например, внести в критерий дополнительную аддитивную составляющую [3]

$$J'(\mathbf{a}) = \sum_{i=2}^n (a_i - a_{i-1})^2.$$

Нетрудно убедиться, что такая квадратичная функция может быть записана в виде [4]

$$J'(\mathbf{a}) = \mathbf{a}^T \tilde{\mathbf{B}} \mathbf{a},$$
$$\tilde{\mathbf{B}}(n \times n) = \begin{pmatrix} 1 & -1 & 0 & \dots & 0 & 0 & 0 \\ -1 & 2 & -1 & \dots & 0 & 0 & 0 \\ 0 & -1 & 2 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 2 & -1 & 0 \\ 0 & 0 & 0 & \dots & -1 & 2 & -1 \\ 0 & 0 & 0 & \dots & 0 & -1 & 1 \end{pmatrix}.$$

Тогда выражения 1,3 можно переписать следующим образом

$$J(\mathbf{a}, b, C) = \mathbf{a}^T \tilde{\mathbf{B}} \mathbf{a} + C \sum_{i=y_i(\mathbf{a}^T \mathbf{x}_i + b) < 1}^N (1 - y_i(\mathbf{a}^T \mathbf{x}_i + b))^2$$

$$\mathbf{a}_{k+1} = \mathbf{a}_k - \alpha_k \nabla J(\mathbf{a}_k) = \mathbf{a}_k - \alpha_k (\tilde{\mathbf{B}} \mathbf{a}_k + C * \sum_{i=y_i(\mathbf{a}_k^T \mathbf{x}_i + b) < 1}^N (1 - y_i(\mathbf{a}_k^T \mathbf{x}_i + b)))$$

Эксперименты

Результаты по некоторым каналам:

