

# Классификация видов физической активности человека по показаниям акселерометра и гироскопа

*О. А. Харациди*

oleg.kharatsidi@gmail.com

Московский государственный университет им. М. В. Ломоносова

Рассматривается задача распознавания видов физической активности человека по показаниям акселерометра и гироскопа портативного устройства на примере открытого набора данных USC-HAD с 12 классами. Предлагается метод, использующий иерархию классов и настраивающий отдельные классификаторы в ее узлах. Ключевую роль играют классификаторы, использующие частотные признаки и представляющие собой смесь из трех принципиально различных моделей: логистической регрессии, метода ближайшего соседа и случайного леса. Итоговое качество классификации соответствует среднему значению F-меры 0,92.

**Ключевые слова:** распознавание физической активности; сенсоры; обработка сигналов

## Human activity recognition based on accelerometer and gyro data

*O. A. Kharatsidi*

M. V. Lomonosov Moscow State University

In the last few years the performance of smartphones has been growing rapidly. They have become capable of carrying out relatively complex computations in real time. In the paper, a problem of human physical activity recognition is considered based on the data from sensors on wearable devices. The classical approach is to split the accelerometer signals into fixed-width windows, classify them independently, and then combine the classification responses into a single one for the whole sample. Classification models may vary from Naïve Bayes classifiers to Neural Nets. The two common types of the features are the statistical metrics (moments, correlations, etc.) and Fourier coefficients.

The method introduced in this paper utilizes the same approach but unlike the most common case, it uses a mixture of three standard classification models: Logistic Regression, Nearest Neighbour, and Random Forest built up on the Fourier coefficients absolute values. Feature selection based on the trained Logistic Regression coefficients is applied to fit the rest two models independently. The method was tested on the USC-HAD open dataset containing measurements of 12 classes from 14 people. Apart from the widely used accelerometer data, it also provides gyro signals which are used in just the same way. The method also exploits a hierarchy of the classes and trains multiple individual classifiers in its nodes.

Since the data consists of the measurements for multiple people, in their experiments, the authors run cross-validation with a single fold per each person. In each iteration, an internal cross-validation was also run to fit hyperparameter. As a result, the algorithm achieves the performance of 0.92 in terms of the mean F-measure. The experiments also show that the mixture of the three models is more stable than each of its components and achieves higher performance. Finally, the method proves to be significantly better than standard  $L_2$ -regularized Logistic Regression built up on the same feature set.

**Keywords:** activity recognition; sensors; signal processing; context recognition

## Введение

За последние несколько лет возможности смартфонов значительно выросли, а их использование стало повсеместным. Появились целые рынки приложений для мобильных операционных систем. На этом фоне возросла потребность в решении целого класса задач (context recognition [1]) по определению местоположения, активности или состояния пользователя, а также анализу его физической деятельности с помощью датчиков устройства: акселерометра, гироскопа, Bluetooth-адаптера, микрофона, датчиков освещения, давления и других. Росту интереса к таким задачам также способствует появление носимой электроники (wearable electronics) — нового класса устройств, таких как, например, «умные часы».

Один из простых примеров конкретных задач, решаемых с помощью показаний датчиков, — подсчет количества шагов пользователя.

В данной работе рассматривается другая известная задача — распознавание видов физической активности пользователя (activity recognition [2, 3, 4, 5, 6, 7]). В качестве исходных данных берутся показания акселерометра и гироскопа (датчика угловой скорости).

## Описание данных

В данной работе рассматривается задача классификации для набора данных USC-HAD [8], находящегося в открытом доступе.

Показания акселерометра и гироскопа (датчика угловой скорости) снимались с прибора, крепящегося в районе пояса в определенном положении, с частотой 100 Гц. Всего участвовало 14 чел., каждый из которых сделал по 5 подходов на каждый из 12 видов физической активности. Каждый подход длился примерно от 10 до 30 с.

Показания снимались в моменты, когда человек:

- (1) идет вперед;
- (2) идет по кругу, поворачивая влево;
- (3) идет по кругу, поворачивая вправо;
- (4) поднимается по лестнице;
- (5) спускается по лестнице;
- (6) бежит вперед;
- (7) прыгает на месте;
- (8) сидит с небольшими движениями;
- (9) стоит на месте;
- (10) лежит;
- (11) поднимается на лифте;
- (12) спускается на лифте.

Каждому из приведенных видов активности соответствует класс. В качестве объекта рассматриваются показания 3-х осей акселерометра и 3-х осей гироскопа (рис. 1).

Таким образом, в выборке 12 классов по  $5 \cdot 14 = 70$  объектов в каждом.

Стоит отметить, что с этим набором данных работать несколько проще, чем с показаниями датчиков современных телефонов: качество и характеристики датчиков телефонов сильно разнятся, и, самое главное, телефоны не носят в каком-то определенном положении.

С другой стороны, на этом наборе данных легко переобучиться, поэтому в данной работе по возможности избегается тонкая настройка гиперпараметров.

Более подробное описание данных и способа их получения можно найти в [8].

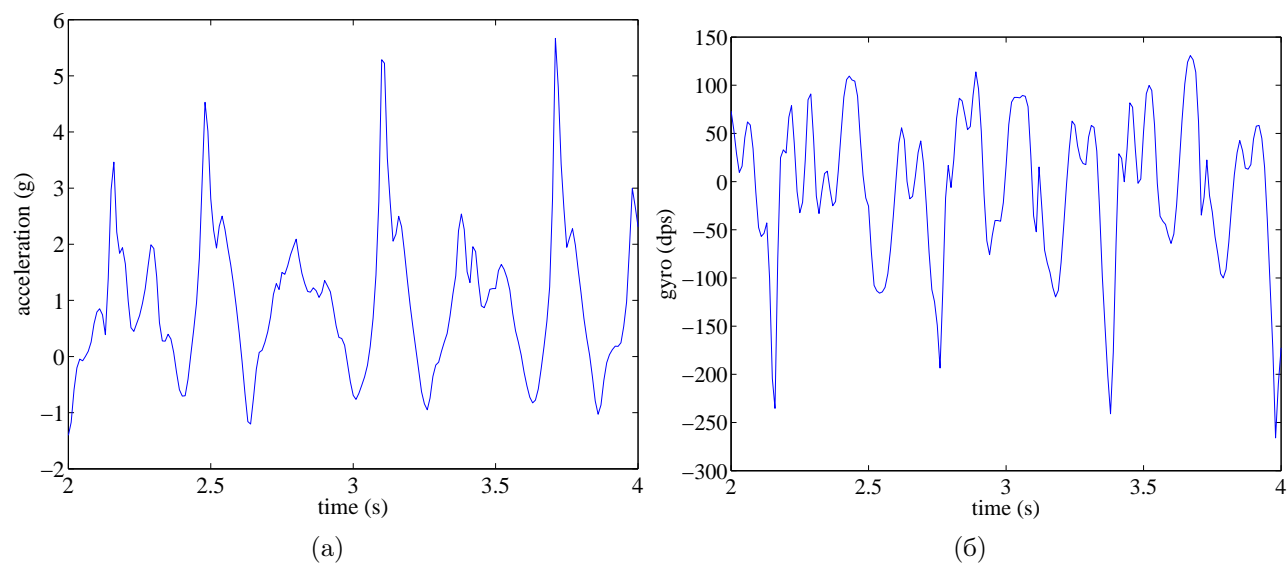


Рис. 1. Фрагмент показаний одной из осей акселерометра (а) и гироскопа (б) для класса «бег»

## Метод решения

Прежде чем перейти непосредственно к описанию моделей, необходимо упомянуть несколько достаточно стандартных идей, нашедших применение также и в предложенном в этой работе методе:

- Для каждого объекта рассматриваются окна шириной в 512 отсчетов, соседние перекрываются на 50%. Каждое такое окно, содержащее по  $512 \cdot 6$  значений, будем рассматривать как отдельный объект одного из 12 исходных классов.

Строго говоря, это уже другая задача классификации, но используя ее результат, можно получить решение для *исходной* (далее везде будет использоваться именно такое название), если построить решающее правило по принципу голосования: с учетом результатов классификации окон строится множество голосов для данного объекта исходной тестовой выборки, затем этот объект относится к тому классу, за который отдано наибольшее количество голосов, т. е.

$$a(x) = \arg \max_{y \in Y} |\{v \in \mathcal{V}(x) | v = y\}|. \quad (1)$$

Здесь и далее  $x$  — объект *исходной* задачи классификации;  $Y$  — множество меток классов;  $\mathcal{V}(x)$  — множество (формально — мультимножество) всех голосов, соответствующих объекту  $x$ . Классификаторы для *исходной* задачи будут обозначаться  $a(\cdot)$ , а классификаторы для окон —  $\tilde{a}(\cdot)$ . Множество  $\mathcal{V}(x)$  будет в разных случаях вводиться несколько по-разному с использованием множества  $\mathcal{W}(x)$  окон объекта  $x$ ;

- по аналогии с [6], все классы делятся на две группы: *активные* и *пассивные*. Классы 1–7 — активные, классы 8–12 — пассивные. Классификатор строится по следующему принципу: сначала тестовый объект относится одним классификатором к одной из двух групп, а затем подается на вход другому алгоритму для классификации внутри данной группы.

Эффективность такого подхода заключается в том, что, как будет показано далее, пассивные и активные действия хорошо разделяются с использованием очень простых

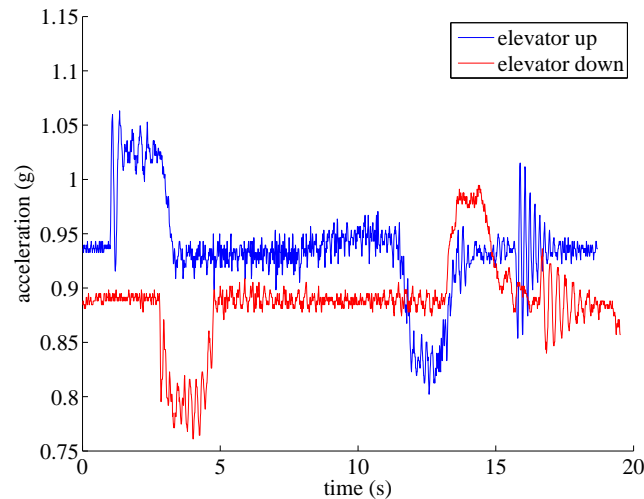


Рис. 2. Пример: передвижение на лифте вверх и вниз, ускорение по вертикали

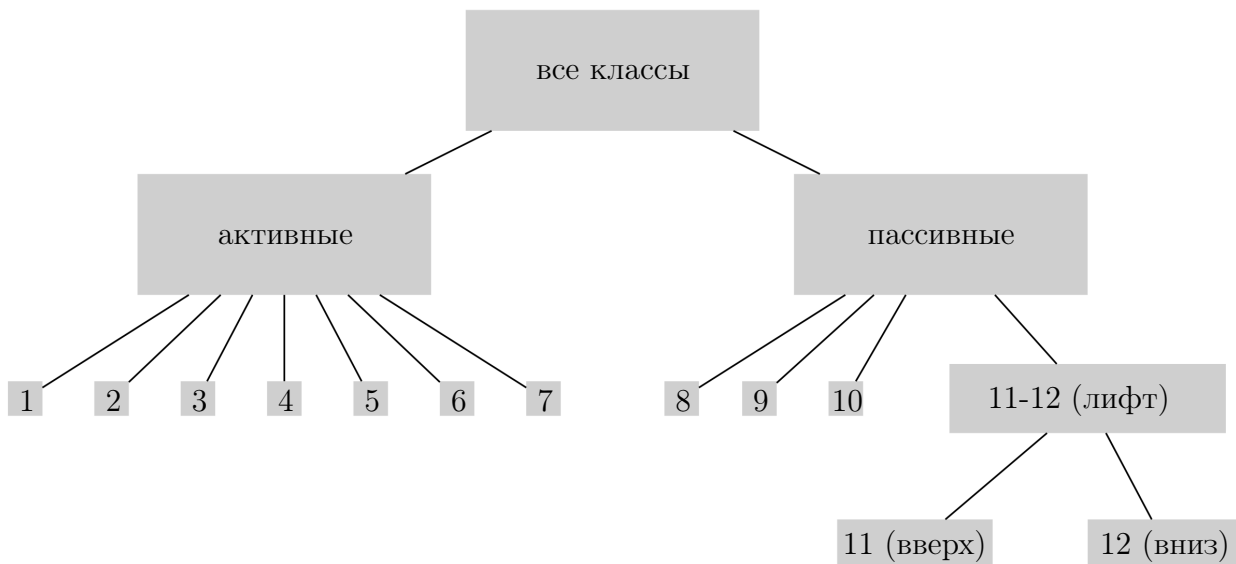
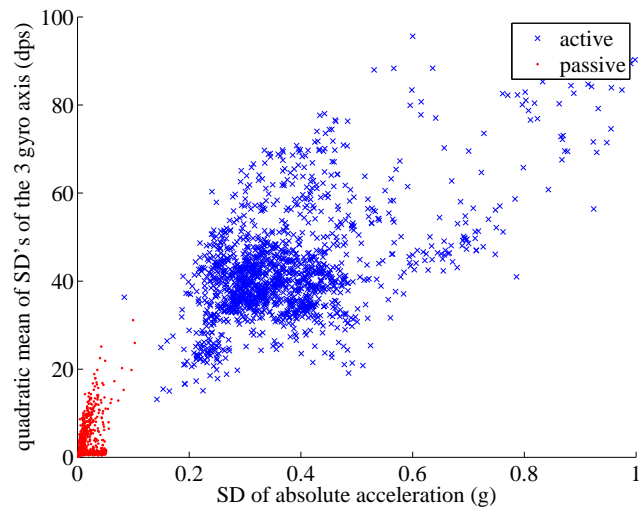


Рис. 3. Иерархия классов

признаков, в то время как для дальнейшей классификации удобно использовать признаки совсем иной природы;

- необходимо отметить, что подход с выделением окон и голосованием не применим для разделения классов 11 и 12 (передвижение на лифте вверх и вниз). Сигналы этих классов имеют характерные «всплески», порядок которых необходимо учесть (рис. 2). Поэтому изначально при обучении и выводе эти классы рассматриваются как один, и только на последней стадии разделяются отдельным классификатором, не использующим окна.



**Рис. 4.** Объекты, соответствующие активным и пассивным группам классов

В пояснение к последним двум идеям на рис. 3 изображена иерархия классов, в соответствии с которой производится классификация. Каждой нелистой вершине дерева соответствует свой классификатор.

Голосование происходит на каждом отдельном этапе, кроме этапа разделения классов 11 и 12, на котором используются не окна, а объекты исходной выборки.

Наибольший интерес представляют алгоритмы классификации, соответствующие узлам всех активных и всех пассивных классов. Другие два классификатора очень просты, и рассматриваются как вспомогательные.

### Разделение активных и пассивных действий

В вершине предложенной иерархии происходит разделение классов на активные и пассивные. Классификатор, используемый для такого разделения, устроен очень просто.

Для каждого окна можно вычислить стандартное отклонение (standard deviation, SD) абсолютного значения ускорения и среднее квадратичное стандартных отклонений трех каналов угловых скоростей. Эти два признака в некоторой мере отражают «степень активности», фиксируемую акселерометром и гироскопом, соответственно.

На рис. 4 видно, что выборка хорошо разделима по первому признаку. Вторым же признаком выглядит менее информативным, поэтому использоваться не будет.

Далее можно воспользоваться пороговым классификатором по первому признаку

$$\tilde{a}(w) = [f(w) - b > 0] = \begin{cases} 1, & \text{если } f(w) - b > 0; \\ 0, & \text{иначе} \end{cases}$$

с функцией потерь

$$\ell(M) = (-M)_+ = \begin{cases} -M, & \text{если } M < 0; \\ 0, & \text{иначе,} \end{cases}$$

где  $M(x)$  — отступ объекта  $x$ ;  $w$  — окно;  $f(w)$  — признак окна  $w$  (среднее квадратичное стандартных отклонений).

Такой классификатор имеет единственный настраиваемый параметр  $b$  — значение порога. Средняя функция потерь на всей выборке может иметь плато, поэтому в качестве

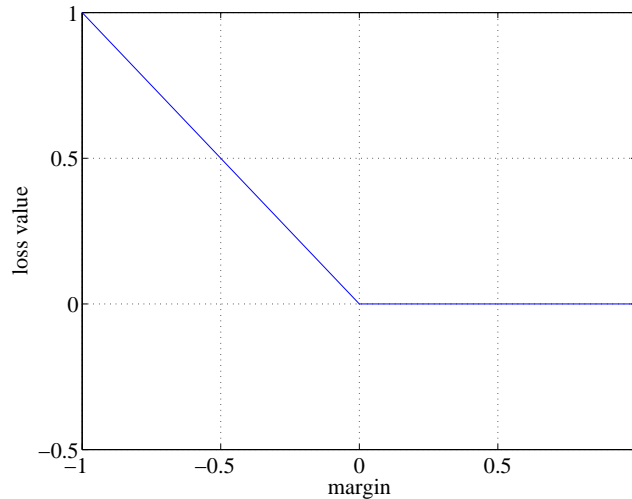


Рис. 5. Функция потерь  $\ell(M) = (-M)_+$

порога берется среднее арифметическое минимального и максимального оптимальных значений.

Голосование строится в соответствии с (1), множество голосов  $\mathcal{V}(x)$  определяется следующим образом:

$$\mathcal{V}(x) = \{\tilde{a}(w) | w \in \mathcal{W}(x)\},$$

где  $\mathcal{W}(x)$  — множество окон объекта  $x$ ;  $\tilde{a}(\cdot)$  — классификатор окон.

Несмотря на примитивность, этот метод показывает качество 100% в рамках исходной задачи классификации.

### Классификация передвижений на лифте

Как видно на рис. 2, для движения лифта вверх характерен «скачок» ускорения по вертикали вначале и «провал» в конце, а для движения вниз — наоборот. Поэтому самый простой способ разделить эти два класса — посчитать среднее ускорение по вертикали в первой половине всего временного интервала и вычесть из него среднее ускорение по вертикали во второй половине.

Таким образом, в данном случае классификатор описывается следующим образом:

$$\tilde{a}(w) = \begin{cases} 0, & \text{если } m_1(x) < m_2(x); \\ 1, & \text{иначе,} \end{cases}$$

где  $m_1(x)$  и  $m_2(x)$  — среднее ускорение по вертикали в первой и во второй половинах временного интервала, соответственно.

Как видно по графику на рис. 6, такой способ дает почти идеальное качество классификации. Кроме того, получившийся классификатор не требует обучения и имеет вполне естественный логический смысл.

### Основная модель

Для классификации внутри групп активных и пассивных классов используется общая модель, но обучение происходит отдельно. При этом в группе пассивных действий классы, соответствующие передвижению на лифте вверх и вниз, на данном этапе объединены.

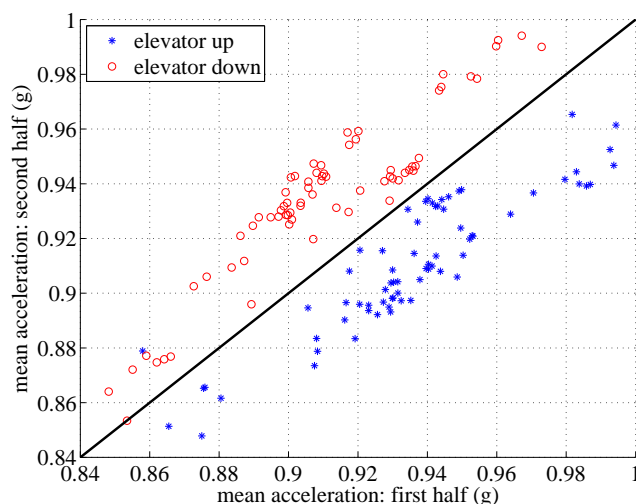


Рис. 6. Движение лифта вверх и вниз — делимость классов

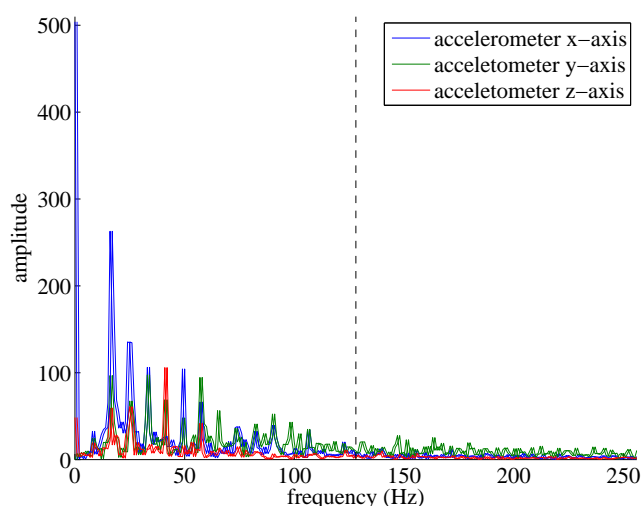


Рис. 7. Амплитудный спектр для показаний акселерометра. Для классификации используются только значения из левой половины графика

В качестве признакового пространства используются низкочастотные половины (128 значений) амплитудных спектров всех шести сигналов, полученные с помощью преобразования Фурье (рис. 7).

Каждый признак центрируется и нормируется на квадратный корень из нормы.

Модель состоит из смеси логистической регрессии, метода ближайшего соседа (Nearest Neighbor) и случайного леса (Random Forest), обозначаемых  $\tilde{a}_1(\cdot)$ ,  $\tilde{a}_2(\cdot)$  и  $\tilde{a}_3(\cdot)$  соответственно. При этом с помощью логистической регрессии также производится отбор признаков, используемый для каждой из двух других подмоделей.

Далее приведено более подробное описание всех подмоделей.

1. **Логистическая регрессия с  $L_1$ -регуляризацией.** Естественно предположить, что спектр содержит много избыточной информации, поэтому имеет смысл рассмотреть линейную логистическую регрессию с  $L_1$ -регуляризацией. В данной работе используется реализация из библиотеки LIBLINEAR [9] с интерфейсом для системы MATLAB.

Это обычная двухклассовая логистическая с добавлением единичного константного признака (bias), обучаемая по принципу «один против всех».

Единственный настраиваемый гиперпараметр — значение  $C$ , обратное коэффициенту регуляризации, — настраивается на скользящем контроле.

С использованием коэффициентов обученной логистической регрессии вводится понятие значимости признаков.

Пусть  $p(\mathbf{x}) = (1 + \exp(-W^T \mathbf{x}))^{-1}$  — вектор оценок вероятности принадлежности объекта  $\mathbf{x}$  к классам (ненормированный);  $W$  — настроенные веса, каждый столбец соответствует своему классу. Тогда мера значимости  $i$ -го признака будет определяться как  $s_i = \max_j |W_{ij}|$ .

2. **Nearest Neighbor.** В качестве следующей подмодели берется обычный метод одного ближайшего соседа с косинусной мерой. В силу того, что признаки уже нормализованы, косинусная мера в данном случае эквивалентна корреляции. Из признаков используется только половина (т. е. 64) наиболее значимых. В качестве ответа возвращается класс самого ближайшего по косинусной мере объекта из обучающей выборки.
3. **Random Forest.** Используется реализация бэггинга над решающими деревьями из пакета Statistics Toolbox системы MATLAB. Количество деревьев — 1000, для остальных параметров используются значения по умолчанию. Здесь, по аналогии с NN, используется только половина наиболее значимых признаков.

Все три подмодели обучаются независимо, если не считать отбор признаков. В рамках исходной задачи, на стадии вывода для каждого объекта тестовой выборки «голоса» моделей объединяются, после чего принимается решение о классификации данного объекта по формуле (1). Формально  $V(x)$  в данном случае определяется так:

$$\mathcal{V}(x) = \mathcal{V}_1(x) \cup \mathcal{V}_2(x) \cup \mathcal{V}_3(x),$$

где  $\mathcal{V}_i(x) = \{\tilde{a}_i(w) | w \in \mathcal{W}(x)\}$ .

Такая композиция — не что иное, как усреднение трех различных подмоделей.

## Эксперименты

### Способ оценки качества

Прогноз модели делается на скользящем контроле, затем итоговое качество вычисляется как усредненная F-мера по всем классам. *Каждый фолд скользящего контроля соответствует отдельному человеку.* Таким образом, всего имеется 14 фолдов: в  $i$ -й входят данные, соответствующие  $i$ -му человеку.

Параметры нормировки признаков (средние значения и нормы) вычисляются на каждой итерации скользящего контроля по текущей обучающей подвыборке.

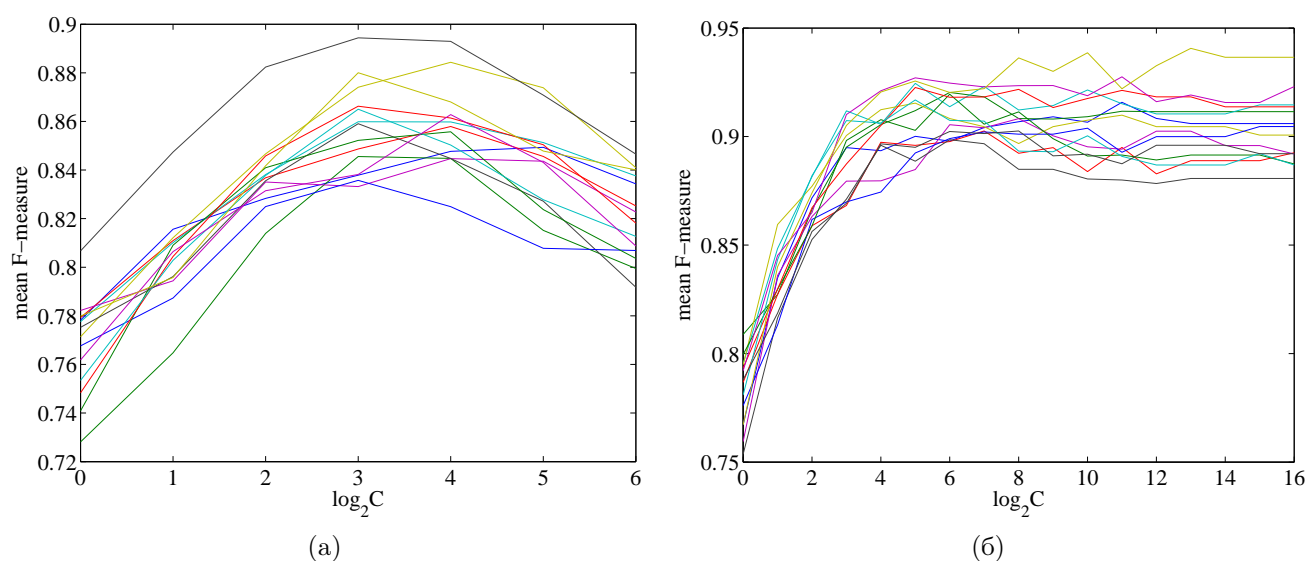
Везде оценивается качество в рамках исходной задачи классификации.

### Результаты

Далее изложены результаты экспериментов для отдельных классификаторов, для подмоделей и для всей композиции. Некоторые из этих результатов уже упоминались выше. Также приводится сравнение с простой логистической регрессией.

- классификация активных и пассивных действий дает идеальное качество: 1,00;
- классификация направлений движения лифта также показывает очень хорошее качество: 0,99;





**Рис. 8.** Зависимость среднего качества от логарифма величины  $C$  для активных (а) и пассивных (б) классов. Каждая точка на графике — среднее качество для конкретного значения  $C$  на «локальном» скользящем контроле, в рамках одной итерации «глобального». Точки, соответствующие одной итерации «глобального» контроля, объединены в одноцветные ломаные

— одна из трех компонент основной модели — логистическая регрессия с  $L_1$ -регуляризацией — использует настройку гиперпараметра  $C$ . Этот гиперпараметр настраивается на «локальном» скользящем контроле по схеме, описанной выше, но уже с 13-ю фолдами. Таким образом, получается, своего рода, «вложенный» скользящий контроль. Это необходимо потому, что гиперпараметр  $C$  в данной задаче существенно влияет на качество, и его необходимо настроить, и в то же время не переобучиться.

Результаты скользящего контроля приведены на графиках на рис. 8.

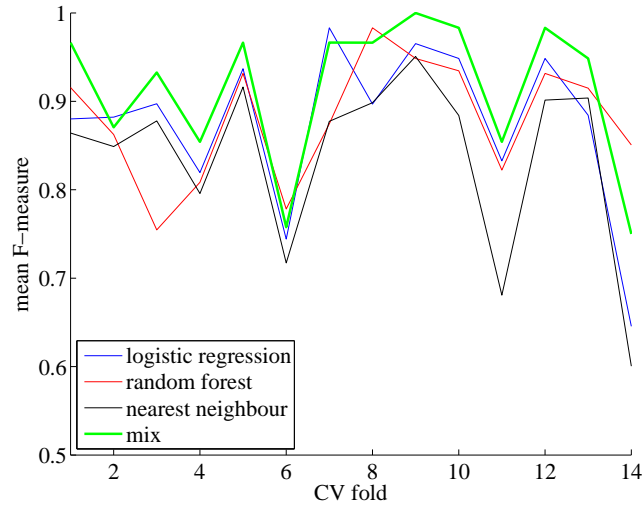
Видно, впрочем, что регуляризация для группы пассивных классов дает слабый выигрыш, поскольку большие значения  $C$  (и, соответственно, малые значения коэффициента регуляризации  $C^{-1}$ ) обеспечивают почти оптимальное качество;

— смесь из трех подмоделей (логистическая регрессия, random forest и NN) часто показывает лучшее качество, чем каждая подмодель в отдельности. Это видно по графикам на рис. 9, отображающим качество классификации для разных итераций скользящего контроля, соответствующих отдельным людям;

— **вся композиция** на трех независимых запусках показала качество 0,9202, 0,9216 и 0,9227. Полные результаты классификации приведены в табл. 1.

Как было сказано выше, решающее правило действует по принципу голосования (1). Этот принцип имеет недостаток: могут возникать неопределенности, когда наибольшее количество голосов набирают 2 или более варианта. Однако в данном эксперименте это произошло всего в 9 случаях из 840.

— автору известен только один опубликованный результат классификации набора данных USC-HAD [10] (он также был указан автору рецензентом), однако сравнение с ним не представляется возможным в силу отсутствия описания деталей эксперимента и предложенного метода. Кроме того, насколько можно судить из текста работы, в ходе эксперимента окна, соответствующие одному и тому же замеру, и даже перекрываю-



**Рис. 9.** Качество подмоделей и их смеси на разных итерациях скользящего контроля

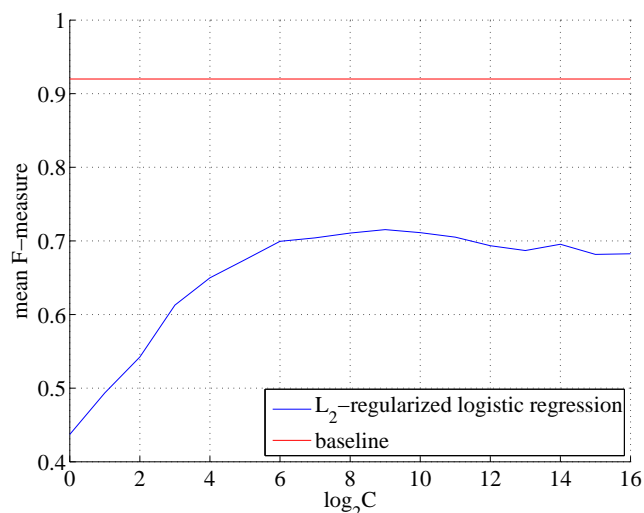
**Таблица 1.** Результаты классификации (confusion matrix). Столбцы соответствуют меткам классов, строки — прогнозам. В каждой ячейке — абсолютное количество соответствующих объектов

	1	2	3	4	5	6	7	8	9	10	11	12
1	70	1	2	6	8	1	0	0	0	0	0	0
2	0	64	2	0	0	0	0	0	0	0	0	0
3	0	2	65	0	0	0	0	0	0	0	0	0
4	0	0	1	64	1	0	1	0	0	0	0	0
5	0	3	0	0	57	1	5	0	0	0	0	0
6	0	0	0	0	3	68	1	0	0	0	0	0
7	0	0	0	0	1	0	63	0	0	0	0	0
8	0	0	0	0	0	0	0	62	8	0	0	0
9	0	0	0	0	0	0	0	3	59	0	2	3
10	0	0	0	0	0	0	0	4	0	70	0	0
11	0	0	0	0	0	0	0	0	2	0	67	1
12	0	0	0	0	0	0	0	1	1	0	1	66
F-мера	0,89	0,94	0,95	0,93	0,84	0,96	0,94	0,89	0,86	0,97	0,96	0,95

щиеся окна могли попасть в обучающую и контрольную выборки, что, как правило, значительно поднимает качество классификации.

Поэтому для сравнения выбран простой метод — логистическая регрессия с  $L_2$ -регуляризацией и добавлением константного признака, работающая по принципу «один против всех» на всех 12 классах и всех признаках (всем амплитудном спектре). Иерархия классов никак не используется, но принцип выбора окон и голосования используется тот же.

Такая логистическая регрессия также имеет единственный гиперпараметр  $C$ , равный обратному значению коэффициента регуляризации. Метод протестирован всех для значений  $C = 2^i$ ,  $i = 0..16$ . Результаты приведены на рис. 10. Максимальное качество не превысило значения 0.72.



**Рис. 10.** Качество классификации логистической регрессии с  $L_2$ -регуляризацией для различных значений параметра  $C$  в сравнении с качеством предложенного метода (baseline)

## Заключение

С использованием достаточно простых признаков и относительно сложного по структуре классификатора получено достаточно высокое качество классификации.

Метод получился вполне устойчивым: при небольших изменениях параметров (таких как доля отбираемых признаков, степень перекрытия соседних окон, значения  $C$  и т. д.) качество остается в пределах от 0,91 до 0,93. Это можно объяснить тем, что в основной модели используется смесь трех принципиально разных подходов: линейного классификатора, бэггинга над решающими деревьями и метрического классификатора. Эти подмодели по-разному реагируют на изменение параметров, поэтому усредненный результат остается почти неизменным. Остальные же классификаторы композиции очень просты по своей природе и имеют простую интерпретацию.

Наконец, отдельного внимания заслуживает вопрос о применимости предложенного метода на практике. Тут можно отметить следующее:

- предложенный классификатор направления движения лифтов вряд ли представляет практический интерес, но в условиях поставленной задачи это приемлемое решение. Автор постарался лишь удовлетворить формальному требованию классификации, чтобы оценить качество всей композиции. Задача классификации сигналов с единичными «всплесками» несколько выходит за рамки данной работы;
- метод требует достаточно трудоемких вычислений. Однако, во-первых, мощность процессоров постоянно растет, а во-вторых, логистическая регрессия – самая «легковесная» подмодель – очень эффективна с точки зрения вычислительных ресурсов и показывает неплохое качество: 0,89.

Работа выполнена в рамках спецсеминара «Алгебра над алгоритмами и эвристический поиск закономерностей» кафедры Математических методов прогнозирования факультета ВМК МГУ под научным руководством д.ф.-м.н., профессора Дьяконова Александра Геннадьевича. Автор также выражает благодарность Илье Владимировичу Сафонову (Nokia Research) за ценные советы и замечания.

## Литература

- [1] *Hoseini-Tabatabaei S. A., Gluhak A., Tafazolli R.* A survey on smartphone-based systems for opportunistic user context recognition // *ACM Computing Surveys (CSUR)*, 2013. Vol. 45, No. 3. P. 27.
- [2] *Avci A., Bosch S., Marin-Perianu M., Marin-Perianu R., Havinga P.* Activity recognition using inertial sensing for healthcare, wellbeing and sports applications: A survey // *23rd Conference (International) on Architecture of Computing Systems (ARCS)*, 2010. P. 1–10.
- [3] *Zhao Z., Chen Y., Liu J., Shen Z., Liu M.* Cross-people mobile-phone based activity recognition // *22nd Joint Conference (International) on Artificial Intelligence Proceedings*. AAAI Press, 2011. Vol. 3. P. 2545–2550.
- [4] *Dernbach S., Das B., Krishnan N. C., Thomas B. L., Cook D. J.* Simple and complex activity recognition through smart phones // *8th IEEE Conference (International) on Intelligent Environments*, 2012. P. 214–221.
- [5] *Yan Z., Subbaraju V., Chakraborty D., Misra A., Aberer K.* Energy-efficient continuous activity recognition on mobile phones: An activity-adaptive approach // *16th IEEE Symposium (International) on Wearable Computers (ISWC)*, 2012. P. 17–24.
- [6] *Siirtola P., Roning J.* Recognizing Human Activities User-independently on Smartphones Based on Accelerometer Data // *Int. J. Interactive Multimedia Artificial Intelligence*, 2012. Vol. 1, No. 5.
- [7] *Incel O. D., Kose M., Ersoy C.* A review and taxonomy of activity recognition on mobile phones // *BioNanoScience*, 2013. Vol. 3, No. 2. P. 145–171.
- [8] *Zhang M., Sawchuk A. A.* USC-HAD: A Daily Activity Dataset for Ubiquitous Activity Recognition Using Wearable Sensors // *ACM Conference (International) on Ubiquitous Computing (UbiComp) Workshop on Situation, Activity and Goal Awareness (SAGAware)*. Pittsburgh, Pennsylvania, USA, 2012.
- [9] *Fan R. E., Chang K W., Hsieh C J, et al.* LIBLINEAR: A library for large linear classification // *J. Machine Learning Res.*, 2008. Vol. 9. P. 1871–1874.
- [10] *Cui J., Xu B.* Cost-effective activity recognition on mobile devices // *8th Conference (International) on Body Area Networks Proceedings*. ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), 2013. P. 90–96.