

Классификация временных рядов в пространстве параметров порождающих моделей*

Карасиков М. Е.

karasikov@phystech.edu

Московский физико-технический институт

Работа посвящена задаче многоклассовой признаковой классификации временных рядов. Признаковая классификация временных рядов заключается в сопоставлении каждому временному ряду его краткого признакового описания, позволяющего решать задачу классификации в пространстве признаков. В работе исследуются методы построения пространства признаков временных рядов. При этом временной ряд рассматривается как последовательность сегментов, аппроксимируемых некоторой параметрической моделью, параметры которой используются в качестве их признаковых описаний. Построенное так признаковое описание сегмента временного ряда наследует от модели аппроксимации такие полезные свойства, как инвариантность относительно сдвига. Для решения задачи классификации в качестве признаковых описаний временных рядов предлагается использовать распределения параметров аппроксимирующих сегменты моделей, что обобщает базовые методы, использующие непосредственно сами параметры аппроксимирующих моделей. Проведен ряд вычислительных экспериментов на реальных данных, показавших высокое качество решения задачи многоклассовой классификации. Эксперименты показали превосходство предлагаемого метода над базовым и многими распространенными методами классификации временных рядов на всех рассмотренных наборах данных.

Последние изменения: 18:15, 27 июня 2015 г.

Ключевые слова: *временные ряды, многоклассовая классификация, сегментация временных рядов, гиперпараметры аппроксимирующей модели, модель авторегрессии, дискретное преобразование Фурье, дискретное вейвлет-преобразование.*

Time series feature-based classification*

Karasikov M. E.

Moscow Institute of Physics and Technology

Введение

Временные ряды являются результатом проведения любых повторяющихся во времени измерений. Обзор по методам и проблемам анализа временных рядов дается в [1, 2]. Одними из основных методов анализа временных рядов являются прогнозирование [3, 4, 5], обнаружение аномалий [6, 7, 8], сегментация [9, 10, 11], кластеризация [12, 13, 14] и классификация [15, 16, 17, 18]. Последние годы связаны с ростом интереса к данной области, проявляющимся в предложениях новых методов анализа временных рядов — метрик [19, 9, 20, 21, 22, 23], алгоритмов сегментации [24, 25, 26, 27], кластеризации [28, 29, 30], и др.

Временным рядом x будем называть конечную упорядоченную последовательность:

$$x = [x^{(1)}, \dots, x^{(t)}].$$

В данной работе рассматривается задача классификации временных рядов, которая возникает во многих приложениях (медицинская диагностика по ЭКГ [31, 32] и ЭЭГ [33], классификация физической активности по данным с акселерометра [34, 35], верификация динамических подписей [36, 37] и т. д.) и заключается в определении неизвестных классов временных рядов рассматриваемого множества.

Формально задача классификации в общем виде может быть поставлена следующим образом. Пусть X — множество описаний объектов произвольной природы, Y — конечное множество меток классов. Предполагается существование целевой функции — отображения $y : X \rightarrow Y$, значения которого известны только на объектах обучающей выборки

$$\mathfrak{D} = \{(x_1, y_1), \dots, (x_m, y_m)\} \subset X \times Y.$$

Требуется построить алгоритм $a : X \rightarrow Y$ — отображение, приближающее целевую функцию y на множестве X . Задачей классификации временных рядов будем называть задачу классификации, в которой объектами классификации являются временные ряды.

Этап построения информативного пространства признаков, позволяющего добиться заданной точности классификации, является одним из важнейших этапов решения задачи классификации.

Одним из способов построения пространства признаков в задаче классификации временных рядов является задание функции расстояния [19, 9, 20, 21, 22, 23] между временными рядами, позволяющего в качестве признаков взять расстояния до опорных объектов. Данный метод чрезвычайно распространен в силу того, что позволяет свести исходную задачу классификации временных рядов к задаче выбора метрики — функции расстояния. При удачном выборе метрики дальнейшая классификация может происходить при помощи простейших метрических алгоритмов классификации, например, методом ближайшего соседа [38]. Данный подход так же зарекомендовал себя в self-training [18] и graph-based [39, 40] методах частичного обучения.

Другой способ состоит в извлечении из каждого временного ряда набора признаков — его информативного описания, позволяющего строить точные классификаторы с хорошей обобщающей способностью. В качестве признаков могут браться буквально произвольные функции $f : X \rightarrow \mathbb{R}^n$ исходных объектов. Например, признаки могут задаваться экспертом. Так в работе [41] предлагается использовать в качестве признаков статистические функции (среднее, отклонения от среднего, коэффициенты эксцесса и др.). Стоит заметить, что при таком подходе к построению пространства признаков часто удается добиться необходимого качества классификации путем выбора соответствующих конкретной задаче признаков (см. пример [42]). Второй метод построения пространства признаков заключается в задании параметрической регрессионной или аппроксимирующей модели временного ряда. Таким образом, в качестве признаков временных рядов будут выступать параметры заданной модели. При этом возникает задача выбора модели для описания временного ряда []. Так в работе [43] в качестве признаков предлагается использовать коэффициенты дискретного преобразования Фурье (DFT). В [43, 44] предлагается использовать дискретное вейвлет-преобразование (DWT), которое сравнивается с предыдущими методами. Особенно эффективно DFT и DWT работают на квазипериодических рядах, где проявляется периодическая структура.

Метод извлечения признаков на основе регрессионных моделей является общим и может быть использован при работе с произвольными объектами. Однако, временные ряды всегда являются некоторыми конечными последовательностями, из которых можно выделять фрагменты — некоторые подпоследовательности. Каждый из фрагментов тоже

является временным рядом, а, следовательно, для него тоже может быть получено признаковое описание. Таким образом, каждый временной ряд может быть описан распределением признаков его фрагментов. Заметим, что формально данный подход является обобщением выше изложенного, в котором из ряда выделяется единственный фрагмент — сам ряд, а функции распределения признаков вырождены.

В частном случае, когда временной ряд является квазипериодическим, существует возможность выделения в нем характерных сегментов — периодов, то есть возможность представления каждого временного ряда $x = [x^{(1)}, \dots, x^{(t)}]$ последовательностью в определенном смысле похожих его сегментов $s^{(1)}, \dots, s^{(p)}$:

$$s^{(1)} = [x^{(1)}, \dots, x^{(t_1)}], \dots, s^{(k)} = [x^{(t_{k-1}+1)}, \dots, x^{(t_k)}], \dots, s^{(p)} = [x^{(t_{p-1}+1)}, \dots, x^{(t)}].$$

В таком случае будем писать

$$x = (s^{(1)}, \dots, s^{(p)}). \quad (1)$$

Для выделения периодов могут быть использованы, например, алгоритмы [9, 10, 11, 24, 25, 26, 27].

В нашей работе предлагается алгоритм классификации квазипериодических временных рядов на основе распределения параметров модели, описывающей сегменты временных рядов. Предлагаемый подход к классификации квазипериодических временных рядов в общем виде изложен в разделе 1. В разделе 1 представлены эксперименты на реальных данных по сравнению предложенного подхода с подходом, предложенным в [34].

Постановка задачи

Поставим задачу в общем виде.

Пусть (X, d) — метрическое пространство временных рядов, Y — множество меток классов. Дано конечное множество временных рядов $X^\ell = \{x_1, \dots, x_\ell\} \subset X$ и обучающая выборка $\mathfrak{D} \subset X^\ell \times Y$.

Заданы алгоритм выделения фрагментов временного ряда

$$s : x \mapsto (s^{(1)}, \dots, s^{(p)}) \in X^p, \text{ где } p = p(x), \quad (2)$$

и параметрическая регрессионная модель фрагментов временного ряда

$$g : \mathbb{R}^n \times X \rightarrow X. \quad (3)$$

Каждому фрагменту $s^{(k)}$ временного ряда x поставим в соответствие его вектор признаков

$$\mathbf{f}(s^{(k)}) = \arg \min_{\mathbf{w} \in \mathbb{R}^n} d(g(\mathbf{w}, s^{(k)}), s^{(k)}). \quad (4)$$

Таким образом, каждому временному ряду x соответствует набор векторов признаков

$$\mathbf{f} = (\mathbf{f}(s^{(1)}), \dots, \mathbf{f}(s^{(p)}))$$

его фрагментов.

Задано некоторое семейство алгоритмов классификации

$$A = \{a : \mathbf{f} \mapsto y \in Y\} \quad (5)$$

и функция потерь

$$\mathcal{L} : X \times Y \rightarrow \mathbb{R}.$$

Найти алгоритм классификации $a^* \in A$, доставляющий минимум функционалу качества $Q(a, \mathfrak{D}) \in \mathbb{R}$, $a \in A$:

$$\begin{aligned} a^* &= \arg \min_{a \in A} Q(a, \mathfrak{D}) = \\ &= \arg \min_{a \in A} \frac{1}{|\mathfrak{D}|} \sum_{(x_i, y_i) \in \mathfrak{D}} \mathcal{L}(a(\mathbf{f}_i), y_i). \end{aligned}$$

Выделение фрагментов

В данном разделе рассматривается проблема выделения фрагментов из временного ряда (см. 2). Фрагментом временного ряда $x = [x^{(1)}, \dots, x^{(t)}]$ будем называть любую его подпоследовательность $s = [x^{(i_1)}, \dots, x^{(i_k)}]$, где $1 \leq i_1 \leq \dots \leq i_k \leq t$. Например, сам временной ряд x является своим фрагментом. То есть в качестве алгоритма выделения временного ряда можно взять

$$s : x \mapsto x \in X. \quad (6)$$

Пусть теперь временной ряд является квазипериодическим. То есть справедливо представление 1. Тогда возьмем в качестве фрагментов ряда его сегменты

$$s : x \mapsto (s^{(1)}, \dots, s^{(p)}) \in X^p. \quad (7)$$

Сегментация временных рядов может проводиться согласно алгоритмам [9, 10, 11, 24, 25, 26, 27]. Общие требования к процедуре выделения фрагментов из временных рядов сформулированы в разделе 1.

Регрессионная модель временного ряда

В качестве признаков фрагментов временного ряда используется вектор параметров аппроксимирующей модели 3. Рассмотрим несколько простых примеров.

— **Простая линейная регрессия.** Пусть задан многокомпонентный временной ряд (например, время и 3 пространственные координаты):

$$x = [\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(t)}], \quad \text{где} \quad \mathbf{x}^{(k)} = [z_0^{(k)}, \dots, z_r^{(k)}]^\top.$$

Тогда нашей аппроксимирующей моделью, например, может быть модель линейной регрессии одной из компонент временного ряда на остальные компоненты:

$$g(\mathbf{w}, x) = [\hat{\mathbf{x}}^{(1)}, \dots, \hat{\mathbf{x}}^{(t)}], \quad \text{где} \quad \hat{\mathbf{x}}^{(k)} = [\hat{z}_0^{(k)}, z_1^{(k)}, \dots, z_r^{(k)}]^\top, \quad k = 1, \dots, t,$$

$$\hat{z}_0 = \begin{bmatrix} \hat{z}_0^{(1)} \\ \vdots \\ \hat{z}_0^{(t)} \end{bmatrix} = \begin{bmatrix} \hat{z}_1^{(1)} & \dots & \hat{z}_r^{(1)} \\ \vdots & \ddots & \vdots \\ \hat{z}_1^{(t)} & \dots & \hat{z}_r^{(t)} \end{bmatrix} \begin{bmatrix} w_1 \\ \vdots \\ w_r \end{bmatrix}.$$

— **Авторегрессионная модель $\text{AR}(k)$.** Задан временной ряд

$$x = [x^{(1)}, \dots, x^{(t)}], \quad x^{(k)} \in \mathbb{R}.$$

Выберем в качестве модели порождения авторегрессионную модель порядка p :

$$g(\mathbf{w}, x) = [\hat{x}^{(1)}, \dots, \hat{x}^{(t)}], \quad \text{где} \quad \hat{x}^{(k)} = w_0 + \sum_{i=1}^p w_i x^{k-i}, \quad k = 1, \dots, t. \quad (8)$$

— **Дискретное преобразование Фурье.** Задан временной ряд

$$x = [x^{(0)}, \dots, x^{(t-1)}], \quad x^{(k)} \in \mathbb{R}.$$

Очевидно, что задание признаковов описания временного ряда коэффициентами Фурье эквивалентно заданию порождающей модели

$$g(\mathbf{w}, x) = [\hat{x}^{(1)}, \dots, \hat{x}^{(t)}], \quad \text{где} \quad \hat{x}^{(k)} = \frac{1}{t} \sum_{j=0}^{t-1} w_j e^{\frac{2\pi i}{t} k j}, \quad k = 0, \dots, t-1$$

и евклидова расстояния в пространстве временных рядов.

Приведенные примеры демонстрируют большую общность построения пространства признаков при помощи модели типа 3 и решения оптимизационной задачи 4. Вообще, легко видеть, что любая функция временного ряда $\mathbf{f} : X \rightarrow \mathbb{R}^n$ может быть задана как решение оптимизационной задачи 4 с соответствующей моделью 3 и функцией расстояния d .

Алгоритм классификации

В данном разделе займемся задачей построения семейства алгоритмов 5. Пусть

$$\mathbf{f}_i = (\mathbf{f}_i^{(1)}, \dots, \mathbf{f}_i^{(p)})$$

— полученное признаковое описание временного ряда $x_i \in X^\ell$, где $\mathbf{f}_i^{(k)}$ — вектор признаков k -го фрагмента этого ряда. Признакомое описание \mathbf{f}_i можно рассматривать как реализацию выборки случайных величин

$$\mathbf{F}_i = (\mathbf{F}_i^{(1)}, \dots, \mathbf{F}_i^{(p)}).$$

Примем гипотезу о простоте выборок \mathbf{F}_i , $i = 1, \dots, \ell$. Данная гипотеза формулирует требования на метод фрагментации 2. Тогда временной ряд естественно описывать распределением параметров порождающей модели его сегментов. Пусть задана некоторая параметрическое семейство распределений $\{\mathbf{P}_\theta\}_{\theta \in \Theta}$. Тогда, получив оценку максимального правдоподобия

$$\theta_i = \arg \max_{\theta \in \Theta} L(\theta | \mathbf{f}_i),$$

будем считать её признаковым описанием исходного временного ряда. Таким образом, задача классификации временных рядов свелась к задаче классификации параметров распределений из семейства $\{\mathbf{P}_\theta\}_{\theta \in \Theta}$. При этом, имея априорное распределение параметра $G(v) = \mathbf{P}\{\theta < v\}$, целесообразно использовать вместо оценки максимального правдоподобия апостериорное математическое ожидание параметра

$$\theta_i = \mathbf{E} [\theta | \mathbf{f}_i].$$

Вычислительный эксперимент

Вычислительный эксперимент проводился на данных для задачи классификации типов физической активности человека.

Датасет WISDM [35]

содержит показания акселерометра для 6 видов человеческой активности:

- | | |
|-------------|---------------|
| 1. Jogging | 4. Downstairs |
| 2. Walking | 5. Sitting |
| 3. Upstairs | 6. Standing |

Необработанные временные ряды, представляющие из себя последовательность размеченных показаний акселерометра (по тройке чисел на каждый отсчет времени с интервалом в 50 миллисекунд), были разбиты на временные ряды длиной по 200 отсчетов (10 секунд).

Классы	Jogging	Walking	Upstairs	Downstairs	Sitting	Standing
Число объектов	1624	2087	549	438	276	231

Таблица 1. Распределение временных рядов по классам. Dataset: WISDM.

Ручное выделение признаков

Выбор признаков. В первом эксперименте в качестве признаков описаний временных рядов использовались их статистические функции. Каждой компоненте временного ряда сопоставлялись 40 чисел — её среднее, стандартное отклонение, средний модуль отклонения от среднего, гистограмма с 10 областями равной ширины. Полученные признаки для каждой компоненты объединялись и к ним добавлялся признак средней величины ускорения.

Классификатор. Задача многоклассовой классификации сводилась к задаче бинарной классификации при помощи подхода One-vs-One. В качестве бинарного классификатора использовался SVM с гауссовским ядром и параметрами $C = 8.5$, $\gamma = 0.12$.

Результаты. Для оценки качества решения использовалась процедура скользящего контроля. Исходная обучающая выборка \mathfrak{D} случайно разбивается m раз на обучающую и контрольную ($\mathfrak{D} = \mathfrak{L}_i \sqcup \mathfrak{T}_i$). В качестве внешнего критерия качества метода обучения μ бралось

$$\frac{1}{m} \sum_{i=1}^m Q(\mu(\mathfrak{L}_i), \mathfrak{T}_i),$$

где для средней точности (ассигуру) классификации объектов класса $c \in Y$

$$Q_c(a, \mathfrak{T}) = \frac{\sum_{\substack{(x,y) \in \mathfrak{T} \\ y=c}} \mathbf{1}\{a(x) = y\}}{\sum_{\substack{(x,y) \in \mathfrak{T} \\ y=c}} 1}, \quad (9)$$

а для среднего качества решения задачи многоклассовой классификации

$$Q(a, \mathfrak{T}) = \frac{1}{|\mathfrak{T}|} \sum_{(x,y) \in \mathfrak{T}} \mathbf{1}\{a(x) = y\}. \quad (10)$$

На диаграмме ниже приведено качество классификации, усредненное по $m = 50$ случайным разбиениям исходной выборки на тестовую и контрольную в отношении 7 к 3.

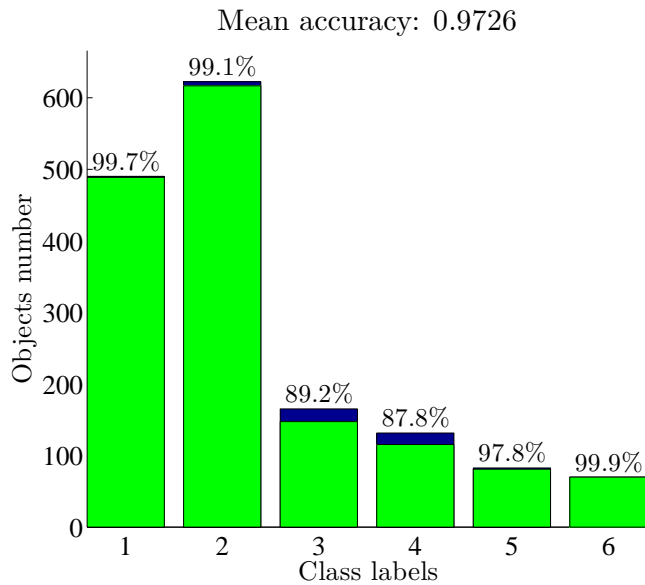


Рис. 1. Точность классификации при ручном выделении признаков. Dataset: WISDM. Под Mean accuracy понимается значение функционала 10. Над столбцами приведены средние точности классификации для каждого класса по формуле 9.

Как видно, классы 3 и 4 хорошо отделяются от остальных. Таким образом, если сначала отделить эту пару классов от остальных, а потом настроить бинарный классификатор для разделения классов 3 и 4, вводя дополнительные признаки (например те, которые будут рассматриваться далее), можно добиться требуемого качества решения задачи, но это не входит в цели данного эксперимента.

Модель авторегрессии 8

Признаковое описание. Во втором эксперименте в качестве признаков описаний временных рядов использовались все статистические функции, что брались в первом эксперименте 1, за исключением гистограммы. Вместо 10 значений для каждого блока гистограммы использовались 7 коэффициентов модели авторегрессии AR(6) (см. 8). Так же, проводилась предварительная нормализация признаков.

Классификатор. Задача многоклассовой классификации сводилась к задаче бинарной классификации при помощи подхода One-vs-All и экспоненциальной функцией потерь. В качестве бинарного классификатора использовался SVM с гауссовским ядром и параметрами $C = 8$, $\gamma = 0.8$.

Результаты. На диаграмме ниже приведено качество классификации, усредненное по $m = 50$ случайным разбиениям исходной выборки на тестовую и контрольную в отношении 7 к 3.

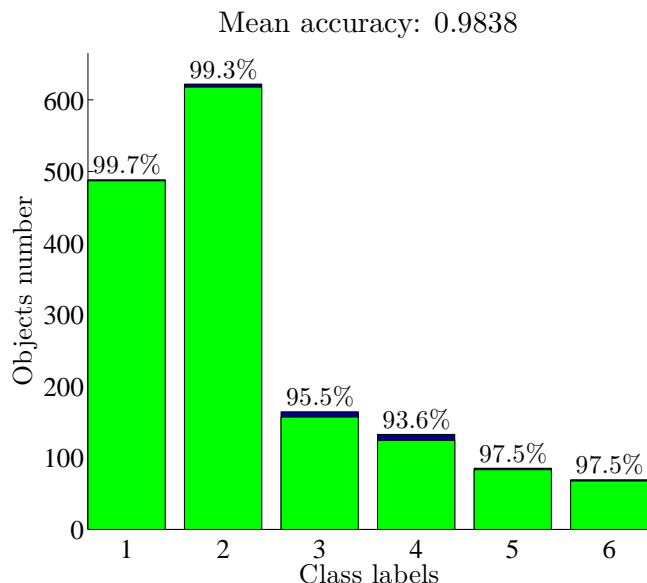


Рис. 2. Точность классификации для параметров модели авторегрессии в качестве признаков описаний. Dataset: WISDM. Под Mean accuracy понимается значение функционала 10. Над столбцами приведены средние точности классификации для каждого класса по формуле 9.

Датасет USC-HAD [45]

содержит показания акселерометра для 12 типов физической активности человека:

- | | |
|------------------|---------------------|
| 1. walk forward | 7. jump up and down |
| 2. walk left | 8. sit and fidget |
| 3. walk right | 9. stand |
| 4. go upstairs | 10. sleep |
| 5. go downstairs | 11. elevator up |
| 6. run forward | 12. elevator down |

Выборка содержит примерно по 70 шести-компонентных временных рядов для каждого класса, а средняя длина временного ряда порядка 3300. Частота записи измерений сенсора 100 Hz.

Ручное выделение признаков

Выбор признаков В качестве признаков брались те же признаки, что и в предыдущем эксперименте 1.

Классификатор Задача многоклассовой классификации сводилась к задаче бинарной классификации при помощи подхода One-vs-One. В качестве бинарного классификатора использовался SVM с гауссовским ядром и параметрами $C = 80$, $\gamma = 0.002$.

Результаты Исходная выборка 100 раз случайно разбивалась на обучающую и контрольную в отношении 7 к 3. В таблице ниже приведен результат — процент верной классификации для объектов каждого класса.

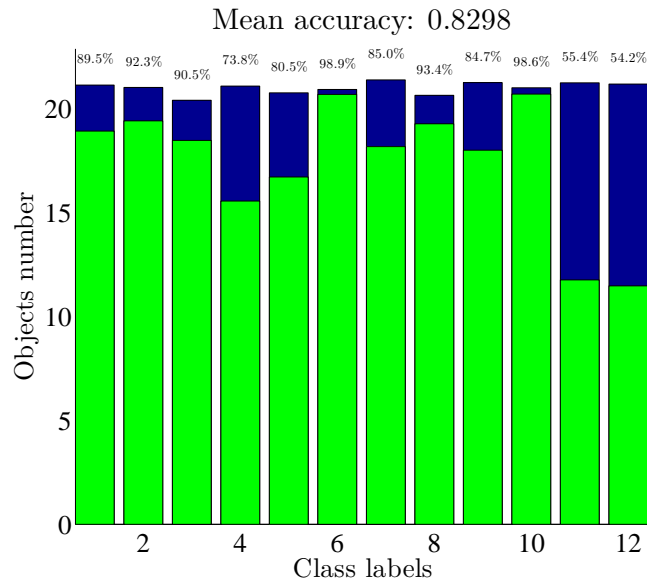


Рис. 3. Точность классификации при ручном выделении признаков. Dataset: USC-HAD. Под Mean accuracy понимается значение функционала 10. Над столбцами приведены средние точности классификации для каждого класса по формуле 9.

Модель авторегрессии 8

Признаковое описание. При записи данных USC-HAD сенсор делал каждую секунду 100 измерений. Предполагая, что на каждое «элементарное движение» человек тратит порядка секунды, приходим к выводу, что параметры авторегрессионной модели малых порядков в данном случае неинформативны. Приведем исходные временные ряды к частоте 10 Hz при помощи осреднения.

В качестве признаков преобразованных временных рядов возьмем статистические функции из 1, за исключением гистограммы. Так же для каждой компоненты отдельно и для модуля результирующего ускорения и поворота добавим по 11 параметров авторегрессионной модели AR(10) (см. 8). Затем проведем нормализацию признаков.

Классификатор. Задача многоклассовой классификации сводилась к задаче бинарной классификации при помощи подхода One-vs-All и экспоненциальной функцией потерь. В качестве бинарного классификатора использовался SVM с гауссовским ядром и параметрами $C = 16$, $\gamma = 0.1$.

Результаты. На диаграмме ниже приведено качество классификации, усредненное по $m = 200$ случайным разбиениям исходной выборки на тестовую и контрольную в отношении 7 к 3.

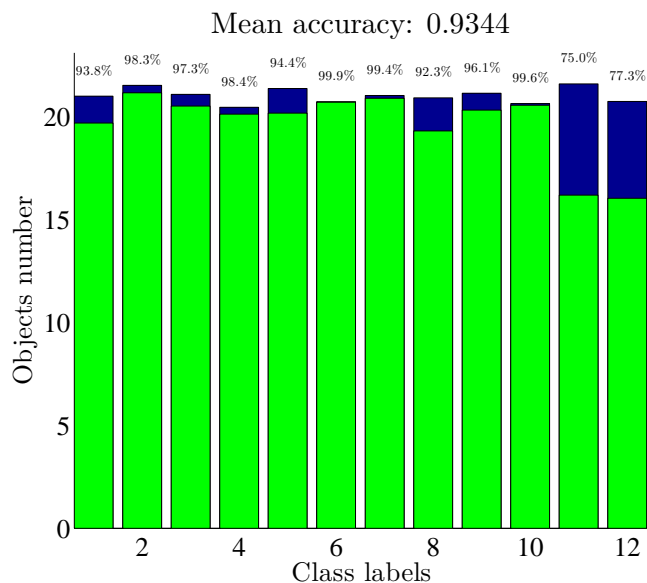


Рис. 4. Точность классификации для параметров модели авторегрессии в качестве признаков описаний. Dataset: USC-HAD. Под Mean accuracy понимается значение функционала 10. Над столбцами приведены средние точности классификации для каждого класса по формуле 9.

Алгоритм голосования

Будем решать задачу классификации только для первых 10 классов (за исключением «elevator up» и «elevator down», которые плохо отделяются друг от друга и от класса «stand») при помощи алгоритма голосования сегментов.

В качестве алгоритма сегментации используем разбиение исходного временного ряда на отрезки заданной длины.

Задача многоклассовой классификации решалась при помощи подхода One-vs-One с регуляризованной логистической регрессии (RLR) с параметром $C = 1$ в качестве бинарного классификатора.

На графике ниже приведены результаты для точности многоклассовой классификации, усредненные по $m = 50$ случайным разбиениям.

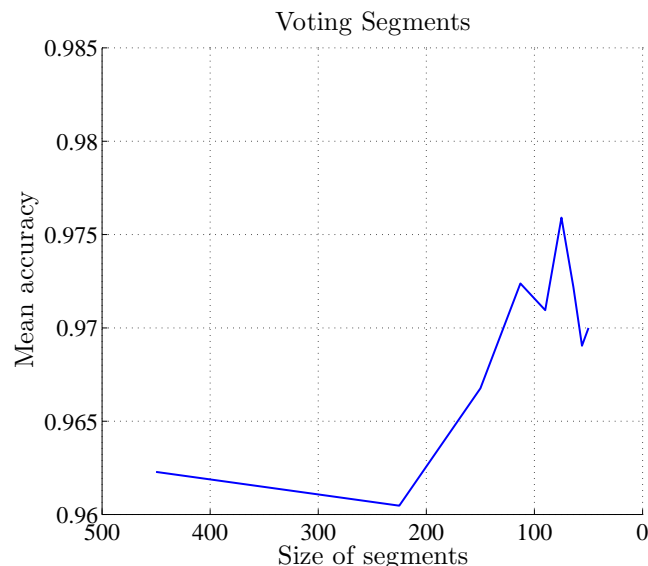


Рис. 5. Зависимость точности классификации от длины сегментов для алгоритма голосования. Dataset: USC-HAD. Под Mean accuracy понимается значение функционала 10.

Из графика можно видеть, что алгоритм голосования позволил повысить точность классификации на 1%.

Датасет MIT-BIH Arrhythmia Database [46]

Заключение

Метод классификации временных рядов в пространстве распределений параметров порождающих моделей обобщает большинство других методов классификации временных рядов и позволяет производить более тонкую настройку алгоритма классификации.

Литература

- [1] *Esling, P.* Time-series data mining / P. Esling, C. Agon // *ACM Comput. Surv.* — 2012. — December. — Vol. 45, no. 1. — Pp. 12:1–12:34. <http://doi.acm.org/10.1145/2379776.2379788>.
- [2] *Fu, T.-c.* A review on time series data mining / T.-c. Fu // *Engineering Applications of Artificial Intelligence.* — 2011. — Vol. 24, no. 1. — Pp. 164–181.
- [3] *Weigend, A. S.* Time series prediction: forecasting the future and understanding the past / A. S. Weigend // *Santa Fe Institute Studies in the Sciences of Complexity.* — 1994.
- [4] *Brockwell, P. J.* Time series: theory and methods / P. J. Brockwell, R. A. Davis. — Springer Science & Business Media, 2009.
- [5] *Tsay, R. S.* Analysis of financial time series / R. S. Tsay. — John Wiley & Sons, 2005. — Vol. 543.
- [6] *Weiss, G. M.* Mining with rarity: a unifying framework / G. M. Weiss // *ACM SIGKDD Explorations Newsletter.* — 2004. — Vol. 6, no. 1. — Pp. 7–19.
- [7] *Chin, S. C.* Symbolic time series analysis for anomaly detection: a comparative evaluation / S. C. Chin, A. Ray, V. Rajagopalan // *Signal Processing.* — 2005. — Vol. 85, no. 9. — Pp. 1859–1868.
- [8] *Yankov, D.* Disk aware discord discovery: finding unusual time series in terabyte sized datasets / D. Yankov, E. Keogh, U. Rebbapragada // *Knowledge and Information Systems.* — 2008. — Vol. 17, no. 2. — Pp. 241–262.

- [9] Segmenting time series: A survey and novel approach / E. Keogh, S. Chu, D. Hart, M. Pazzani // *Data mining in time series databases*. — 2004. — Vol. 57. — Pp. 1–22.
- [10] Geurts, P. Segment and combine approach for non-parametric time-series classification / P. Geurts, L. Wehenkel // *Knowledge Discovery in Databases: PKDD 2005*. — Springer, 2005. — Pp. 478–485.
- [11] Nunthanid, P. Parameter-free motif discovery for time series data / P. Nunthanid, V. Niennatrakul, C. A. Ratanamahatana // *Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON), 2012 9th International Conference on* / IEEE. — 2012. — Pp. 1–4.
- [12] Jain, A. K. Data clustering: A review / A. K. Jain, M. N. Murty, P. J. Flynn // *ACM Comput. Surv.* — 1999. — September. — Vol. 31, no. 3. — Pp. 264–323. <http://doi.acm.org/10.1145/331499.331504>.
- [13] Liao, T. W. Clustering of time series data—a survey / T. W. Liao // *Pattern Recognition*. — 2005. — Vol. 38, no. 11. — Pp. 1857–1874. <http://www.sciencedirect.com/science/article/pii/S0031320305001305>.
- [14] Zolhavarieh, S. A review of subsequence time series clustering / S. Zolhavarieh, S. Aghabozorgi, Y. W. Teh // *The Scientific World Journal*. — 2014. — Vol. 2014.
- [15] Bakshi, B. Representation of process trends—iv. induction of real-time patterns from operating data for diagnosis and supervisory control / B. Bakshi, G. Stephanopoulos // *Computers & Chemical Engineering*. — 1994. — Vol. 18, no. 4. — Pp. 303–332.
- [16] Geurts, P. Pattern extraction for time series classification / P. Geurts // *Principles of Data Mining and Knowledge Discovery*. — Springer, 2001. — Pp. 115–127.
- [17] Human activity recognition using smart phone embedded sensors: A linear dynamical systems method / W. Wang, H. Liu, L. Yu, F. Sun // *Neural Networks (IJCNN), 2014 International Joint Conference on* / IEEE. — 2014. — Pp. 1185–1190.
- [18] Wei, L. Semi-supervised time series classification / L. Wei, E. Keogh // *Proceedings of the 12th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. — KDD '06. — New York, NY, USA: ACM, 2006. — Pp. 748–753. <http://doi.acm.org/10.1145/1150402.1150498>.
- [19] Querying and mining of time series data: Experimental comparison of representations and distance measures / H. Ding, G. Trajcevski, P. Scheuermann et al. // *Proc. VLDB Endow.* — 2008. — August. — Vol. 1, no. 2. — Pp. 1542–1552. <http://dx.doi.org/10.14778/1454159.1454226>.
- [20] Berndt, D. J. Using dynamic time warping to find patterns in time series. / D. J. Berndt, J. Clifford // *KDD Workshop* / Ed. by U. M. Fayyad, R. Uthurusamy. — AAAI Press, 1994. — Pp. 359–370. <http://dblp.uni-trier.de/db/conf/kdd/kdd94.html#BerndtC94>.
- [21] A novel bit level time series representation with implication of similarity search and clustering / C. Ratanamahatana, E. Keogh, A. J. Bagnall, S. Lonardi // *Advances in knowledge discovery and data mining*. — Springer, 2005. — Pp. 771–777.
- [22] Salvador, S. Toward accurate dynamic time warping in linear time and space / S. Salvador, P. Chan // *Intelligent Data Analysis*. — 2007. — Vol. 11, no. 5. — Pp. 561–580.
- [23] Marteau, P.-F. Time warp edit distance with stiffness adjustment for time series matching / P.-F. Marteau // *Pattern Analysis and Machine Intelligence, IEEE Transactions on*. — 2009. — Vol. 31, no. 2. — Pp. 306–318.
- [24] Shatkay, H. Approximate queries and representations for large data sequences / H. Shatkay, S. Zdonik // *Data Engineering, 1996. Proceedings of the Twelfth International Conference on*. — 1996. — Feb. — Pp. 536–545.

- [25] *Li, C.-S.* Malm: a framework for mining sequence database at multiple abstraction levels / C.-S. Li, P. S. Yu, V. Castelli // *Proceedings of the seventh international conference on Information and knowledge management / ACM.* — 1998. — Pp. 267–272.
- [26] *Vasko, K.* Estimating the number of segments in time series data using permutation tests / K. Vasko, H. Toivonen // *Data Mining, 2002. ICDM 2003. Proceedings. 2002 IEEE International Conference on.* — 2002. — Pp. 466–473.
- [27] Streaming time series summarization using user-defined amnesic functions / T. Palpanas, M. Vlachos, E. Keogh, D. Gunopulos // *Knowledge and Data Engineering, IEEE Transactions on.* — 2008. — July. — Vol. 20, no. 7. — Pp. 992–1006.
- [28] *Fröhwrth-Schnatter, S.* Model-based clustering of multiple time series / S. Fröhwrth-Schnatter, S. Kaufmann // *Journal of Business & Economic Statistics.* — 2008. — Vol. 26, no. 1. — Pp. 78–89.
- [29] *Corduas, M.* Time series clustering and classification by the autoregressive metric / M. Corduas, D. Piccolo // *Computational Statistics & Data Analysis.* — 2008. — Vol. 52, no. 4. — Pp. 1860 – 1872. <http://www.sciencedirect.com/science/article/pii/S0167947307002368>.
- [30] *Cormode, G.* Conquering the divide: Continuous clustering of distributed data streams / G. Cormode, S. Muthukrishnan, W. Zhuang // *Data Engineering, 2007. ICDE 2007. IEEE 23rd International Conference on.* — 2007. — April. — Pp. 1036–1045.
- [31] *Bortolan, G.* Diagnostic ecg classification based on neural networks / G. Bortolan, J. Willems // *Journal of electrocardiology.* — 1993. — Vol. 26 Suppl. — P. 75–79. <http://europepmc.org/abstract/MED/8189152>.
- [32] Finding unusual medical time-series subsequences: Algorithms and applications / E. Keogh, J. Lin, A. W. Fu, H. Van Herle // *Information Technology in Biomedicine, IEEE Transactions on.* — 2006. — Vol. 10, no. 3. — Pp. 429–439.
- [33] *Marcel, S.* Person authentication using brainwaves (eeg) and maximum a posteriori model adaptation / S. Marcel, J. d. R. Millán // *Pattern Analysis and Machine Intelligence, IEEE Transactions on.* — 2007. — Vol. 29, no. 4. — Pp. 743–752.
- [34] Human activity recognition using smart phone embedded sensors: A linear dynamical systems method / W. Wang, H. Liu, L. Yu, F. Sun // *Neural Networks (IJCNN), 2014 International Joint Conference on.* — 2014. — July. — Pp. 1185–1190.
- [35] *Kwapisz, J. R.* Activity recognition using cell phone accelerometers / J. R. Kwapisz, G. M. Weiss, S. A. Moore // *SIGKDD Explor. Newsl.* — 2011. — March. — Vol. 12, no. 2. — Pp. 74–82. <http://doi.acm.org/10.1145/1964897.1964918>.
- [36] *Martens, R.* On-line signature verification by dynamic time-warping / R. Martens, L. Claesen // *Pattern Recognition, 1996., Proceedings of the 13th International Conference on.* — Vol. 3. — 1996. — Pp. 38–42 vol.3. <http://dx.doi.org/10.1109/ICPR.1996.546791>.
- [37] *Gruber, C.* Signature verification with dynamic rbf networks and time series motifs / C. Gruber, M. Coduro, B. Sick // *Tenth International Workshop on Frontiers in Handwriting Recognition / Suvisoft.* — 2006.
- [38] Pattern recognition and classification for multivariate time series / S. Spiegel, J. Gaebler, A. Lommatzsch et al. // *Proceedings of the Fifth International Workshop on Knowledge Discovery from Sensor Data. — SensorKDD '11. — New York, NY, USA: ACM, 2011. — Pp. 34–42. <http://doi.acm.org/10.1145/2003653.2003657>.*
- [39] *Nguyen, M. N.* Positive unlabeled learning for time series classification. / M. N. Nguyen, X.-L. Li, S.-K. Ng // *IJCAI / Citeseer.* — Vol. 11. — 2011. — Pp. 1421–1426.
- [40] *Marussy, K.* Success: a new approach for semi-supervised classification of time-series / K. Marussy, K. Buza // *Artificial Intelligence and Soft Computing / Springer.* — 2013. — Pp. 437–447.

- [41] *Nanopoulos, A.* Feature-based classification of time-series data / A. Nanopoulos, R. Alcock, Y. Manolopoulos // *International Journal of Computer Research*. — 2001. — Vol. 10. — Pp. 49–61.
- [42] *Wiens, J.* Patient risk stratification for hospital-associated c. diff as a time-series classification task / J. Wiens, E. Horvitz, J. V. Guttag // *Advances in Neural Information Processing Systems*. — 2012. — Pp. 467–475.
- [43] *Mörchen, F.* Time series feature extraction for data mining using dwt and dft. — 2003.
- [44] *Zhang, H.* A non-parametric wavelet feature extractor for time series classification / H. Zhang, T. B. Ho, M. S. Lin // *Advances in Knowledge Discovery and Data Mining*. — Springer, 2004. — Pp. 595–603.
- [45] *Zhang, M.* Usc-had: A daily activity dataset for ubiquitous activity recognition using wearable sensors / M. Zhang, A. A. Sawchuk // *ACM International Conference on Ubiquitous Computing (Ubicomp) Workshop on Situation, Activity and Goal Awareness (SAGAware)*. — Pittsburgh, Pennsylvania, USA: 2012. — September.
- [46] Physiobank, physiotoolkit, and physionet: Components of a new research resource for complex physiologic signals / A. L. Goldberger, L. A. N. Amaral, L. Glass et al. // *Circulation*. — 2000. — Vol. 101, no. 23. — Pp. e215–e220. <http://circ.ahajournals.org/content/101/23/e215.abstract>.