

# Классификация временных рядов в пространстве параметров порождающих моделей

Карасиков Михаил

Московский физико-технический институт  
Факультет управления и прикладной математики  
Кафедра интеллектуальных систем

*Научный руководитель:* д.ф.-м.н. В. В. Стрижов

Москва, 2015 г.

## Исследуется

Задача построения пространства признаков для многоклассовой классификации временных рядов

## Проблема

Построение краткого интерпретируемого признакового описания временных рядов

## Цели исследования:

- построение алгоритма многоклассовой классификации, использующего в качестве признаков временных рядов параметры моделей временных рядов и их распределения,
- обобщение методов классификации временных рядов, использующих явное признаковое описание,
- повышение качества решения задач классификации временных рядов.

- Time-series data mining / P. Esling, C. Agon // ACM Comput. Surv. — 2012.— December.— Vol. 45, no. 1.— Pp. 12:1–12:34.
- 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.
- 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>.

**Дано:**  $X = \{x_i = [x_i^{(1)}, \dots, x_i^{(t)}] : i \in \mathcal{I}\}$  — временные ряды,  
 $Y$  — множество меток классов,  
 $\mathcal{D} \subset X \times Y$  — обучающая выборка.

**Модель классификации:**  $a = b \circ f \circ S$ , где

$S$  — алгоритм сегментации  $S : x \mapsto \{s^{(1)}, \dots, s^{(p)}\}$ ,

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

$f$  — признакововое описание набора сегментов,

$b$  — алгоритм многоклассовой классификации.

**Метод обучения**  $\mu : (X \times Y)^m \rightarrow A$

выбирается по скользящему контролю:

$$\mu^* = \arg \min_{\mu} CV(\mu, \mathcal{D}).$$

Каждый сегмент есть временной ряд:  $\mathbf{S}(x) \subset X$ .

Сегменты временных рядов описывается моделями вида

$$g(\mathbf{w}, x) \in X, \text{ где } \mathbf{w} \in \mathbb{R}^d.$$

Параметры настроенной модели определяются по формуле

$$\mathbf{w}_g(x) = \arg \min_{\mathbf{w} \in \mathbb{R}^d} \rho(g(\mathbf{w}, x), x).$$

### Исследуются

- Линейная регрессионная модель
- Авторегрессионная модель (AR)
- Модель скользящего среднего (MA)
- Фурье-модель
- Вейвлет-модель

Предлагаются две схемы решения исходной задачи.

- Принцип голосования: обучение алгоритма  $b$  на новой обучающей выборке  $\hat{\mathfrak{D}}$ , составленной из сегментов временных рядов исходной обучающей выборки  $\mathfrak{D}$ :

$$\hat{\mathfrak{D}} = \{(\mathbf{w}_g(s), y) : (x, y) \in \mathfrak{D}, s \in \mathbf{S}(x)\}$$

и последующая классификация

$$a(x; \mathbf{S}, g, b) = h(\{b(\mathbf{w}_g(s)) : s \in \mathbf{S}(x)\}).$$

- Классификация в пространстве гиперпараметров моделей (параметров распределений параметров моделей).

$\mathbf{w}_g \circ \mathbf{S}$  дает множество наборов параметров модели:

$$W(x; \mathbf{S}, g) = \{\mathbf{w}_g(s) : s \in \mathbf{S}(x)\}.$$

## Гипотеза порождения временного ряда

Сегменты временного ряда  $s \in \mathbf{S}(x)$  описываются моделью  $g(\mathbf{w}, s)$  со случайными параметрами  $\mathbf{w}$  из параметрического семейства распределений  $\{P_{\theta}\}_{\theta \in \Theta}$ .

**Предлагается** в качестве признакового описания временного ряда использовать оценку вектора параметров распределения:

$$\mathbf{f}(x; \mathbf{S}, g, \Theta) = \arg \max_{\theta \in \Theta} \mathcal{L}(\theta | W(x; \mathbf{S}, g)).$$

Тогда получим алгоритм классификации временных рядов:

$$a(x) = b(\mathbf{f}(x; \mathbf{S}, g, \Theta)).$$

- One-vs-All approach:

$$a(x) = \arg \max_{i=1,\dots,N} f_i(x), \quad f_i(x) = \begin{cases} \geq 0, & \text{если } y(x) = i, \\ < 0, & \text{если } y(x) \neq i. \end{cases}$$

- One-vs-One approach:

$$a(x) = \arg \max_{i=1,\dots,N} \sum_{\substack{j=1,\dots,N \\ j \neq i}} f_{ij}(x), \quad f_{ij}(x) = \begin{cases} +1, & \text{если } y(x) = i, \\ -1, & \text{если } y(x) = j. \end{cases}$$

- Error-Correcting Output Codes approach:

$$a(x) = \arg \min_{i=1,\dots,N} \sum_{j=1}^F L(M_j^i f_j(x)),$$

где  $M \in \{-1, 0, +1\}^{N \times F}$  — матрица, строки которой состоят из кодов меток классов  $Y$ , а  $L$  — функция потерь.



## Бинарная классификация: $Y = \{\pm 1\}$

Пусть  $\mathbf{f}(x) \in \mathbb{R}^n$  — признаковое описание временного ряда  $x$ .

Тогда для решения задачи бинарной классификации временных рядов необходимо задать метод обучения  $\mu_b : \mathfrak{D} \mapsto f$ .

Например,

- SVM с линейным ядром:

$$f(x; \mathbf{w}) = \text{sign}(\mathbf{w}^T \mathbf{f}(x) - w_0),$$

где  $\mathbf{w}$  и  $w_0$  — решения оптимизационной задачи

$$\frac{1}{2C} \|\mathbf{w}\|^2 + \sum_{(x,y) \in \mathfrak{D}} \left(1 - y(\mathbf{w}^T \mathbf{f}(x) - w_0)\right)_+ \rightarrow \min_{\mathbf{w} \in \mathbb{R}^n, w_0 \in \mathbb{R}};$$

- регуляризованная лог. регрессия (RLR):

$$\frac{1}{2C} \|\mathbf{w}\|^2 + \sum_{(x,y) \in \mathfrak{D}} \log \left(1 + e^{-\mathbf{w}^T \mathbf{f}(x)}\right) \rightarrow \min_{\mathbf{w} \in \mathbb{R}^n}.$$

В качестве приложения рассматривается задача классификации физической активности по данным с акселерометра.

## Особенности

- Классификация физической активности людей с разными физическими характеристиками
- Форма временного ряда существенно зависит от характеристик человека
- Во временных рядах допускаются аномалии

## Предположение

Форма временного ряда сохраняются для конкретного человека и типа физической активности

## Цели эксперимента:

- 1 демонстрация качества предлагаемого алгоритма
- 2 изучение зависимости качества классификации от
  - алгоритма сегментации
  - модели сегмента

Временной ряд из трех компонент:  $\mathbf{x} = [x_t^k]^{k=1,2,3}$ .

## Признаки (по 31 на каждый временной ряд)

1 7 коэффициентов авторегрессии  $AR(6)$  :

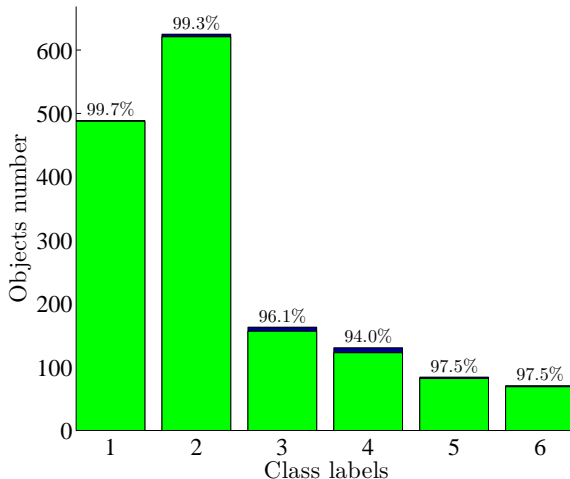
$$\arg \min_{w_0, \dots, w_6} \sum_t \left( x_t^k - w_0 - \sum_{i=1}^6 w_i x_{t-i}^k \right)^2.$$

2 Статистики:

- $\bar{x}^k = \frac{1}{T} \sum_t x_t^k,$
- $\frac{1}{T} \sum_t |x_t^k - \bar{x}^k|,$
- $\sqrt{\frac{1}{T-1} \sum_t (x_t^k - \bar{x}^k)^2},$
- $\frac{1}{T} \sum_t \|\mathbf{x}_t\|.$

Без сегментации, классификатор RBF SVM ( $\gamma = 0.8$ ,  $C = 4$ ),  
подход One-vs-All, 50 случайных разбиений в отношении 7 к 3.

Mean accuracy: 0.9849



Class labels:

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

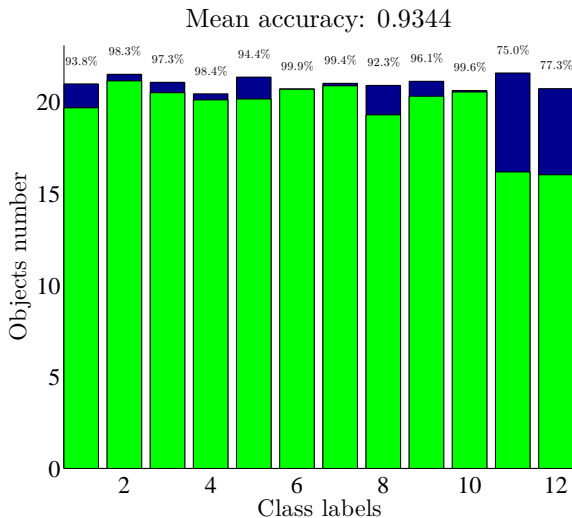
**Выборка содержит:** показания акселерометра и гироскопа для 12 типов физической активности.

Частота исходной выборки 100 Hz.

Приведем данные к частоте 10 Hz.

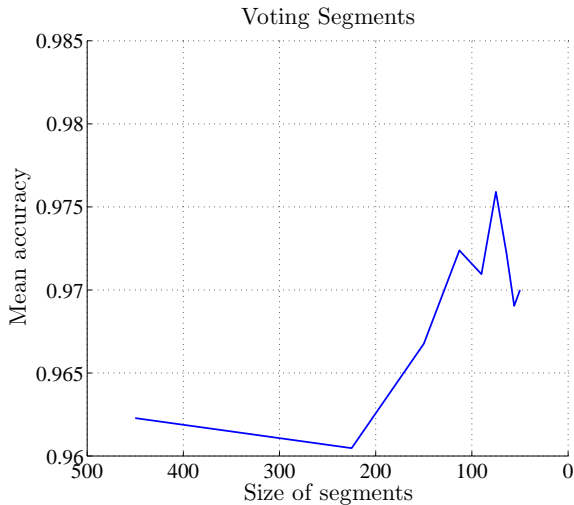
Будем решать задачу многоклассовой классификации, классификатор RBF SVM ( $\gamma = 0.1$ ,  $C = 16$ ), подход One-vs-All, 200 случайных разбиений в отношении 7 к 3.

В качестве признаков берутся коэффициенты модели авторегрессии.



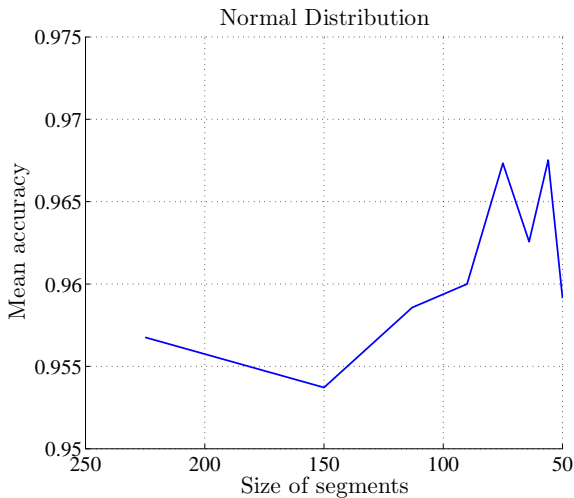
## Class labels:

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

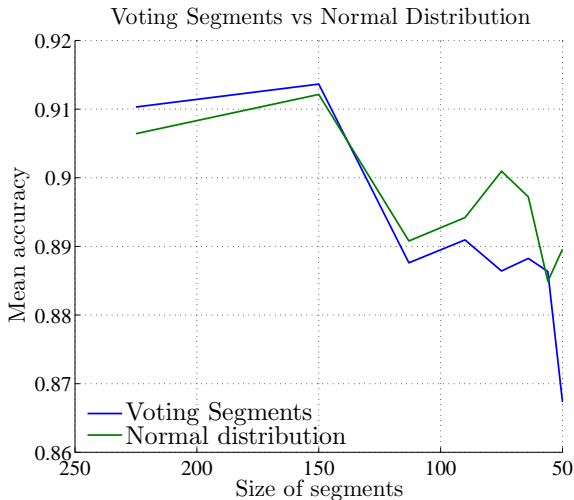


**Рис.:** Эксперимент на первых 10 классах. One-vs-One, RLR ( $C = 1$ ).





**Рис.:** На первых 10 классах. One-vs-One, RLR ( $C = 0.25$ ).



**Рис.:** Сравнение на всей выборке (12 классов).

- Метод классификации временных рядов в пространстве параметров моделей показывает высокие результаты.
- Использование модели авторегрессии не накладывает ограничений на алгоритм сегментации.
- Алгоритм голосования менее устойчив к изменению длине сегмента.

- Предложен алгоритм построения пространства признаков.
- Предложены алгоритмы классификации временных рядов
- Выполнена программная реализация и проведены численные эксперименты, показавшие повышение качества решения задачи классификации типов физической активности.