# 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.

# Описание исследования

## Исследуется

Задача построения пространства признаков в задаче многоклассовой классификации 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.

# Литература


- 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:** $X^\ell = \{x_1, \dots, x_\ell\} \subset X$ — временные ряды,
$Y$ — множество меток классов,
$\mathfrak{D} \subset X^\ell \times Y$ — обучающая выборка.

**Модель алгоритма классификации:** $a = b \circ \mathbf{f} \circ S$, где
$S$ — алгоритм фрагментации,
$\mathbf{f}$ — признаковое описание набора фрагментов,
$b$ — алгоритм многоклассовой классификации.

**Метод обучения** $\mu : \ (X \times Y)^m \to A$
выбирается по скользящему контролю:

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

# Разбиение временного ряда на фрагменты

## Определения

- Временной ряд: $x = [x^{(1)}, \dots, x^{(t)}] \in X$.
- Фрагмент временного ряда: $s = [x^{(i_1)}, \dots, x^{(i_k)}]$,
  где $1 \leqslant i_1 \leqslant \ldots \leqslant i_k \leqslant t$.

Пусть $\mathbf{S}(x)$ — множество всех фрагментов временного ряда $x$.
Тогда алгоритм фрагментации есть отображение $S : x \mapsto \mathbf{S}(x)$.

**Примеры:**

- тождественное отображение

$$S: \ x \mapsto \{x\},$$

- сегментация

$$S: \ x \mapsto \{s^{(1)}, \dots, s^{(p)}\}, \text{ где } x = (s^{(1)}, \dots, s^{(p)}).$$

# Модель фрагмента временного ряда

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

Фрагменты временных рядов описывается моделями вида

$$g : \mathbb{R}^d \times X \to X.$$

Параметры настроенной модели определяются по формуле

$$\mathbf{w}_g(x) = \operatorname*{arg\,min}_{\mathbf{w} \in \mathbb{R}^d} \rho\left(g(\mathbf{w}, x), x\right).$$

## Примеры.

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

# Параметры моделей фрагментов

Предлагаются две схемы решения исходной задачи.

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

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

  и последующая классификация

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

- Классификация в пространстве параметров распределений параметров моделей.

# Классификация в пространстве параметров распределений параметров моделей

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

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

### Гипотеза порождения временного ряда

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

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

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

Тогда получим алгоритм классификации временных рядов:

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

## Сведение задачи многоклассовой классификации к бинарным

- One-vs-All approach:

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

- One-vs-One approach:

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

- Error-Correcting Output Codes approach:

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

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, для SVM с линейным ядром

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

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

$$\frac{1}{2C}\|\boldsymbol{w}\|^2 + \sum_{(x,y)\in\mathfrak{D}} \Big(1 - y(\boldsymbol{w}^\mathsf{T}\mathbf{f}(x) - w_0)\Big)_+ \to \min_{\boldsymbol{w}\in\mathbb{R}^n,\, w_0\in\mathbb{R}}.$$

# Приложение

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

# Вычислительный эксперимент

**Цели эксперимента:**

1. демонстрация качества предлагаемого алгоритма
2. изучение зависимости качества классификации от
   - алгоритма фрагментации
   - модели фрагмента

# Dataset WISDM

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

## Признаки (по 31 на каждый временной ряд)

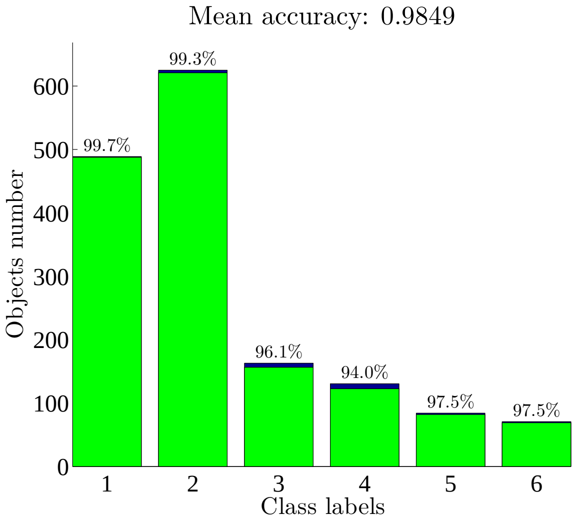
1. 7 коэффициентов авторегрессии $AR(6)$ :

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

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




**Class labels:**

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

# Dataset USC-HAD

# Dataset USC-HAD: сегментация

# Dataset USC-HAD: голосование

# Dataset USC-HAD: распределения

# Заключение

# Результаты, выносимые на защиту

- Предложен алгоритмы построения пространства признаков
- Предложен алгоритм классификации временных рядов
- Выполнена программная реализация и проведены численные эксперименты, показавшие повышения качества решения задачи классификации