

Локальные модели для классификации объектов сложной структуры

Исаченко Роман, Жариков Илья,
Бочкарёв Артём

Научный руководитель: Стрижов В. В.
Московский физико-технический институт
Математические методы распознавания образов

Таганрог
10 октября, 2017

Задача

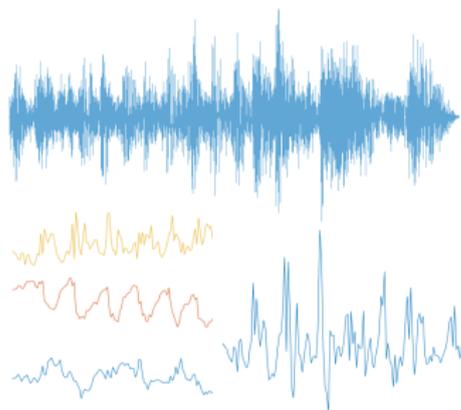
Построить модель машинного обучения для объектов сложной структуры.

Проблема

Исходные объекты не имеют явного признакового описания

Применение:

- обработка сигналов;
- классификация сигналов;
- *анализ временных рядов.*



- 1 Wang W. et al. Human activity recognition using smart phone embedded sensors // *International Joint Conference on Neural Networks*. 2014.
- 2 Kwapisz, Jennifer R., Gary M. Weiss, and Samuel A. Moore. Activity recognition using cell phone accelerometers. // *ACM SigKDD Explorations Newsletter*. 12(2): 74-82. 2011.
- 3 Кузнецов М. П., Ивкин Н. П. Алгоритм классификации временных рядов акселерометра по комбинированному признаковому описанию // *Машинное обучение и анализ данных*. 1(11): 1471-1483. 2015.
- 4 Karasikov M. E., Strijov V. V. Feature-based time-series classification // *Intelligence*. 24(1): 164-181. 2016.
- 5 Motrenko, A. P., Strijov V. V. Extracting fundamental periods to segment human motion time series // *Journal of Biomedical and Health Informatics*. 2015.

Постановка задачи

Заданы $s \in \mathcal{S}$ — объект сложной структуры;
 $y \in Y$ — метка класса;
 $\mathcal{D} = \{(s_i, y_i)\}_{i=1}^m$ — размеченная выборка.

Задача

Определить функцию f^* такую что

$$f^* = \arg \min_f L(f, \mathcal{D}),$$

где $L(\cdot, \cdot)$ — функция ошибки и $f : \mathcal{S} \rightarrow Y$.

Предположение

Предположим что $f = g \circ h$, где

- 1 $h(s) : \mathcal{S} \rightarrow H \subset \mathbb{R}^n$ отображение из \mathcal{S} в признаковое пространство H ;
- 2 $g(\mathbf{h}, \theta) : H \rightarrow Y$ параметрическое отображение (модель классификации).

$h(s)$

Выбор функции $h(s)$ обусловлен

- априорными (экспертными) знаниями;
- *минимизацией функционала ошибки.*

$g(\mathbf{h}, \theta)$

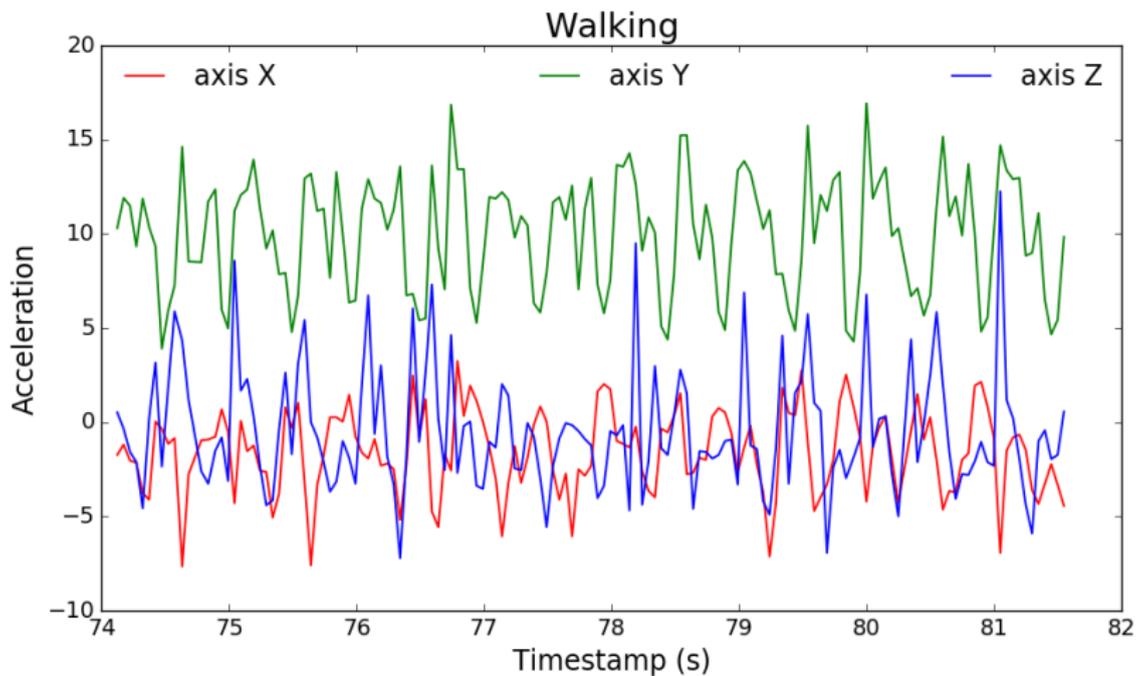
Классификация $\{(\mathbf{h}_i, y_i)\}_{i=1}^m$, $\mathbf{h}_i = h(s_i)$:

$$\theta^* = \arg \min_{\theta} L(g, \theta, \mathcal{D}),$$

- $g(\mathbf{h}, \theta)$ - модель классификации;
- θ - параметры модели;
- $L(g, \theta, \mathcal{D})$ - функция ошибки классификации.

$$s \xrightarrow{h(\cdot)} \mathbf{h} \xrightarrow{g(\cdot, \theta)} y$$

Пример временного ряда



Априорное знание об объектах позволяет извлечь признаки.

Признаковое описание

$\mathbf{h}_i = h(s_i) \in \mathbb{R}^{40}$ — различные статистики:

- 3 среднее ускорение по каждой оси;
- 3 стандартное отклонение;
- 3 среднее абсолютное отклонение;
- 1 среднее ускорение;
- 30 значение гистограммы для каждой оси.

Проблема: необходимы экспертные знания.

Гипотеза порождения данных

Предположим, что $s = (x_1, \dots, x_T)$ порождён с помощью авторегрессионной модели:

$$\hat{x}_t = w_0 + \sum_{j=1}^n w_j x_{t-j}.$$

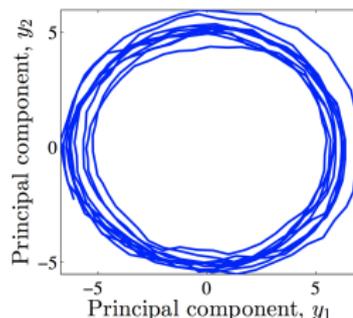
Признаковое описание

$$h(s) = \mathbf{w}^* = \arg \min_{\mathbf{w} \in \mathbb{R}^{n+1}} \sum_{j=n+1}^T \|x_j - \hat{x}_j\|^2.$$

Анализ сингулярного спектра (SSA)

Траекторная матрица временного ряда $s = (x_1, \dots, x_T)$:

$$\mathbf{X} = \begin{pmatrix} x_1 & x_2 & \dots & x_n \\ x_2 & x_3 & \dots & x_{n+1} \\ \vdots & \vdots & \ddots & \vdots \\ x_{T-n+1} & x_{T-n+2} & \dots & x_T \end{pmatrix}$$

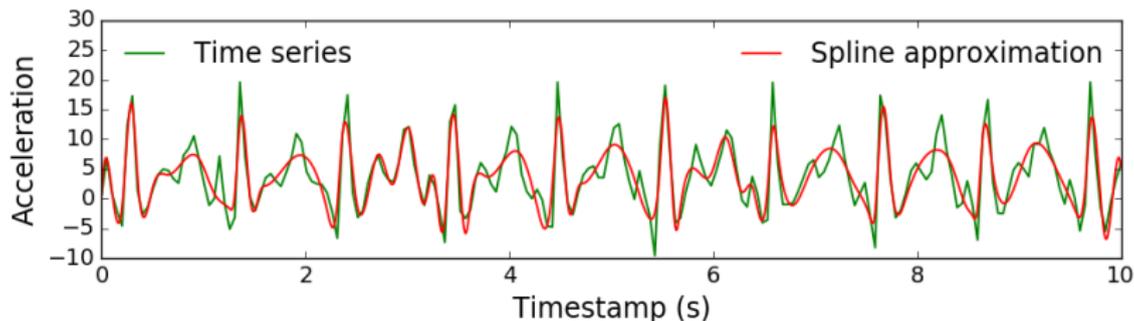


Признаковое описание

$$h(s) = (\lambda_1, \dots, \lambda_n),$$

где $\{\lambda_i\}_{i=1}^n$ — собственные значения матрицы $\mathbf{X}^T \mathbf{X}$, полученные с помощью SVD: $\mathbf{X}^T \mathbf{X} = \mathbf{V} \cdot \text{diag}(\lambda_1, \dots, \lambda_n) \cdot \mathbf{V}^T$.

Кубические сплайны



Сплайн определяется $\{\xi_\ell\}_{\ell=1}^L$ — набором узлов и $\{\mathbf{w}_\ell\}_{\ell=1}^L$ — параметрами моделей, построенными на отрезках $[\xi_\ell; \xi_{\ell+1}]$.

Признаковое описание

$$\mathbf{h} = h(s) = (\xi_1, \dots, \xi_L, \mathbf{w}_1, \dots, \mathbf{w}_L).$$

WISDM¹

Activity	# objects	
Standing	229	5.3%
Walking	1917	44.4%
Upstairs	466	10.8%
Sitting	277	6.4%
Jogging	1075	24.9%
Downstairs	357	8.3%
Total	4321	

USC-HAD²

Activity	# objects	
Walking-downstairs	951	7%
Walking-upstairs	1018	7.4%
Walking-forward	1874	13.8%
Walking-right	1305	9.6%
Walking-left	1280	9.4%
Elevator-up	764	5.6%
Elevator-down	763	5.6%
Standing	1167	8.6%
Sitting	1294	9.5%
Sleeping	1860	13.7%
Jumping	495	3.6%
Running	849	6.2%
Total	13620	

¹<http://www.cis.fordham.edu/wisdm/>

²<http://sipi.usc.edu/HAD/>

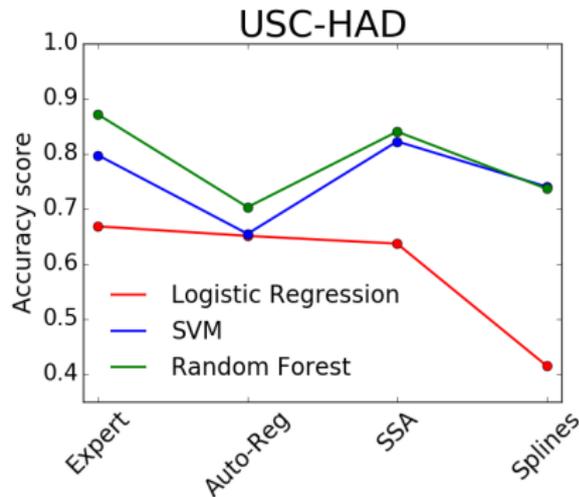
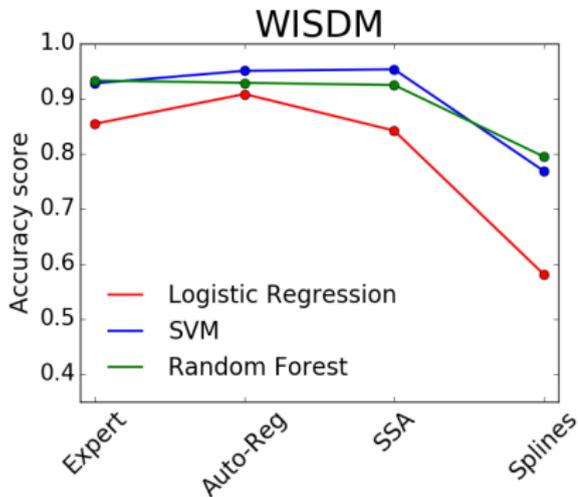
Данные: WISDM, USC-HAD.

Методы генерации признаков: экспертные функции;
авторегрессионная модель;
анализ сингулярного спектра;
сплайны.

Модели классификации: логистическая регрессия;
SVM;
случайный лес.

Оптимизация параметров: кросс-валидация.

Метрика качества: точность.



Объединение всех признаков.

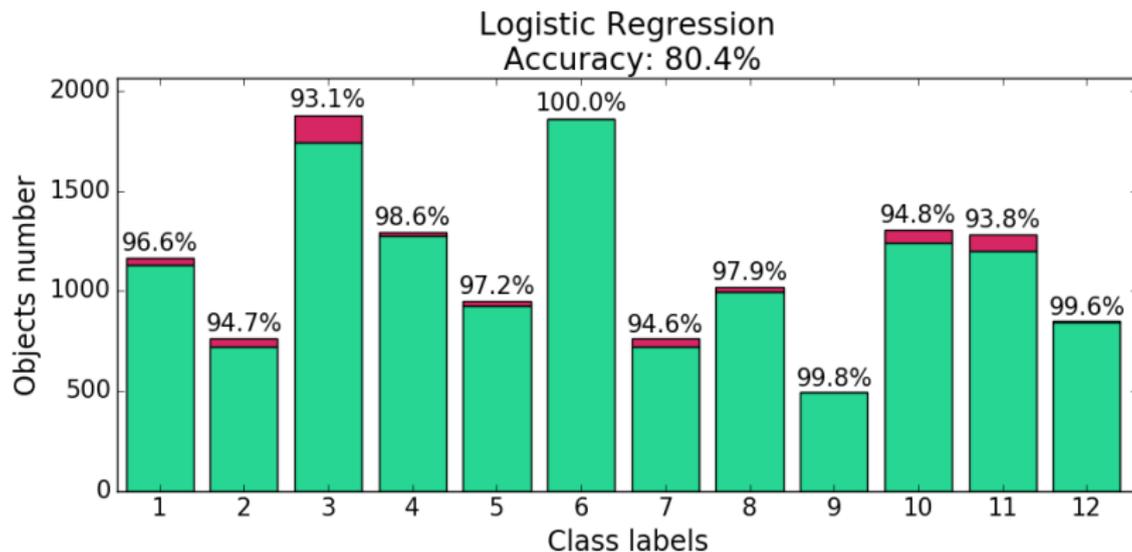
SVM: **0.983**

Случайный лес: **0.882**

USC-HAD

	Logistic Regression				Random Forest				SVM			
	EX	AR	SSA	SPL	EX	AR	SSA	SPL	EX	AR	SSA	SPL
All	0.67	0.65	0.64	0.41	0.87	0.70	0.84	0.74	0.80	0.65	0.82	0.74
Standing	0.94	0.94	0.92	0.89	0.98	0.94	0.97	0.98	0.95	0.94	0.97	0.96
Elevator-up	0.94	0.94	0.93	0.92	0.95	0.95	0.95	0.95	0.93	0.94	0.94	0.93
Walking-forward	0.87	0.87	0.89	0.70	0.97	0.89	0.96	0.88	0.95	0.87	0.97	0.91
Sitting	0.98	0.95	0.94	0.96	0.99	0.96	0.98	0.99	0.98	0.96	0.99	0.99
Walking-downstairs	0.95	0.93	0.93	0.90	0.99	0.96	0.98	0.95	0.98	0.93	0.98	0.96
Sleeping	1.00	0.98	0.99	1.00	1.00	0.98	1.00	1.00	1.00	0.98	1.00	1.00
Elevator-down	0.94	0.94	0.94	0.91	0.95	0.95	0.95	0.95	0.93	0.94	0.94	0.93
Walking-upstairs	0.94	0.95	0.93	0.92	0.98	0.95	0.98	0.96	0.98	0.95	0.98	0.96
Jumping	0.99	0.99	1.00	0.97	1.00	0.99	1.00	0.99	1.00	0.99	0.97	0.99
Walking-right	0.91	0.90	0.91	0.86	0.97	0.92	0.96	0.92	0.96	0.90	0.97	0.93
Walking-left	0.89	0.91	0.90	0.88	0.97	0.93	0.97	0.93	0.95	0.91	0.97	0.93
Running	0.99	0.99	0.99	0.92	1.00	0.99	1.00	0.97	1.00	1.00	0.95	0.98

USC-HAD, объединение признаков



- задача классификации объектов сложной структуры может быть декомпозирована на две процедуры: извлечения признаков и обучение модели;
- параметры локальных моделей могут быть использованы в качестве признакового описания для объектов сложной структуры;
- вычислительные эксперименты на реальных данных акселерометра подтверждают гипотезу о применимости подхода.