

Порождение признаков в задаче классификации физической активности

Исаченко Роман^{1,2}, Жариков Илья^{1,2},
Бочкарёв Артём^{1,2}, Стрижов Вадим¹

¹ Московский физико-технический институт, Институтский пер. 9,
Долгопрудный, 141700, Россия

² Сколковский институт науки и технологий, ул. Нобеля 3, Москва, 143026,
Россия

Аннотация Данная работа посвящена предсказанию физической активности человека по сигналам акселерометра мобильного телефона. Сигнал акселерометра представляет собой временной ряд. В связи с большой размерностью описания объектов —временных рядов и недостатком вычислительных мощностей возникает задача порождения признакового пространства. Авторы предлагают использовать параметры локальных моделей в качестве набора признаков. Эксперимент проводился на реальных данных акселерометра мобильного телефона: WISDM и USD-HAD. В работе произведён анализ различных суперпозиций моделей порождения признаков и классификации.

Ключевые слова: *акселерометр, временные ряды, локальные модели, классификация*

1 Введение

В работе рассматривается задача многоклассовой классификации объектов, которые не имеют явного признакового представления. Данная задача возникает при анализе биологических данных [17], человеческого поведения и социальных взаимоотношений [1]. Временные ряды акселерометра [8,15,20] мобильного телефона позволяют распознать вид физической активности человека, используя концепцию Интернета Вещей [2,14]. Данная задача решалась различными методами от топологического анализа данных [18] до сверточных нейронных сетей [6]. Обзор известных методов и данных для данной задачи представлен в [9].

Рассматриваемые данные содержат временные ряды ускорения по трём осям, полученные со смартфона с акселерометром. Такие временные ряды имеют разные длины, являются не выравненными и мультишкалированными [4]. Задача состоит в предсказании вида физической активности человека. Список активностей включает в себя бег, ходьбу, сидение, стояние и тд. При такой постановке временной ряд представляет собой объект сложной структуры без явного признакового представления. Данное предположение позволяет использовать технологии моделирования временных

рядов акселерометра. Основная проблема состоит в нехватке вычислительных ресурсов и памяти смартфонов. Данное исследование предлагает подход к генерации признаков временных рядов, рассматривая их как объекты сложной структуры. Сгенерированные признаки позволяют получить приемлемое качество классификации и требуют умеренных ресурсов.

Проблема классификации объектов сложной структуры разбивается на две несвязанные процедуры. Первая извлекает информативные признаки. Вторая классифицирует объекты, используя порожденные признаки.

Данная работа нацелена на сравнение различных методов генерации признаков: [11,12]: экспертные функции, авторегрессионная модель and анализ сингулярного спектра. Экспертные функции [13] включают в себя среднее ускорение, стандартное отклонение, среднее абсолютное отклонение и значения гистограммы распределения. Авторегрессионная модель [16] строит параметрическую функцию для каждого временного ряда и использует параметры функции в качестве признаков для классификации. Анализ сингулярного спектра [7] использует собственные значения траекторной матрицы в качестве порождаемых признаков.

Авторы предлагают новый метод порождения признаков. Мы аппроксимируем сегменты временного ряда кубическими сплайнами [3]. Аппроксимация сплайнами — это кусочная кривая третьего порядка с заданными узлами. Условия дополнительной гладкости делают кривую, её первую и вторую производные непрерывными. Сплайны генерируют гладкую кривую с достаточным качеством аппроксимации.

Эксперимент проводился на двух датасетов с акселерометра: WISDM [21], USC-HAD [19]. Сравнивались качества упомянутых методов извлечения признаков и различных моделей классификации.

2 Problem Statement

Временные ряды акселерометра образуют множество \mathcal{S} сегментов s фиксированной длины T :

$$s = [x_1, \dots, x_T]^T \in \mathbb{R}^T. \quad (1)$$

Необходимо построить модель классификации $f : \mathbb{R}^T \rightarrow Y$, которая будет ставить в соответствие каждому сегменту из множества \mathcal{S} метку класса из конечного множества Y . Обозначим за

$$\mathcal{D} = \{(s_i, y_i)\}_{i=1}^m \quad (2)$$

исходная выборка, где $s_i \in \mathcal{S}$ и $y_i = f(s_i) \in Y$.

Авторы предлагают построить модель f в виде суперпозиции $f = f(g)$. Функция $g : \mathbb{R}^T \rightarrow \mathbb{R}^n$ является отображением из пространства \mathbb{R}^T в признаковое пространство $G \subset \mathbb{R}^n$. Имея функцию порождения признаков g , преобразуем исходную выборку (2) в

$$\mathcal{D}_G = \{(g_i, y_i)\}_{i=1}^m,$$

где $\mathbf{g}_i = \mathbf{g}(s_i) \in G$.

Модель классификации $f = f(\mathbf{g}, \boldsymbol{\theta})$ параметризована вектором $\boldsymbol{\theta}$. Оптимальные параметры $\hat{\boldsymbol{\theta}}$ определяются оптимизацией функции ошибки классификации

$$\hat{\boldsymbol{\theta}} = \arg \min_{\boldsymbol{\theta}} L(\boldsymbol{\theta}, \mathcal{D}_G, \boldsymbol{\mu}). \quad (3)$$

Вектор $\boldsymbol{\mu}$ является внешним параметром для заданной модели классификации. Примеры таких параметров и функций ошибки для различных моделей классификации приведены ниже.

Чтобы сравнить качество классификации с прошлыми результатами [11,12], в качестве метрики качества используется точность классификации:

$$\text{accuracy} = \frac{1}{m} \sum_{i=1}^m \left[f(\mathbf{g}(s_i), \hat{\boldsymbol{\theta}}) = y_i \right]. \quad (4)$$

3 Порождение признаков

Цель данной работы — провести сравнение различных подходов к генерации признаков. В этом разделе проводится анализ рассматриваемых методов.

Экспертные функции. В качестве базового подхода будем использовать экспертные функции как функции порождения признаков. Экспертные функции — это некоторые статистики g_j , где $g_j : \mathbb{R}^T \rightarrow \mathbb{R}$. Признаковым описанием $\mathbf{g}(s)$ объекта s являются значения заданных экспертных статистик для данного объекта

$$\mathbf{g}(s) = [g_1(s), \dots, g_n(s)]^\top.$$

В работе [13] авторы предлагают использовать экспертные функции, приведенные в таблице 1. Такая процедура порождения признаков генерирует признаковое описание временного ряда $\mathbf{g}(s) \in \mathbb{R}^{40}$.

Таблица 1: Expert functions

Function description	Formula
Mean	$\bar{x} = \frac{1}{T} \sum_{t=1}^T x_t$
Standard deviation	$\sqrt{\frac{1}{T} \sum_{t=1}^T (x_t - \bar{x})^2}$
Mean absolute deviation	$\frac{1}{T} \sum_{t=1}^T x_t - \bar{x} $
Distribution	Histogram values with 10 bins

Авторегрессионная модель. Авторегрессионная модель [16] порядка n использует параметрическую модель для аппроксимации временного ряда s . Каждое значение временного ряда приближается линейной комбинацией предыдущих $n - 1$ значений

$$x_t = w_0 + \sum_{j=1}^{n-1} w_j x_{t-j} + \epsilon_t,$$

где ϵ_t — регрессионные остатки. Оптимальные параметры $\hat{\mathbf{w}}$ авторегрессионной модели используются как признаки $\mathbf{g}(s)$. Данные параметры минимизируют квадратичную ошибку аппроксимации временного ряда и предсказания модели

$$\mathbf{g}(s) = \hat{\mathbf{w}} = \arg \min_{\mathbf{w} \in \mathbb{R}^n} \left(\sum_{t=n}^T \|x_t - \hat{x}_t\|^2 \right). \quad (5)$$

Задача (5) эквивалентна задаче линейной регрессии. Поэтому для каждого временного ряда s необходимо решить задачу линейной регрессии размера n . Пример аппроксимации временного ряда авторегрессионной моделью представлен на Рис. 1.

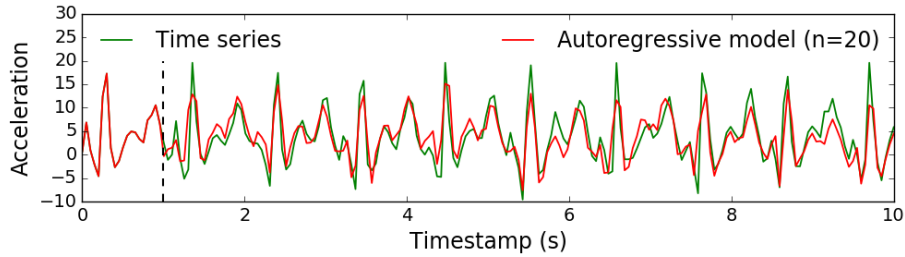


Рис. 1: Time series approximation using autoregressive model with order $n = 20$

Анализ сингулярного спектра. Альтернативной гипотезой порождения признаков пространства для временного ряда является анализ сингулярного спектра (Singular Spectrum Analysis, SSA) [7]. Для каждого временного ряда s из выборки \mathcal{D} строится траекторная матрица:

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

Здесь ширина окна n является внешним структурным параметром. Сингулярное разложение [5] матрицы $\mathbf{X}^\top \mathbf{X}$:

$$\mathbf{X}^\top \mathbf{X} = \mathbf{U} \mathbf{\Lambda} \mathbf{U}^\top,$$

где \mathbf{U} — унитарная матрица и $\mathbf{\Lambda} = \text{diag}(\lambda_1, \dots, \lambda_n)$ причём λ_i собственные значения $\mathbf{X}^\top \mathbf{X}$. Признаковое описание объекта s задаётся спектром матрицы $\mathbf{X}^\top \mathbf{X}$:

$$\mathbf{g}(s) = [\lambda_1, \dots, \lambda_n]^\top.$$

Spline Approximation. Предлагаемый метод аппроксимирует временные ряды с помощью сплайнов [3]. Сплайн определяется его параметрами: узлами и коэффициентами. Предполагается, что узлы сплайна $\{\xi_\ell\}_{\ell=0}^M$ равномерно распределены по временной оси. Кусочные модели, построенные на отрезках $[\xi_{\ell-1}; \xi_\ell]$, заданы коэффициентами $\{\mathbf{w}_\ell\}_{\ell=1}^M$. Оптимальные пара-

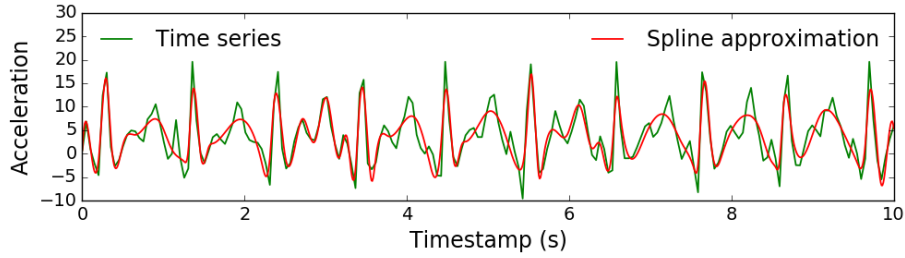


Рис. 2: Time series approximation using three order splines

метры сплайна являются решением системы с дополнительными условиями равенства производных до второго порядка включительно на концах отрезков. Обозначим каждый отрезок-сегмент $p_i(t)$ $i = 1, \dots, M$ и весь сплайн $S(t)$. Тогда система уравнений принимает вид

$$S(t) = \begin{cases} p_1(t) = w_{10} + w_{11}t + w_{12}t^2 + w_{13}t^3, & t \in [\xi_0, \xi_1], \\ p_2(t) = w_{20} + w_{21}t + w_{22}t^2 + w_{23}t^3, & t \in [\xi_1, \xi_2], \\ \dots & \dots \\ p_M(t) = w_{L0} + w_{M1}t + w_{M2}t^2 + w_{M3}t^3, & t \in [\xi_{M-1}, \xi_M], \end{cases}$$

$$\begin{aligned} S(\xi_t) &= x_t, \quad t = 0, \dots, M, \\ p'_i(\xi_i) &= p'_{i+1}(\xi_i), \quad p''_i(\xi_i) = p''_{i+1}(\xi_i), \quad i = 1, \dots, M-1, \\ p_i(\xi_{i-1}) &= x_{i-1}, \quad p_i(\xi_i) = x_i, \quad i = 1, \dots, M. \end{aligned}$$

Объединение всех параметров сплайна задаёт признаковое описание временного ряда:

$$\mathbf{g}(s) = [\mathbf{w}_1, \dots, \mathbf{w}_M]^\top.$$

Рис. 2 показывает аппроксимацию временного ряда с использованием модели сплайнов. По сравнению с авторегрессионной моделью сплайны строят более гладкую аппроксимацию, используя такое же количество параметров.

4 Классификация временных рядов

Для классификации временных рядов будем использовать подход один против всех. Для каждого класса обучается бинарный классификатор, и на стадии предсказания объект классифицируется согласно наиболее уверенному классификатору. Использовались три модели классификации: логистическая регрессия, SVM и случайный лес.

Логистическая регрессия. Оптимальные параметры модели $\hat{\mathbf{w}}, \hat{b}$ в случае логистической регрессии определяются минимизацией функции ошибки (3)

$$L(\boldsymbol{\theta}, \mathcal{D}_G, \mu) = \sum_{i=1}^m \log(1 + \exp(-y_i[\mathbf{w}^\top \mathbf{g}_i + b])) + \frac{\mu}{2} \|\mathbf{w}\|^2, \text{ where } \boldsymbol{\theta} = \begin{bmatrix} \mathbf{w} \\ b \end{bmatrix}.$$

Решающее правило $f(\mathbf{g}, \boldsymbol{\theta})$ — знак линейной комбинации описания объекта \mathbf{g} и параметров $\hat{\boldsymbol{\theta}}$

$$\hat{y} = f(\mathbf{g}, \hat{\boldsymbol{\theta}}) = \text{sgn}(\mathbf{g}^\top \hat{\mathbf{w}} + \hat{b}).$$

SVM. Оптимизационная задача метода SVM имеет вид

$$\hat{\boldsymbol{\theta}} = \begin{pmatrix} \hat{\mathbf{w}} \\ \hat{b} \\ \hat{\boldsymbol{\xi}} \end{pmatrix} = \arg \min_{\mathbf{w}, b, \boldsymbol{\xi}} \frac{1}{2} \|\mathbf{w}\|^2 + \mu \sum_{i=1}^m \xi_i, \text{ s.t. } y_i (\mathbf{w}^\top \mathbf{g}_i + b) \geq 1 - \xi_i, \\ \xi_i \geq 0, \quad 1 \leq i \leq m.$$

Целевая функция соответствует функции ошибки классификации $L(\boldsymbol{\theta}, \mathcal{D}_G, \mu)$. Предсказание для нового объекта вычисляется аналогично $\hat{y} = \text{sgn}(\mathbf{g}^\top \hat{\mathbf{w}} + \hat{b})$.

Случайный лес. Случайный лес использует идею бэггинга. Идея состоит в построении многих слабых, неустойчивых классификаторов на подвыборках с возвращениями и усреднения их предсказаний. Метод предполагает использование в качестве базовых классификаторов моделей с низким смещением и высокой дисперсией. Усреднение позволяет уменьшить дисперсию. В случае случайного леса базовой моделью выступают решающие деревья. Идея

бэггинга используется не только для самих объектов, но и для множества признаков. В данном случае предсказание для нового объекта получается усреднением всех предсказаний отдельных деревьев:

$$\hat{y} = \text{sgn} \left(\frac{1}{B} \sum_{i=1}^B \text{pred}(\mathbf{g}_i) \right),$$

где B — количество деревьев в композиции.

5 Эксперимент

В данной работе эксперименты проводились на двух датасетах временных рядов акселерометра мобильного телефона: WISDM [21] и USC-HAD [19]. Акселерометр мобильного телефона проводит измерение ускорения по трём осям с частотой 100 Hz. Данные WISDM содержат 4321 временной ряд. Каждый временной ряд принадлежит к одному из 6 классов. Данные USC-HAD содержат 13620 временных рядов, принадлежащих одному из 12 классов. В таблице 2 представлено распределение временных рядов по классам для каждого датасета. Длина временного ряда равна 200. На рис. 3 представлен пример одного из временных рядов.

Таблица 2: Distributions of the classes
(a) WISDM (b) USC-HAD

Activity	# objects	Activity	# objects
1 Standing	229 5.30 %	1 Standing	1167 8.57 %
2 Walking	1917 44.36 %	2 Elevator-up	764 5.61 %
3 Upstairs	466 10.78 %	3 Walking-forward	1874 13.76 %
4 Sitting	277 6.41 %	4 Sitting	1294 9.50 %
5 Jogging	1075 24.88 %	5 Walking-downstairs	951 6.98 %
6 Downstairs	357 8.26 %	6 Sleeping	1860 13.66 %
Total	4321	7 Elevator-down	763 5.60 %
		8 Walking-upstairs	1018 7.47 %
		9 Jumping	495 3.63 %
		10 Walking-right	1305 9.58 %
		11 Walking-left	1280 9.40 %
		12 Running	849 6.23 %
		Total	13620

В эксперименте для каждого датасета были порождены признаки одним из методов: экспертные функции, авторегрессионная модель, SSA и сплайны. Для каждой процедуры порождения признакового описания строились три модели классификации: логистическая регрессия, SVM и

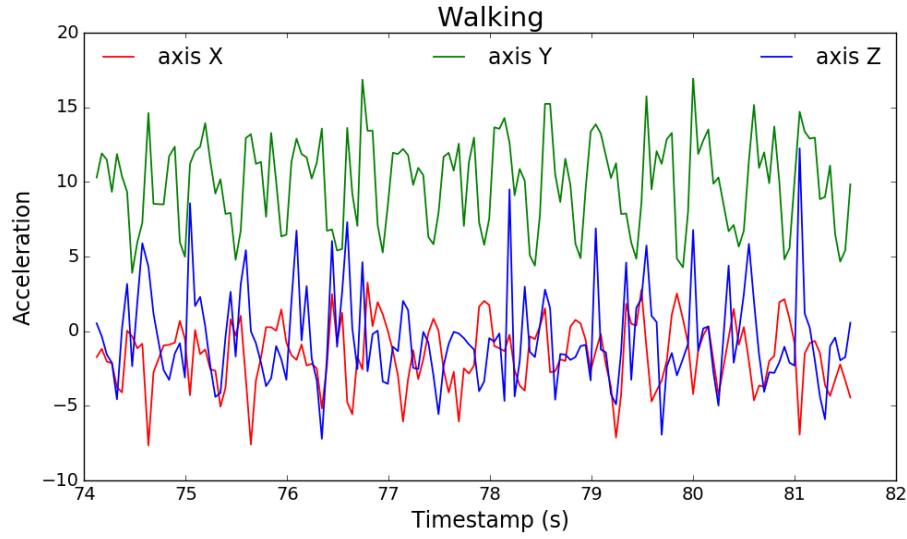


Рис. 3: Time series example

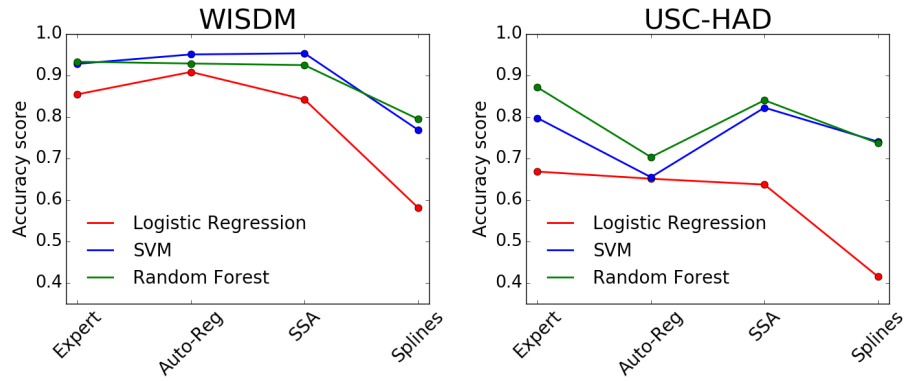


Рис. 4: Multiclass accuracy score

случайный лес. Внешние структурные параметры (длина авторегрессионной модели n , ширина окна SSA n , число узлов сплайна M) настраивались процедурой кросс-валидации:

$$CV(K) = \frac{1}{K} \sum_{k=1}^K L(f_k, \mathcal{D} \setminus \mathcal{C}_k), \quad (6)$$

где $C_k = \frac{K-1}{K}$ — доля от всей выборки, используемая для обучения модели f_k . Гиперпараметры μ моделей классификации были настроаны той же процедурой кросс-валидации.

Первый подход к порождению признаков временных рядов — экспертные функции. Основной недостаток такого подхода необходимость экспертного задания функций и возможности вычисления их для конкретного датасета.

Авторегрессионная модель требует задания параметра длины модели n . Процедура кросс-валидации дала наибольшее качество при $n = 20$ для обоих датасетов.

Модель SSA была настроена аналогичной процедурой выбора оптимальных гиперпараметров. Конечная модель имела ширину окна $n = 20$.

Для аппроксимации временных рядов кубическими сплайнами [3] использовалась библиотека *scipy* [10]. Узлы сплайнов $\{\xi_\ell\}_{\ell=1}^M$ были распределены равномерно по временной оси. Значение параметра M было подобрано на кросс-валидации.

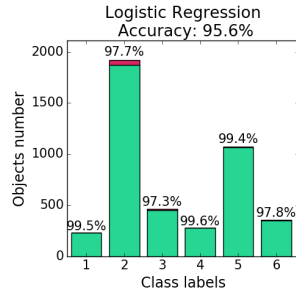
Для обоих датасетов процедуры порождения признаковов описаний дали следующие количества признаков: экспертные функции — 40; авторегрессионная модель — 60; анализ сингулярного спектра — 60; сплайны — 33.

Таблица 3: Binary accuracy scores for WISDM using different feature generation methods: EX — Expert, AR — Auto-Reg, SSA and SPL for Splines

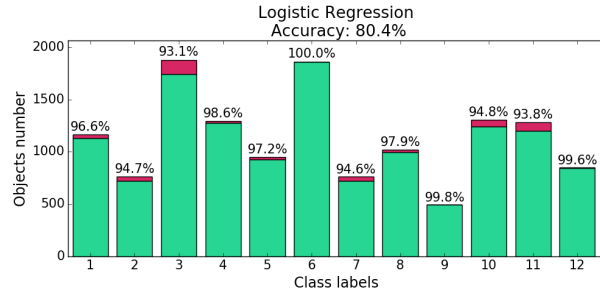
	Logistic Regression				Random Forest				SVM			
	EX	AR	SSA	SPL	EX	AR	SSA	SPL	EX	AR	SSA	SPL
All	0.85	0.91	0.84	0.58	0.93	0.93	0.92	0.79	0.93	0.95	0.95	0.77
Standing	0.99	0.98	1.00	0.95	1.00	0.99	1.00	0.99	0.99	0.98	1.00	0.96
Walking	0.91	0.96	0.86	0.61	0.96	0.97	0.95	0.86	0.96	0.98	0.98	0.84
Upstairs	0.91	0.95	0.91	0.89	0.96	0.96	0.96	0.90	0.96	0.98	0.97	0.89
Sitting	0.99	0.98	1.00	0.99	1.00	0.99	1.00	1.00	0.99	0.98	1.00	1.00
Jogging	0.98	0.99	0.99	0.80	0.99	0.99	0.99	0.92	0.99	0.99	0.99	0.93
Downstairs	0.93	0.96	0.94	0.92	0.96	0.97	0.96	0.92	0.96	0.98	0.97	0.92

На рис. 4 показано качество классификации (4) для двух датасетов. Для данных WISDM сплайны дали самое слабое качество классификации. Результаты для экспертных функций, авторегрессионной модели и SSA схожи. Для данных USC-HAD результат более восприимчив к выбору модели классификации. Для обоих датасетов логистическая регрессия продемонстрировала наименьшее качество, SVM и случайный лес показали почти одинаковое качество. Для датасета USC-HAD модель с использованием аппроксимации сплайнами показала сравнимое с другими методами качество.

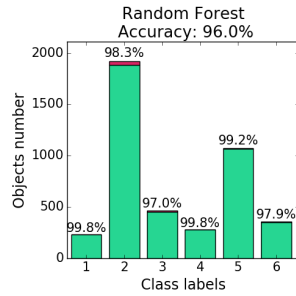
В Табл. 3 и Табл. 4 представлены результаты классификации (4) для каждого класса в отдельности. Первая строка в обеих таблицах демонстрирует точность по всем классам для каждой модели и процедуры генерации признаков. Следующие строки соответствуют бинарным точностям по каждому из классов. Для данных WISDM лучшее качество имеют наименее



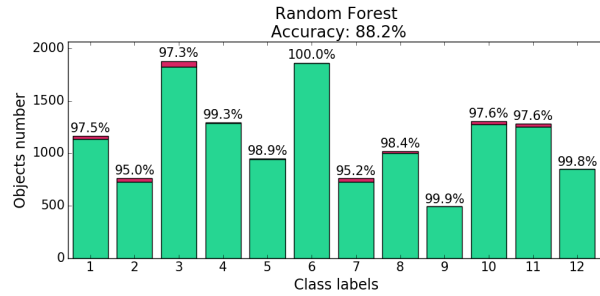
(a) WISDM dataset



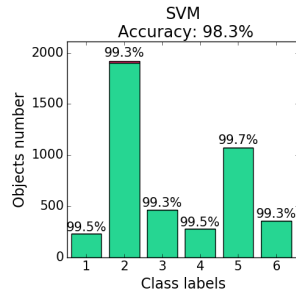
(b) USC-HAD dataset



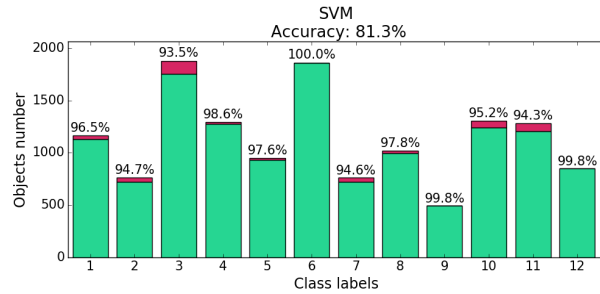
(c) WISDM dataset



(d) USC-HAD dataset



(e) WISDM dataset



(f) USC-HAD dataset

Рис. 5: Accuracy scores of classification of each class using all features

активные классы, такие как Standing и Sitting. Для USC-HAD заметного выделения качества для определенных классов не наблюдается.

Также был проведён эксперимент с использованием объединённого множества всех 193 сгенерированных признаков. Результаты представлены на Рис. 5. Соответствие между номера классов и видами активности приведено в Табл. 2. Объединение признаков для обучения одной модели позволило

Таблица 4: Binary accuracy scores for USC-HAD using different feature generation methods: EX — Expert, AR — Auto-Reg, SSA and SPL for Splines

	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

увеличить качество. Для данных WISDM все точности классификации по классам больше 97%, а для USC-HAD выше 93%.

6 Заключение

В работе рассматривается задача классификации объектов сложной структуры. В эксперименте было произведено сравнение различных подходов к извлечению признаков, таких как экспертно заданные функции и вычисление параметров локальных моделей. Результаты эксперимента показали, что рассмотренные методы позволяют восстановить метки классов с высокой точностью. Предложенный метод аппроксимации сплайнами позволяет получить гладкое приближение временного ряда, используя меньшее число параметров. Качество классификации с использованием сплайнами сравнимо с другими рассмотренными методами. Конкатенация всех признаков позволяет получить прирост в качестве.

Список литературы

1. Bao, L., Intille, S.S.: Activity recognition from user-annotated acceleration data. In: International Conference on Pervasive Computing. pp. 1–17. Springer (2004)
2. Budnik, M., Gutierrez-Gomez, E.L., Safadi, B., Pellerin, D., Quénot, G.: Learned features versus engineered features for multimedia indexing. Multimedia Tools and Applications 76(9), 11941–11958 (2017)
3. De Boor, C.: A practical guide to splines, vol. 27. Springer-Verlag (1978)
4. Geurts, P.: Pattern extraction for time series classification. In: European Conference on Principles of Data Mining and Knowledge Discovery. pp. 115–127. Springer (2001)

5. Golub, G.H., Reinsch, C.: Singular value decomposition and least squares solutions. *Numerische mathematik* 14(5), 403–420 (1970)
6. Hammerla, N.Y., Halloran, S., Ploetz, T.: Deep, convolutional, and recurrent models for human activity recognition using wearables. *arXiv preprint arXiv:1604.08880* (2016)
7. Hassani, H.: Singular spectrum analysis: methodology and comparison. *Journal of Data Science* 5(2), 239–257 (2007)
8. Ignatov, A.D., Strijov, V.V.: Human activity recognition using quasiperiodic time series collected from a single tri-axial accelerometer. *Multimedia tools and applications* 75(12), 7257–7270 (2016)
9. Incel, O.D., Kose, M., Ersoy, C.: A review and taxonomy of activity recognition on mobile phones. *BioNanoScience* 3(2), 145–171 (2013)
10. Jones, E., Oliphant, T., Peterson, P.: SciPy: Open source scientific tools for Python (2001), <http://www.scipy.org/>, [Online; accessed today]
11. Karasikov, M., Strijov, V.: Feature-based time-series classification. *Intelligence* 24(1), 164–181 (2016)
12. Kuznetsov, M., Ivkin, N.: Time series classification algorithm using combined feature description. *Machine Learning and Data Analysis* 1(11), 1471–1483 (2015)
13. Kwapisz, J.R., Weiss, G.M., Moore, S.A.: Activity recognition using cell phone accelerometers. *ACM SigKDD Explorations Newsletter* 12(2), 74–82 (2011)
14. Lu, L., Qing-ling, C., Yi-Ju, Z.: Activity recognition in smart homes. *Multimedia Tools and Applications* pp. 1–18 (2016)
15. Lu, Y., Wei, Y., Liu, L., Zhong, J., Sun, L., Liu, Y.: Towards unsupervised physical activity recognition using smartphone accelerometers. *Multimedia Tools and Applications* 76(8)
16. Lukashin, Y.P.: Adaptive methods of short-term forecasting of time series. *M.: Finance and statistics* (2003)
17. Motrenko, A., Strijov, V.: Extracting fundamental periods to segment biomedical signals. *IEEE journal of biomedical and health informatics* 20(6), 1466–1476 (2016)
18. Umeda, Y.: Time series classification via topological data analysis. *Transactions of the Japanese Society for Artificial Intelligence* 32(3), D–G72_1 (2017)
19. The usc human activity dataset. <http://www-scf.usc.edu/~mizhang/datasets.html>
20. Wang, W., Liu, H., Yu, L., Sun, F.: Human activity recognition using smart phone embedded sensors: A linear dynamical systems method. In: *Neural Networks (IJCNN), 2014 International Joint Conference on*. pp. 1185–1190. IEEE (2014)
21. The wisdm dataset. <http://www.cis.fordham.edu/wisdm/dataset.php>