

# Порождение признаков с помощью локально-аппроксимирующих моделей.\*

Садиев А. А.<sup>1</sup>, Мотренко А. П.<sup>1</sup>, Стрижов В. В.<sup>1</sup>

sadiev.aa@phystech.edu, anastasiya.motrenko@phystech.edu, strijov@ccas.ru

<sup>1</sup>Московский физико-технический институт (МФТИ)

Рассматриваются методы определения вида деятельности человека по измерениям акселерометра. Статья посвящена исследованию проблемы порождения признаков с использованием локально-аппроксимирующих моделей. В работе строится набор локально-аппроксимирующих моделей и проверяется корректность применения гипотезы о простоте выборки для порожденных признаков. Также внимание уделено выбору оптимального способа порождения признаков временного ряда. В контексте данной работы предполагается метод построения метрического пространства описаний элементарных движений.

**Ключевые слова:** *временной ряд, локально-аппроксимирующая модель, метрическое пространство.*

## 1 Введение

Работа посвящена поиску оптимальных признаков для задачи классификации видов деятельности человека. Исследование проводится с целью автоматизации порождения признаков слабоструктурированных данных, таких как временные ряды. Оптимальный выбор признаков должен удовлетворять выборкам временных рядов с различными частотами. Также предлагаемый в данной работе метод должен обеспечивать минимальное расхождение в точности задачи классификации с различными множествами ответов.

Проблема оптимального порождения признаков решается множеством способов: в работе [?] выделяются фундаментальные периоды временных рядов, в [?] внимание уделено сегментации временного ряда различными способами. В данной работе задача решается с помощью построения локально-аппроксимирующих моделей исходной выборки. Предлагаемый метод не дает наилучшую точность среди уже имеющихся способов, однако является универсальным для данных с различными параметрами выборок.

Исследование проводится на данных временных рядов акселерометра WISDM с целью решения задачи классификации.

## 2 Название раздела

Данный документ демонстрирует оформление статьи, подаваемой в электронную систему подачи статей <http://jmla.org/papers> для публикации в журнале «Машинное обучение и анализ данных». Более подробные инструкции по стилевому файлу `jmla.sty` и использованию издательской системы  $\text{\LaTeX}$  находятся в документе `authors-guide.pdf`. Работу над статьёй удобно начинать с правки  $\text{\TeX}$ -файла данного документа.

### 2.1 Название параграфа.

Нет ограничений на количество разделов и параграфов в статье. Разделы и параграфы не нумеруются.

---

\*Работа выполнена при финансовой поддержке РФФИ, проект №00-00-00000. Научный руководитель: Стрижов В. В. Задачу поставил: Эксперт И. О. Консультант: Мотренко А. П.

## 2.2 Теоретическую часть работы

желательно структурировать с помощью окружений Def, Axiom, Hypothesis, Problem, Lemma, Theorem, Corollary, State, Example, Remark.

**Определение 1.** Математический текст хорошо структурирован, если в нём выделены определения, теоремы, утверждения, примеры, и т. д., а неформальные рассуждения (мотивации, интерпретации) вынесены в отдельные параграфы.

**Утверждение 1.** Мотивации и интерпретации наиболее важны для понимания сути работы.

**Теорема 1.** Не менее 90% коллег, заинтересовавшихся Вашей статьёй, прочитают в ней не более 10% текста.

**Доказательство.** Причём это будут именно те разделы, которые не содержат формул. ■

**Замечание 1.** Выше показано применение окружений Def, Theorem, State, Remark, Proof.

## 3 Некоторые формулы

Образец формулы:  $f(x_i, \alpha^\gamma)$ .

Образец выключной формулы без номера:

$$y(x, \alpha) = \begin{cases} -1, & \text{если } f(x, \alpha) < 0; \\ +1, & \text{если } f(x, \alpha) \geq 0. \end{cases}$$

Образец выключной формулы с номером:

$$y(x, \alpha) = \begin{cases} -1, & \text{если } f(x, \alpha) < 0; \\ +1, & \text{если } f(x, \alpha) \geq 0. \end{cases} \quad (1)$$

Образец выключной формулы, разбитой на две строки с помощью окружения align:

$$R'_N(F) = \frac{1}{N} \sum_{i=1}^N \left( P(+1 | x_i) C(+1, F(x_i)) + \right. \\ \left. + P(-1 | x_i) C(-1, F(x_i)) \right). \quad (2)$$

Образцы ссылок: формулы (1) и (2).

## 4 Пример иллюстрации

Рисунки вставляются командой `\includegraphics`, желательно с выравниванием по ширине колонки: `[width=\linewidth]`.

Практически все популярные пакеты рисуют графики с подписями, которые трудно читать на бумаге и на слайдах из-за малого размера шрифта. Шрифт на графиках (подписи осей и цифры на осях) должны быть такого же размера, что и основной текст.

При значительном количестве рисунков рекомендуется группировать их в одном окружении `{figure}`, как это сделано на рис. 1.

## 5 Пример таблицы

Подпись делается *над таблицей*, см. таблицу 1.

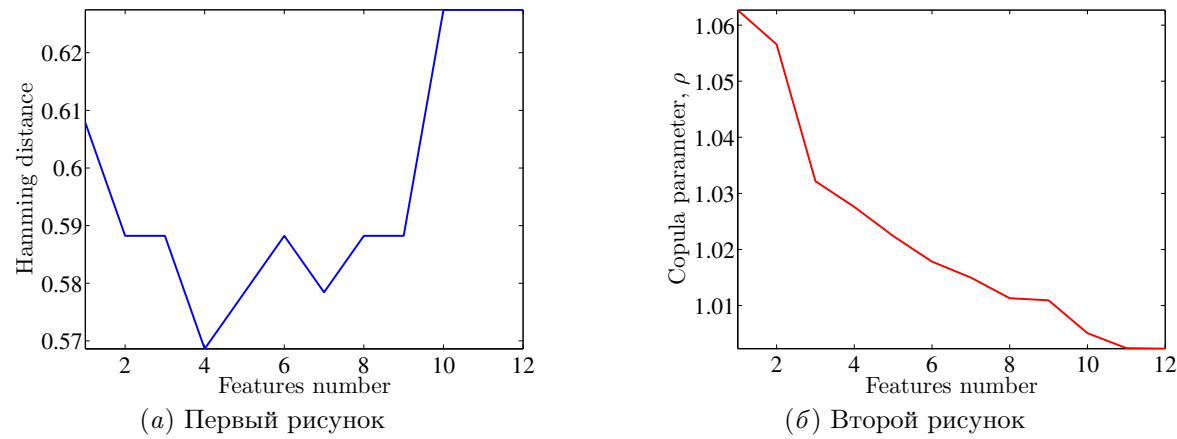


Рис. 1 Подпись должна размещаться под рисунком.

Таблица 1 Подпись размещается над таблицей.

Задача	CCEL	boosting
Cancer	$3.46 \pm 0.37$ (3.16)	$4.14 \pm 1.48$
German	$25.78 \pm 0.65$ (1.74)	$29.48 \pm 0.93$
Hepatitis	$18.38 \pm 1.43$ (2.87)	$19.90 \pm 1.80$

6 Заключение

Желательно, чтобы этот раздел был, причём он не должен дословно повторять аннотацию. Обычно здесь отмечают, каких результатов удалось добиться, какие проблемы остались открытыми.