

Порождение признаков с помощью локально-аппроксимирующих моделей*

М. Е. Христолюбов¹

khristolyubov.me@phystech.edu

¹Московский физико-технический институт

В работе рассматривается многоклассовая классификация временных рядов. Классификация производится методом порождения признаков с помощью локально аппроксимирующих моделей. Предполагается, что временные ряды измерений содержат кластеры в пространстве описаний временных рядов. Предлагается разбить временные ряды на сегменты, а в качестве признаков использовать параметры моделей, обученных на этих сегментах. Изучается выполнимость гипотезы о простоте выборки для порожденных признаков и информативность порожденных признаков.

Ключевые слова: *временной ряд; многоклассовая классификация; сегментация временных рядов; локально аппроксимирующая модель*

DOI: 00.00000/000000000

1 Введение

В статье изучается задача идентификации движений человека по временным рядам. Данные представляют измерения акселерометра и гироскопа, встроенных в мобильное устройство iPhone 6s, хранящегося в переднем кармане участника. Метками классов служат: подъем по лестнице вверх, спуск по лестнице вниз, ходьба, бег трусцой, сидение, лежание. В дополнении к этому исследуется возможность выделения атрибутивных паттернов, которые могут быть использованы для определения пола или личности субъектов данных в дополнение к их деятельности.

Базовый метод решения задачи берется из [1] и [2]. В работе [1] аппроксимирующая модель обучается на всем временном ряду без сегментации. В качестве признаков используются параметры модели авторегрессии, а так же собственные числа траекторной матрицы, в случае модели сингулярного спектра. В работе [2] проводится сегментация временных рядов, а так же сравнивается алгоритм голосования сегментов и классификация в пространстве параметров распределений признаков их сегментов. В работе [7] используются скрытые марковские модели, но особенность этой работы заключается в классификации действий в каждый момент времени, а не в классификация отдельных временных рядов.

Классификация временных рядов осложнена тем, что исходное описание временного ряда (последовательность значений, снятых с определенной частотой) нельзя использовать в качестве признаков, в связи с тем, что отдельные значения несут мало информации о характерном виде временного ряда. Требуется найти адекватное признаковое пространство, удобное для применения моделей классификации. В качестве признаков могут использоваться значения некоторых функций, задаваемые экспертом (дисперсия, минимальное и максимальное значения, среднее значения и т.п.). Однако это качество этого

*Работа выполнена при финансовой поддержке РФФИ, проекты № №00-00-00000 и 00-00-00001.

метода зависит от содержательности экспертных функций, и, зачастую, дает посредственные результаты. Кроме того, в качестве признаков могут использоваться коэффициенты разложения в ряд Фурье, сингулярный спектр.

Еще одним подходом является обучение некоторой модели, аппроксимирующей временной ряд, и отображение временного ряда в пространство параметров обученной модели. В качестве модели, параметры которой берутся в качестве признаков, можно взять линейную модель, модель авторегрессии. Для этого метода требуется сегментация временного ряда, и способ сегментации является важной проблемой.

В работе предлагается оптимальный способ сегментации и метод выделения некоторых элементарных движений, по признаковому описанию которых можно будет идентифицировать род деятельности человека соответствующий временному ряду.

2 Название раздела

2.1 Название параграфа

Разделы и параграфы, за исключением списков литературы, нумеруются.

3 Заключение

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

Литература

- [1] Карасиков М.Е., Стрижов В.В. Классификация временных рядов в пространстве параметров порождающих моделей // Информатика и ее применения, 2016. doi: <http://dx.doi.org/10.3114/S187007708007>.

Поступила в редакцию 00.00.0000

Machine Learning and Data Analysis journal paper template*

F. S. Author¹, F. S. Co-Author², and F. S. Name^{1,2}

khristolyubov.me@phystech.edu

¹Organization, address; ²Organization, address

This is the template of the paper submitted to the journal “Machine Learning and Data Analysis”.

The title should be concise and informative. Titles are often used in information-retrieval systems. Avoid abbreviations and formulae where possible.

A concise and factual abstract is required.

Background: One paragraph about the problem, existent approaches and its limitations.

Methods: One paragraph about proposed method and its novelty.

Results: One paragraph about major properties of the proposed method and experiment results if applicable.

Concluding Remarks: One paragraph about the place of the proposed method among existent approaches.

Keywords: *keyword; keyword; more keywords, separated by “;”*

*The research was supported by the Russian Foundation for Basic Research (grants 00-00-0000 and 00-00-00001).

DOI: 00.00000/00000000

References

- [1] N. P. Ivkin, M. P. Kuznetsov.. 2015. Time series classification algorithm using combined feature description. . *Machine Learning and Data Analysis* (11):1471–1483.
- [2] V. V. Strijov, M. E. Karasikov. 2016. Feature-based time-series classification *Informatics*
- [3] V.V. Strijov, R.V. Isachenko.. 2016. Metric learning in multiclass time series classification problem. *Informatics and Applications* (10(2)):48–57.
- [4] V.V. Strijov, Andrew Zadayanchuk, Maria Popova.. 2016. Selection of optimal physical activity classification model using measurements of accelerometer. *Information Technologies* (22(4)):313–318.
- [5] Strijov V.V., Motrenko A.P.. 2016. Extracting fundamental periods to segment human motion time series. *Journal of Biomedical and Health Informatics* 20(6):1466 – 1476.
- [6] Strijov V.V., Ignatov A.. 2015. Human activity recognition using quasiperiodic time series collected from a single triaxial accelerometer. *Multimedia Tools and Applications* pages 1–14.
- [7] Dafne van Kuppevelt, Joe Heywood, Mark Hamer, Séverine Sabia, Emla Fitzsimons, Vincent van Hees. 2019. Segmenting accelerometer data from daily life with unsupervised machine learning. *PLOS ONE* doi: <http://dx.doi.org/10.5255/UKDA-SN-8156-3>.
- [8] Malekzadeh, Mohammad and Clegg, Richard G. and Cavallaro, Andrea and Haddadi, Hamed. 2019. *Mobile Sensor Data Anonymization* pages 49–58. Proceedings of the International Conference on Internet of Things Design and Implementation doi: <http://dx.doi.org/10.1145/3302505.3310068>.

Received January 00, 0000