

**МЕТРИЧЕСКОЕ ОБУЧЕНИЕ В ЗАДАЧАХ
МУЛЬТИКЛАССОВОЙ КЛАССИФИКАЦИИ
ВРЕМЕННЫХ РЯДОВ**

Исаченко Роман Владимирович

Студент

Московский физико-технический институт, Долгопрудный, Россия

E-mail: isa-ro@yandex.ru

Решается задача мультиклассовой классификации временных рядов. Для решения этой задачи ранее использовались метод опорных векторов, нейронные сети, байесовский подход. В данной работе для классификации временных рядов используется идея ближайших соседей.

Для улучшения качества классификации предлагается использовать методы метрического обучения [1]. Метрическое обучение позволяет модифицировать расстояния между временными рядами с помощью линейного преобразования признакового пространства объектов. В результате преобразования временные ряды одного класса оказываются ближе друг к другу по отношению к выбранной метрике, а временные ряды, принадлежащие разным классам, отдаляются друг от друга. Методы метрического обучения применяются при ранжировании поисковой выдачи, идентификации лиц, распознавании рукописных цифр. В данной работе в качестве алгоритма метрического обучения был выбран алгоритм LMNN [2]. Данный алгоритм основан на идеях метода k ближайших соседей. Алгоритм для каждого объекта минимизирует расстояния до k ближайших соседей, принадлежащих тому же классу и штрафует объекты из других классов, попавшие на расстояние порядка расстояния до k -ого ближайшего соседа.

В качестве начальной предобработки временных рядов в данной работе производится их выравнивание относительно центроидов классов. Задача поиска оптимального центроида класса решается с помощью DBA метода [3]. Выравнивание производится методом динамической трансформации [4]. Классификация, основанная на идее ближайших соседей чувствительна к масштабам имеющихся признаков. Для повышения устойчивости классификации выравненные временные ряды были отнормированы.

Тем самым полученная модель классификации представляет собой суперпозицию алгоритмов: выравнивания временных рядов относительно центроидов классов, метрического обучения алгоритмом

LMNN, классификации.

В данной работе вычислительный эксперимент проводится на синтетических данных временных рядов, представляющих аналитические функции, и реальных данных показаний акселерометра мобильного телефона. Цель эксперимента — определить вид активности человека по форме сигнала.

В данной работе предложен новый подход к классификации временных рядов. Вычислительный эксперимент проиллюстрировал работоспособность алгоритма и улучшение качества классификации при использовании метрического обучения.

Литература

1. Bellet A., Habrard A., Sebban M., A survey on metric learning for feature vectors and structured data: arXiv:1306.6709, 2013.
2. Weinberger K. Q., Blitzer J., Saul L. K. Gançarski P. Distance metric learning for large margin nearest neighbor classification. 2005. Advances in neural information processing systems. P.1473–1480.
3. Petitjean F., Ketterlin A., Gançarski P. A global averaging method for dynamic time warping, with applications to clustering. 2011. Pattern Recognition. Vol. 44, № 3. P.678–693.
4. Berndt D. J., Clifford J. Using dynamic time warping to find patterns in time series. 1994. KDD workshop. Vol. 10, № 3. P.359–370.