

Metric learning in multiclass time series classification problem

R. V. Isachenko¹, V. V. Strijov²

Annotation: This paper is devoted to the problem of multiclass time series classification. We propose to align time series in relation to class centroids. Building of the centroids and alignment of time series carried out by dynamic time warping algorithm. The accuracy of classification depends significantly on the metric used to compute distances between time series. In this paper to improve classification accuracy we use distance metric learning approach. The metric learning procedure modifies distances between objects to make objects from the same cluster closer and from the different clusters more distant. The distance between time series is measured by Mahalanobis metric. Distance metric learning procedure finds optimal transformation matrix for Mahalanobis metric. To calculate quality of classification we carried out computational experiment on synthetic data and real data of human activity recognition and conclude about performance of the algorithm.

Keywords: time series classification; time series alignment; distance metric learning; LMNN algorithm.

References

- [1] Popova M. S., V. V. Strijov. 2015. Vybor optimal'noy modeli klassifikatsii fizicheskoy aktivnosti po izmereniyam akselerometra. *Informatika i ee primeneniya* 9(1):79–89.
- [2] Ignatov A. D., V. V. Strijov. 2015. Human activity recognition using quasiperiodic time series collected from a single tri-axial accelerometer. *Multimedia Tools and Applications* 1–14.

¹Moscow Institute of Physics and Technology, isa-ro@yandex.ru

²Dorodnicyn Computing Centre of the Russian Academy of Sciences, strijov@ccas.com

- [3] Güler, I., E. D. Übeyli. 2007. Multiclass support vector machines for eeg-signals classification. *Information Technology in Biomedicine, IEEE Transactions on* 11(2):117–126.
- [4] Übeyli, E. D. 2007. Ecg beats classification using multiclass support vector machines with error correcting output codes. *Digital Signal Processing* 17(3):675–684.
- [5] Anand, R., K. Mehrotra, C. K. Mohan, and S. Ranka. 1995. Efficient classification for multiclass problems using modular neural networks. *Neural Networks, IEEE Transactions on* 6(1):117–124.
- [6] Kafai, M., B. Bhanu. 2012. Dynamic bayesian networks for vehicle classification in video. *Industrial Informatics, IEEE Transactions on* 8(1):100–109.
- [7] Chaovalitwongse, W. A., Y.-J. Fan, and R. C. Sachdeo. 2007. On the time series k-nearest neighbor classification of abnormal brain activity. *Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on* 37(6):1005–1016.
- [8] Bellet, A., A. Habrard, and M. Sebban. 2013. A survey on metric learning for feature vectors and structured data. *arXiv preprint arXiv:1306.6709* .
- [9] Yang, L., R. Jin. 2006. Distance metric learning: A comprehensive survey. *Michigan State University* .
- [10] Wang, F., J. Sun. 2015. Survey on distance metric learning and dimensionality reduction in data mining. *Data Mining and Knowledge Discovery* 29(2):534–564.
- [11] McFee, B., G. R. Lanckriet. 2010. Metric learning to rank. *Proceedings of the 27th International Conference on Machine Learning (ICML-10)*. 775–782.
- [12] Guillaumin, M., J. Verbeek, and C. Schmid. 2009. Is that you? metric learning approaches for face identification. *Computer Vision, 2009 IEEE 12th international conference on*. 498–505.
- [13] Weinberger, K. Q., L. K. Saul. 2008. Fast solvers and efficient implementations for distance metric learning. *Proceedings of the 25th international conference on Machine learning*. 1160–1167.
- [14] Weinberger, K. Q., J. Blitzer, and L. K. Saul. 2005. Distance metric learning for large margin nearest neighbor classification. *Advances in neural information processing systems*. 1473–1480.
- [15] Berndt, D. J., J. Clifford. 1994. Using dynamic time warping to find patterns in time series. *KDD workshop*. 10(16):359–370.
- [16] Petitjean, F., A. Ketterlin, and P. Gançarski. 2011. A global averaging method for dynamic time warping, with applications to clustering. *Pattern Recognition* 44(3):678–693.

- [17] Goncharov, A. B., M. S. Popova, V. V. Strijov. 2015. Metricheskaya klassifikatsiya vremennykh ryadov s vyravnivaniem odnositel'no tsentroidov klassov. *Sistemy i sredstva informatiki* 25(4):52–64.
- [18] Vandenberghe, L., and S. Boyd. 1996. Semidefinite programming. *SIAM review* 38(1):49–95.
- [19] Isachenko R. V. Project code. // Sourceforge.net. 2015. Available at: <https://sourceforge.net/p/mlalgorithms/code/HEAD/tree/Group274/Isachenko2015TimeSeries/code/> (accessed March 18, 2016).
- [20] UCI repository. Human Activity Recognition Using Smartphones Dataset. Available at: <https://archive.ics.uci.edu/ml/datasets/Human+Activity+Recognition+Using+Smartphones> (accessed March 18, 2016).

Метрическое обучение в задачах мультиклассовой классификации временных рядов*

Р. В. Исаченко¹, В. В. Стрижов²

Аннотация: Работа посвящена построению модели многоклассовой классификации временных рядов. Предлагается выравнить временные ряды относительно центроидов классов. Процедура нахождения центроидов и выравнивания временных рядов осуществляется с помощью алгоритма динамической трансформации времени. Для повышения качества классификации в данной работе используются методы метрического обучения. Метрическое обучение позволяет модифицировать расстояния между временными рядами, сближая временные ряды из одного класса и отдаляя временные ряды из разных классов. Расстояние между временными рядами измеряется с помощью метрики Махаланобиса. Процедура метрического обучения состоит в определении оптимальной матрицы трансформаций в метрике Махаланобиса. Для анализа качества построенного алгоритма произведён вычислительный эксперимент на синтетических и реальных данных показаний с акселерометра мобильного телефона.

Ключевые слова: классификация временных рядов; выравнивание; метрическое обучение; алгоритм LMNN.

1 Введение

Решается задача мультиклассовой классификации временных рядов [1, 2]. Для решения этой задачи ранее использовались метод опорных векторов [3, 4], нейронные сети [5], байесовский подход [6]. В данной работе для классификации временных рядов используется идея ближайших соседей [7].

*Работа выполнена при финансовой поддержке РФФИ (проект 13-07-01163).

¹Московский физико-технический институт, isa-ro@yandex.ru

²Вычислительный центр им. А. А. Дородницына ФИЦ ИУ РАН, strijov@ccas.ru

8 Для повышения качества классификации предлагается использовать методы метрического обучения [8–10]. 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 [11], идентификации лиц [12], распознавании рукописных цифр [13]. В данной работе в качестве алгоритма метрического обучения был выбран алгоритм LMNN (Large Margin Nearest Neighbor) [14]. 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.

21 Алгоритм LMNN позволяет произвести отбор признаков. С помощью линейного преобразования алгоритм помещает объекты в новое пространство. Если размерность нового пространства меньше размерности исходного пространства, то происходит снижение размерности, т. е. отбор признаков.

25 Для вычисления расстояний между временными рядами в данной работе производится их выравнивание относительно центроидов классов. Выравнивание производится методом динамической трансформации времени DTW (Dynamic Time Warping) [15]. Задача поиска оптимального центроида класса решается с помощью метода выравненного взвешенного усреднения DBA (DTW Barycenter Averaging) [16]. Классификация, основанная на идее ближайших соседей чувствительна к изменению масштабов признаков. Для повышения устойчивости классификации выравненные временные ряды были отнормированы.

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

36 В данной работе вычислительный эксперимент проводится на синтетических временных рядах, представляющих аналитические функции, и реальных данных показаний акселерометра мобильного телефона. Цель эксперимента — определить вид активности человека по форме сигнала акселерометра. Получена оценка качества работы построенного алгоритма и произведён анализ его свойств.

41 2 Постановка задачи

42 Пусть объект $\mathbf{x}_i \in \mathbb{R}^n$ — временной ряд, последовательность измерений некоторой ис-
43 следуемой величины в различные моменты времени. Пусть \mathbf{X} — множество всех вре-
44 менных рядов фиксированной длины n , $Y = \{1, \dots, K\}$ — множество меток классов.
45 Пусть задана выборка $\mathfrak{D} = \{(\mathbf{x}_i, y_i)\}_{i=1}^\ell$ — множество объектов с известными метками
46 классов $y_i \in Y$.

Требуется построить точную, простую, устойчивую модель классификации

$$a : \mathbf{X} \rightarrow Y.$$

Данную модель представим в виде суперпозиции

$$a(\mathbf{x}) = b \circ \mathbf{f} \circ G(\mathbf{x}, \{\mathbf{c}_e\}_{e=1}^K), \quad (1)$$

47 где G — процедура выравнивания временных рядов относительно центроидов клас-
48 сов $\{\mathbf{c}_e\}_{e=1}^K$, \mathbf{f} — алгоритм метрического обучения, b — алгоритм многоклассовой клас-
49 сификации.

50 **Выравнивание временных рядов.** Для повышения качества и устойчивости ал-
51 горитма классификации предлагается провести выравнивание временных рядов каж-
52 дого класса относительно центроида.

Пусть \mathbf{X}_e — множество объектов обучающей выборки \mathfrak{D} , принадлежащих одно-
му классу $e \in \{1, \dots, K\}$. Центроидом множества объектов $\mathbf{X}_e = \{\mathbf{x}_i | y_i = e\}_{i=1}^\ell$ по
расстоянию ρ назовем вектор $\mathbf{c}_e \in \mathbb{R}^n$ такой, что:

$$\mathbf{c}_e = \operatorname{argmin}_{\mathbf{c} \in \mathbb{R}^n} \sum_{\mathbf{x}_i \in \mathbf{X}_e} \rho(\mathbf{x}_i, \mathbf{c}). \quad (2)$$

53 Для нахождения центроида предлагается в качестве расстояния между временны-
54 ми рядами использовать путь наименьшей стоимости [17], найденный методом дина-
55 мической трансформации времени. Псевдокод решения оптимизационной задачи (2)
56 приведён в алгоритме 1.

57 Общая процедура выравнивания имеет вид:

- 58 1) построить множество центроидов классов $\{\mathbf{c}_e\}_{e=1}^K$;
- 59 2) по множеству центроидов найти пути наименьшей стоимости между каждым
60 временным рядом \mathbf{x}_i и центроидом его класса \mathbf{c}_{y_i} ;

Алгоритм 1 Нахождение центроида. DBA(\mathbf{X}_e, n_iter).

Вход: \mathbf{X}_e — множество временных рядов, принадлежащих одному и тому же классу,
 n_iter — количество итераций алгоритма.

Выход: \mathbf{c} — центроид множества \mathbf{X}_e .

- 1: задать начальное приближение приближение центроида \mathbf{c} ;
- 2: **для** $i = 1, \dots, n_iter$
- 3: **для** $\mathbf{x} \in \mathbf{X}_e$
- 4: вычислить выравнивающий путь между \mathbf{c} и \mathbf{x}
 $\text{alignment}(\mathbf{x}) := \text{DTWalignment}(\mathbf{c}, \mathbf{x})$;
- 5: объединить поэлементно множества индексов для каждого отсчёта времени
 $\text{alignment} := \bigcup_{\mathbf{x} \in \mathbf{X}_e} \text{alignment}(\mathbf{x})$;
- 6: $\mathbf{c} = \text{mean}(\text{alignment})$

ПРОЦЕДУРА DTWalignment(\mathbf{c}, \mathbf{x})

Вход: \mathbf{c}, \mathbf{x} — временные ряды.

Выход: alignment — выравнивающий путь. // *каждый индекс временного ряда \mathbf{x} поставлен в однозначное соответствие индексу временного ряда \mathbf{c}*

- 1: построить $n \times n$ -матрицу деформаций DTW
 $\text{cost} := \text{DTW}(\mathbf{c}, \mathbf{x})$;
- 2: вычислить выравнивающий путь по матрице деформаций
 $\text{alignment} := \text{DTWpath}(\text{cost})$;

61 3) по каждому пути восстановить выравненный временной ряд;

62 4) привести множества выравненных временных рядов к нулевому среднему и нор-
63 мировать на дисперсию.

64 Результатом выравнивания является множество выравненных временных рядов.

Метрическое обучение. Введём на множестве выравненных временных рядов расстояние Махаланобиса

$$d_{\mathbf{A}}(\mathbf{x}_i, \mathbf{x}_j) = \sqrt{(\mathbf{x}_i - \mathbf{x}_j)^{\top} \mathbf{A} (\mathbf{x}_i - \mathbf{x}_j)},$$

65 где матрица трансформаций $\mathbf{A} \in \mathbb{R}^{n \times n}$ является симметричной и неотрицательно
66 определенной ($\mathbf{A}^{\top} = \mathbf{A}$, $\mathbf{A} \succeq 0$). Представим матрицу \mathbf{A} в виде разложения $\mathbf{A} = \mathbf{L}^{\top} \mathbf{L}$.
67 Матрица $\mathbf{L} \in \mathbb{R}^{p \times n}$ — матрица линейного преобразования, где p задаёт размерность

68 преобразованного пространства. Если параметр $p < n$, то происходит снижение раз-
 69 мерности признакового пространства.

Расстояние $d_{\mathbf{A}}(\mathbf{x}_i, \mathbf{x}_j)$ есть евклидово расстояние между $\mathbf{L}\mathbf{x}_i$ и $\mathbf{L}\mathbf{x}_j$

$$d_{\mathbf{A}}(\mathbf{x}_i, \mathbf{x}_j) = \sqrt{(\mathbf{x}_i - \mathbf{x}_j)^\top \mathbf{L}^\top \mathbf{L} (\mathbf{x}_i - \mathbf{x}_j)} = \sqrt{(\mathbf{L}(\mathbf{x}_i - \mathbf{x}_j))^\top (\mathbf{L}(\mathbf{x}_i - \mathbf{x}_j))} = \|\mathbf{L}(\mathbf{x}_i - \mathbf{x}_j)\|_2.$$

В качестве алгоритма метрического обучения в данной работе был выбран алгоритм LMNN. Данный алгоритм сочетает в себе идеи метода k ближайших соседей. Первая идея заключается в минимизации расстояний между k ближайшими объектами, находящимися в одном классе. Запишем функционал качества в виде

$$Q_1(\mathbf{L}) = \sum_{j \rightsquigarrow i} \|\mathbf{L}(\mathbf{x}_i - \mathbf{x}_j)\|^2 \rightarrow \min_{\mathbf{L}},$$

где $j \rightsquigarrow i$ означает, что \mathbf{x}_j является одним из k ближайших соседей для \mathbf{x}_i . Вторая идея состоит в максимизации расстояния между каждым объектом и его объектами-нарушителями. Объектом-нарушителем для \mathbf{x}_i назовём объект \mathbf{x}_l такой, что

$$\|\mathbf{L}(\mathbf{x}_i - \mathbf{x}_l)\|^2 \leq \|\mathbf{L}(\mathbf{x}_i - \mathbf{x}_j)\|^2 + 1, \quad \text{где } j \rightsquigarrow i. \quad (3)$$

Тем самым необходимо минимизировать следующий функционал

$$Q_2(\mathbf{L}) = \sum_{j \rightsquigarrow i} \sum_l (1 - y_{il}) [1 + \|\mathbf{L}(\mathbf{x}_i - \mathbf{x}_j)\|^2 - \|\mathbf{L}(\mathbf{x}_i - \mathbf{x}_l)\|^2]_+ \rightarrow \min_{\mathbf{L}},$$

70 где $y_{il} = 1$, если $y_i = y_l$, и $y_{il} = 0$ в противном случае. Положительная срезка позволяет
 71 штрафовать только те объекты, которые удовлетворяют условию (3).

Задача метрического обучения состоит в нахождении линейного преобразования $\mathbf{f}(\mathbf{x}) = \mathbf{L}\mathbf{x}$, то есть нахождении матрицы \mathbf{L} в виде решения оптимизационной задачи:

$$Q(\mathbf{L}) = \mu Q_1(\mathbf{L}) + (1 - \mu) Q_2(\mathbf{L}) \rightarrow \min_{\mathbf{L}}, \quad (4)$$

72 где $\mu \in (0, 1)$ — весовой параметр, определяющий вклад каждого из функционалов.
 73 Задача (4) представляет собой задачу полуопределенного программирования [18] и
 74 может быть решена существующими оптимизационными пакетами.

Классификация временных рядов. Пусть $\mathbf{x} \in \mathbf{X}$ — неразмеченный временной ряд. Выравним временной ряд \mathbf{x} относительно всех центроидов классов

$$\hat{\mathbf{x}}_e = G(\mathbf{x}, \mathbf{c}_e), \quad \text{где } e = \{1, \dots, K\}.$$

Отнесём временной ряд к классу, для которого минимально расстояние до соответствующего центроида. В качестве расстояния используем обученную метрику Махаланобиса с фиксированной матрицей \mathbf{A}

$$\hat{y} = \operatorname{argmin}_{e \in \{1, \dots, K\}} d_{\mathbf{A}}(\hat{\mathbf{x}}_e, \mathbf{c}_e).$$

75 После нахождения оптимальных центроидов классов и нахождения оптимальной мат-
76 рицы трансформаций, процедура классификации заключается в измерении расстоя-
77 ния между найденными центроидами и новыми неразмеченными объектами.

Для оценки качества работы алгоритма будем вычислять ошибку классификации как долю неправильно классифицированных объектов тестовой выборки \mathcal{U}

$$\text{error} = \frac{1}{|\mathcal{U}|} \sum_{i=1}^{|\mathcal{U}|} [a(\mathbf{x}_i) \neq y_i].$$

78 3 Вычислительный эксперимент

79 Цель вычислительного эксперимента — проверить работоспособность предложенного
80 подхода. Предполагается, что построенный алгоритм мультиклассовой классифика-
81 ции способен определить тип активности человека по форме сигнала акселерометра
82 мобильного телефона.

83 Для проведения базового вычислительного эксперимента были подготовлены син-
84 тетические временные ряды, принадлежащие двум классам [19]. Первый класс — си-
85 нусы вида $\sin(x + b)$, где параметр b определяет сдвиг каждого временного ряда.
86 Второй класс — пилообразные функции с различными сдвигами по временной шкале.
87 На каждый временной ряд был наложен нормальный шум. Число временных рядов
88 каждого класса = 60. Длина каждого временного ряда $n = 50$.

89 Построенные центроиды классов проиллюстрированы на рис. 1. Из рисунка видно,
90 что процедура корректно определяет сдвиги временных рядов.

91 Для того чтобы убедиться в целесообразности применения метрического обучения,
92 данные временные ряды классифицировались в пространстве с евклидовой метрикой
93 и в пространстве с метрикой Махаланобиса. Число ближайших соседей $k = 5$, размер-
94 ность преобразованного пространства $p = 40$. Полученные ошибки классификации
95 составили:

96 евклидова метрика: 27%

97 метрика Махаланобиса: 6%.

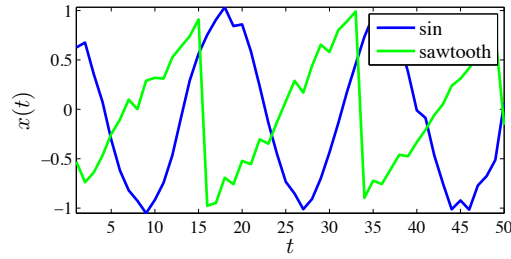


Рис. 1: Центроиды синтетических временных рядов

98 Реальные данные [20] представляли собой временные ряды акселерометра мобиль-
 99 ного телефона. Каждый из шести классов соответствует определенной физической
 100 активности испытуемых. Для проведения вычислительного эксперимента было вы-
 101 брано по 200 объектов каждого класса. Длина каждого временного ряда равна $n = 128$
 102 отсчетам времени.

103 Построенные центроиды классов изображены на рис. 2. Найденные центроиды об-
 104 ладают периодичностью, свойственной временным рядам показаний активности че-
 ловека. На рис. 3 показаны примеры временных рядов каждого класса. Эти же вре-

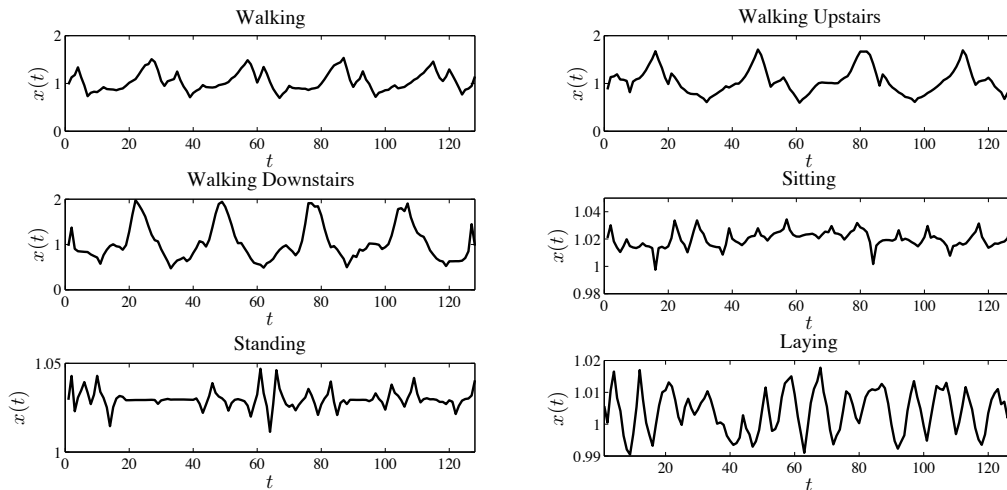


Рис. 2: Центроиды временных рядов акселерометра

105
 106 менные ряды после процедуры выравнивания относительно построенных центроидов
 107 изображены на рис. 4.

108 Ошибка классификации без использования метрического обучения составила 37,5%.
 109 Алгоритм LMNN позволяет настроить параметры: число ближайших соседей k , раз-
 110 мерность преобразованного евклидова пространства p . Для выбора оптимальных па-

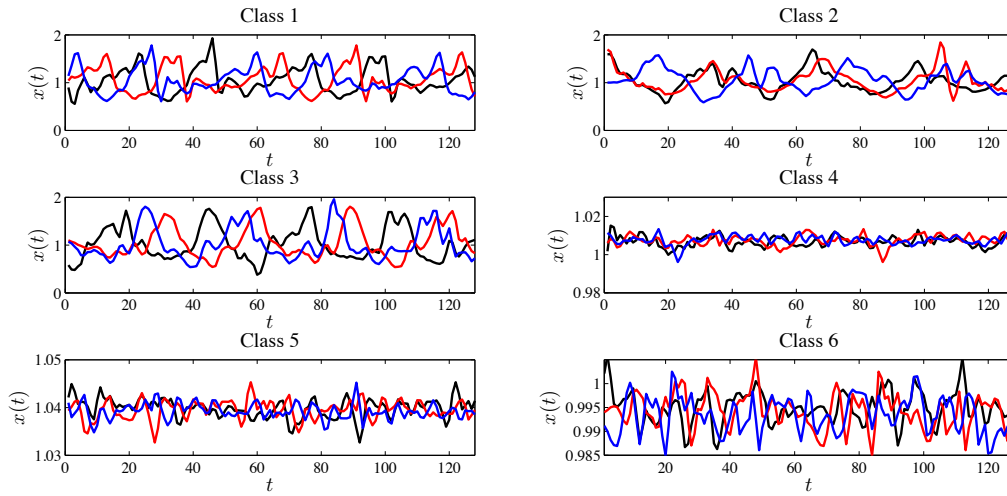


Рис. 3: Временные ряды акселерометра

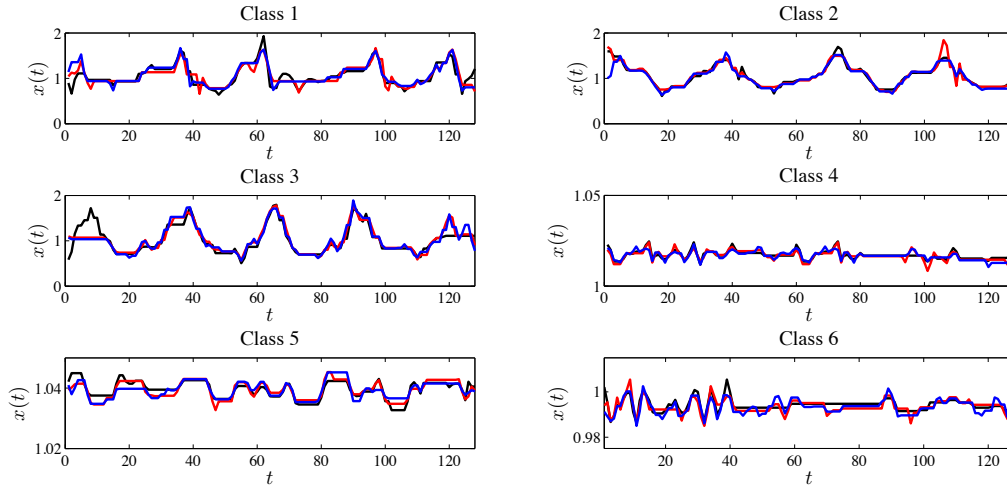


Рис. 4: Выравненные временные ряды акселерометра

111 раметров воспользуемся процедурой кросс-проверки. На рис. 5 цветом показана ошиб-
 112 ка классификации алгоритма в зависимости от параметров алгоритма. На данной
 113 выборке алгоритм LMNN оказывается слабо чувствителен к количеству ближайших
 114 соседей, при этом при уменьшении размерности пространства объектов ошибка клас-
 115 сификации растёт.

116 Настроим алгоритм LMNN со следующими параметрами: число ближайших сосе-
 117 дей $k = 30$, размерность выходного пространства $p = 128$. Ошибка классификации

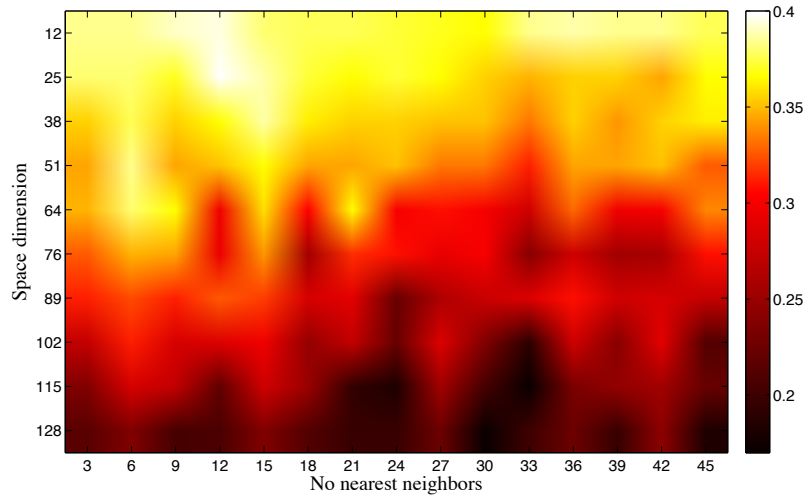


Рис. 5: Ошибка классификации в зависимости от параметров

118 составила 17.25%, что вдвое меньше ошибки классификации с использованием евкли-
 119 довой метрики.

Таблица 1: Матрицы несоответствий

(a) Евклидова метрика

	Истинные метки классов					
	1	2	3	4	5	6
1	80	0	5	0	0	0
2	4	56	33	0	0	0
3	5	5	86	0	0	0
4	7	8	5	168	4	21
5	51	61	57	12	192	11
6	53	70	14	20	2	168

(b) Метрика Махаланобиса

	Истинные метки классов					
	1	2	3	4	5	6
1	151	12	13	0	0	0
2	10	142	14	0	0	0
3	9	10	171	0	0	0
4	10	7	0	173	9	21
5	2	11	0	12	186	9
6	18	18	2	15	5	170

120 В таблице 1 изображены матрицы несоответствий результатов классификации при
 121 использовании евклидовой метрики и метрики Махаланобиса. Столбцы соответствую-
 122 т истинным меткам классов объектов, строки — предсказанным меткам. Диагональ-
 123 ное преобладание матрицы несоответствий указывает на высокую предсказательную
 124 способность алгоритма.

125 В таблице 2 продемонстрировано увеличение точности классификации при исполь-
 126 зовании в качестве меры расстояния метрику Махаланобиса. Пересечение i -го столбца

127 и j -й строки отвечает изменению доли объектов класса i , отнесенных к классу j . Поло-
 128 жительное суммарное значение диагональных элементов таблицы соответствует уве-
 129 личению качества классификации. Значительное улучшение предсказания происходит
 130 при классификации первых трех классов. Данные классы соответствуют следующим
 131 видам физической активности: ходьба, ходьба вверх, ходьба вниз.

Таблица 2: Увеличение точности классификации при использовании адекватной оцен-
 ки матрицы трансформаций

	Истинные метки классов					
	1	2	3	4	5	6
1	0,355	0,06	0,04	0	0	0
2	0,03	0,43	-0,095	0	0	0
3	0,02	0,025	0,425	0	0	0
4	0,015	-0,005	-0,025	0,025	0,025	0
5	-0,245	-0,25	-0,28	0	-0,03	-0,01
6	-0,175	-0,26	-0,06	-0,025	0,005	-0,01

132 4 Заключение

133 В данной работе предложен новый подход к решению задачи многоклассовой клас-
 134 сификации временных рядов. Сравнивались результаты классификации множества
 135 временных рядов, основанных на измерении расстояний с помощью евклидовой мет-
 136рики и обученной метрики Махаланобиса. Проведён вычислительный эксперимент на
 137 реальных данных показаний акселерометра мобильного телефона. Построенная мо-
 138дель классификации показала высокое качество распознавания активности человека
 139 по форме сигнала акселерометра.

Список литературы

- [1] Попова М. С., В. В. Стрижов. 2015. Выбор оптимальной модели классификации физической активности по измерениям акселерометра. *Информатика и её применения*, 9(1):79–89.
- [2] Ignatov A. D., V. V. Strijov. 2015. Human activity recognition using quasiperiodic time series collected from a single tri-axial accelerometer. *Multimedia Tools and Applications*, 1–14.
- [3] Güler, I., E. D. Übeyli. 2007. Multiclass support vector machines for eeg-signals classification. *Information Technology in Biomedicine, IEEE Transactions on*, 11(2):117–126.
- [4] Übeyli, E. D. 2007. Ecg beats classification using multiclass support vector machines with error correcting output codes. *Digital Signal Processing*, 17(3):675–684.
- [5] Anand, R., K. Mehrotra, C. K. Mohan, and S. Ranka. 1995. Efficient classification for multiclass problems using modular neural networks. *Neural Networks, IEEE Transactions on*, 6(1):117–124.
- [6] Kafai, M., B. Bhanu. 2012. Dynamic bayesian networks for vehicle classification in video. *Industrial Informatics, IEEE Transactions on*, 8(1):100–109.
- [7] Chaovalitwongse, W. A., Y.-J. Fan, and R. C. Sachdeo. 2007. On the time series k-nearest neighbor classification of abnormal brain activity. *Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on*, 37(6):1005–1016.
- [8] Bellet, A., A. Habrard, and M. Sebban. 2013. A survey on metric learning for feature vectors and structured data. *arXiv preprint arXiv:1306.6709*.
- [9] Yang, L., R. Jin. 2006. Distance metric learning: A comprehensive survey. *Michigan State University*, 2.
- [10] Wang, F., J. Sun. 2015. Survey on distance metric learning and dimensionality reduction in data mining. *Data Mining and Knowledge Discovery*, 29(2):534–564.
- [11] McFee, B., G. R. Lanckriet. 2010. Metric learning to rank. In *Proceedings of the 27th International Conference on Machine Learning (ICML-10)*, 775–782.

- 168 [12] Guillaumin, M., J. Verbeek, and C. Schmid. 2009. Is that you? metric learning
169 approaches for face identification. In *Computer Vision, 2009 IEEE 12th international*
170 *conference on*, 498–505. IEEE.
- 171 [13] Weinberger, K. Q., L. K. Saul. 2008. Fast solvers and efficient implementations
172 for distance metric learning. In *Proceedings of the 25th international conference on*
173 *Machine learning*, 1160–1167. ACM.
- 174 [14] Weinberger, K. Q., J. Blitzer, and L. K. Saul. 2005. Distance metric learning for large
175 margin nearest neighbor classification. In *Advances in neural information processing*
176 *systems*, 1473–1480.
- 177 [15] Berndt, D. J., J. Clifford. 1994. Using dynamic time warping to find patterns in time
178 series. In *KDD workshop*, volume 10, 359–370. Seattle, WA.
- 179 [16] Petitjean, F., A. Ketterlin, and P. Gançarski. 2011. A global averaging method for
180 dynamic time warping, with applications to clustering. *Pattern Recognition*, 44(3):678–
181 693.
- 182 [17] Гончаров А. В., М. С. Попова, В. В. Стрижов. 2015. Метрическая классификация
183 временных рядов с выравниванием относительно центроидов классов. *Системы*
184 *и средства информатики*, 25(4):52–64.
- 185 [18] Vandenberghe, L., and S. Boyd. 1996. Semidefinite programming. *SIAM review*,
186 38(1):49–95.
- 187 [19] Исаченко Р. В. 2015. Реализация алгоритма классификации временных ря-
188 дов // Sourceforge.net. [http://sourceforge.net/p/mlalgorithms/code/HEAD/
189 tree/Group274/Isachenko2015TimeSeries/code/](http://sourceforge.net/p/mlalgorithms/code/HEAD/tree/Group274/Isachenko2015TimeSeries/code/).
- 190 [20] UCI repository. Human Activity Recognition Using Smartphones Dataset.
191 [https://archive.ics.uci.edu/ml/datasets/Human+Activity+Recognition+](https://archive.ics.uci.edu/ml/datasets/Human+Activity+Recognition+Using+Smartphones)
192 [Using+Smartphones](https://archive.ics.uci.edu/ml/datasets/Human+Activity+Recognition+Using+Smartphones).