*Abstract*—**We address the problem of segmenting nearly periodic time series into period-like segments. We introduce a definition of nearly periodic time series via triplets ⟨basic shape, shape transformation, time scaling⟩ that covers a wide range of time series. To split the time series into periods we select a pair of principal components of the Hankel matrix. We then cut the trajectory of the selected principal components by its symmetry axis, thus obtaining half-periods that are merged into segments. We describe a method of automatic selection of periodic pairs of principal components, corresponding to the fundamental periodicity.**

**We demonstrate the application of the proposed method to the problem of period extraction for accelerometric time series of human gait. We see the automatic segmentation into periods as a problem of major importance for human activity recognition problem, since it allows to obtain interpretable segments: each extracted period can be seen as an ultimate entity of gait.**

**The method we propose is more general compared to the application specific methods and can be used for any nearly periodical time series. We compare its performance to classical mathematical methods of period extraction and find that it is not only comparable to the alternatives, but in some cases performs better.**

"Energy" models for continuous domains can be applied to many problems, but often suffer from high computational expense in training, due to the need to repeatedly minimize the energy function to high accuracy. This paper considers a modified setting, where the model is trained in terms of results after optimization is truncated to a fixed number of iterations. We derive "backpropagating" versions of gradient descent, heavy-ball and LBFGS. These are simple to use, as they require as input only routines to compute the gradient of the energy with respect to the domain and parameters. Experimental results on denoising and image labeling problems show that learning with truncated optimization greatly reduces computational expense compared to "full" fitting. http://users.cecs.anu.edu.au/~jdomke/papers/2012aistats.pdf

# Abstract

In regression problems over $\mathbb{R}^d$, the unknown function $f$ often varies more in some coordinates than in others. We show that weighting each coordinate $i$ according to an estimate of the variation of $f$ along coordinate $i$ – e.g. the $L_1$ norm of the $i$th-directional derivative of $f$ – is an efficient way to significantly improve the performance of distance-based regressors such as kernel and $k$-NN regressors. The approach, termed Gradient Weighting (GW), consists of a first pass regression estimate $f_n$ which serves to evaluate the directional derivatives of $f$, and a second-pass regression estimate on the re-weighted data. The GW approach can be instantiated for both regression and classification, and is grounded in strong theoretical principles having to do with the way regression bias and variance are affected by a generic feature-weighting scheme. These theoretical principles provide further technical foundation for some existing feature-weighting heuristics that have proved successful in practice.

We propose a simple estimator of these derivative norms and prove its consistency. The proposed estimator computes efficiently and easily extends to run online. We then derive a classification version of the GW approach which evaluates on real-worlds datasets with as much success as its regression counterpart.

http://jmlr.org/papers/volume17/13-351/13-351.pdf

**Annotation:** The paper provides a guidance on deep learning net construction and optimization using GPU. The paper proposes to use GPU-instances on the cloud platform Amazon Web Services. The problem of time series classification is considered. The paper proposes to use a deep learning net, i.e. a multilevel superposition of models, belonging to the following classes: Restricted Boltzman Machines, autoencoders and neural nets with softmax-function in output. The proposed method was tested on a dataset containing time segments from mobile phone accelerometer. The analysis of relation between classification error, dataset size and superposition parameter amount is conducted.

Many machine learning algorithms can be formulated as the minimization of a training criterion which involves a hyper-parameter. This hyper-parameter is usually chosen by trial and error with a model selection criterion. In this paper we present a methodology to optimize several hyper-parameters, based on the computation of the gradient of a model selection criterion with respect to the hyper-parameters. In the case of a quadratic training criterion, the gradient of the selection criterion with respect to the hyper-parameters is efficiently computed by back-propagating through a Cholesky decomposition. In the more general case, we show that the implicit function theorem can be used to derive a formula for the hyper-parameter gradient involving second derivatives of the training criterion.

**Abstract**: The paper addresses the problem of designing Brain-Computer Interfaces. We solve the problem of feature selection in regression models in application to ECoG-based motion decoding. The task is to predict hand trajectories from the voltage time series of cortical activity. Feature description of a each point resides in spatial-temporal-frequency domain and include the voltage time series themselves and their spectral characteristics. Feature selection is crucial for adequate solution of this regression problem, since electrocorticographic data is highly dimensional and the measurements are correlated both in time and space domains. We propose a multi-way formulation of quadratic programming feature selection (QPFS), a recent approach to filtering-based feature selection proposed by Katrutsa and Strijov, “Comprehensive study of feature selection methods to solve multicollinearity problem according to evaluation criteria”. QPFS incorporates both estimates of similarity between features, and their relevance to the regression problem, and allows an effective way to leverage them by solving a quadratic program. Our modification allows to apply this approach to multi-way data. We show that this modification improves prediction quality of resultant models.

**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.

При решении задач планирования в системах железнодорожного транспорта возникают проблемы связанные с нестационарностью, неравномерностью и высокой зашумленностью данных о грузоперевозках. Для повышения эффективности управления необходимо создание интеллектуальных систем, опирающихся на математические модели, исторические данные и формализованный опыт экспертов. Данная статья посвящена описанию проекта по созданию системы прогнозирования, направленной на повышение качества управления грузовыми железнодорожными перевозками путем выявления взаимосвязи объемов погрузки и спроса на грузовые железнодорожные перевозки с учетом экзогенных факторов.