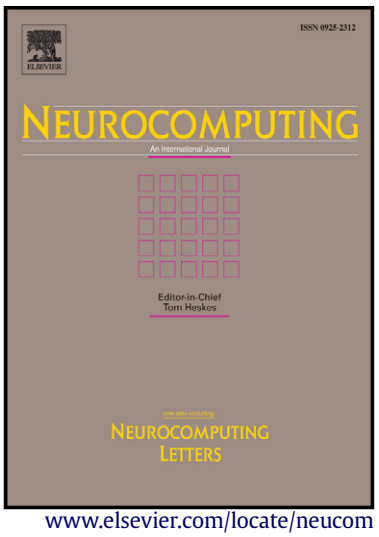
Developments in computational intelligence and machine learning

Biehl, Michael; Ghio, Alessandre; Schleif, Frank-michael

DOI:
10.1016/j.neucom.2015.03.062



www.elsevier.com/locate/neucom

Developments in computational intelligence and machine learning

Michael Biehl, Alessandre Ghio, Frank-Michael Schleif

PII: S0925-2312(15)00383-5
DOI: http://dx.doi.org/10.1016/j.neucom.2015.03.062
Reference: NEUCOM15309

To appear in: *Neurocomputing*

Cite this article as: Michael Biehl, Alessandre Ghio, Frank-Michael Schleif, Developments in computational intelligence and machine learning, *Neurocomputing*, http://dx.doi.org/10.1016/j.neucom.2015.03.062

Editorial

Developments in computational intelligence and machine learning

This special issue of Neurocomputing comprises 17 original articles, which are extended versions of selected papers from the 22*nd* European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning, ESANN 2014.

ESANN is a single-track conference held annually in the medieval town Bruges in Belgium. The fascinating location and inspiring atmosphere facilitates efficient scientific exchange and offers a most enjoyable environment in which to discuss recent developments. The conference is organized by Prof. Michel Verleysen from Universite Catholique de Louvain, Belgium. In addition to regular contributions, the conference welcomes special sessions relating to specific topics of current interest. At ESANN 2014, for instance, special sessions addressed topics like spiking neural networks, learning in real world computing architectures, structured and non-standard data, big data, weightless neural networks, and classification in the presence of label noise.

The 22*nd* ESANN was attended by 145 researchers who presented 114 oral and poster contributions, corresponding to an acceptance rate of 60%. Based on the reviews of the conference papers, taking also into account the quality of the conference presentation and individual recommendations made by special session organizers, a number of authors were invited to submit an extended version of their contribution for this special issue of Neurocomputing. The selected articles were, once more, reviewed independently by at least two expert referees. In the following, we provide a brief overview of the 17 articles that were accepted for publication.

A number of contributions deal with kernel methods, suggesting extensions and modifications of the basic concepts or addressing specific applications:

The paper *Kernel methods for heterogeneous feature selection* by Paul *et al.* introduces two feature selection methods to deal with heterogeneous data that include continuous and categorical variables. A dedicated kernel is suggested that handles both kind of variables within a recursive feature elimination procedure.

In the article by Chrysanthos et *al.* entitled *Theoretical properties and implementation of the one-sided mean kernel for time series*, a novel kernel for time series modeling is introduced. The authors show that it

is positive definite and reduces to a product kernel when comparing two sequences of the same length. The implementation is realized using dynamic programming techniques providing a fast algorithm with low complexity.

In their paper *Scalable, accurate image annotation with joint SVMs and output kernels*, Xiong et *al.* discuss the problem of efficient image annotation by joint training of multiple support vector machines. The authors exploit the dependencies between tags and show how the joint training strategy of SVMs can yield to substantial improvements, in terms of accuracy and efficiency, over training them independently.

Multiple kernel learning (MKL) is the central topic of the contribution by Donini and Aiolli, entitled *Easy MKL: a scalable multiple kernel learning algorithm.* The authors propose a new time and space effective algorithm, that can deal with large amounts of kernels by scaling more efficiently than other state-of-the-art approaches.

Resource aware and energy efficient methods are addressed in two contributions, explicitly:

In *Learning resource-aware models for mobile devices: from regularization to energy efficiency,* Ghio et *al.* consider how modern machine learning approaches can be used in environments with limited ressource by means of regularization and self-adaptation. The authors propose to design a Human Activity Recognition algorithm, which takes into account that only limited resources are available for its execution. They show how the hypothesis space can be restricted by applying advanced concepts from Statistical Learning Theory.

Energy efficiency is also the key topic of the paper *Optimal resource usage in ultra-low-power sensor interfaces through context- and resource-cost-aware machine learning.* Lauwereins et *al.* propose a novel approach to combine the C4.5 decision tree algorithm with adaptive hardware approaches. Only the most relevant features are activated in learning, exploiting adaptive feature extraction circuits and hardware embedded learning.

The analysis of very high-dimensional data requires special measures in practice. A number of articles address the processing of high-dimensional data and their low-dimensional representation:

Lee et *al.* contributed the paper entitled *Multi-scale similarities in stochastic neighbourhood embedding: reducing dimensionality while preserving both local and global structure.* In comparison with standard, single-scale dissimilarities, the suggested approach aims at good embedding quality on all scales. Experimental results illustrate the method and demonstrate its usefulness.

A framework for performing PCA over a network of computing nodes is studied in the contribution by Fellus et *al.*, which is entitled *Asyn-*

chronous gossip principal component analysis. A decentralized algorithm is proposed, analysed mathematically, and studied experimentally. It can be applied to large scale problems and incorporates a dimension reduction step into the so-called *gossip protocol*.

The work of Kerdels et *al.* is focused on the *Analysis of high-dimensional data using local input space histograms*. Local input space histograms were introduced as a means to augment prototype-based vector quantization methods in order to gather more information about the structure of the respective input space. The authors investigate this approach in the context of clustering and visualization of high-dimensional data.

The hubness phenomenon is discussed by Flexer et *al.* in their paper *Choosing l_p norms in high-dimensional spaces based on hub analysis*. Hub objects have a small distance to an exceptionally large number of data points. To avoid the negative effects of distance concentration, the previously suggested use of fractional norms is studied by the authors in an empirical analysis. They propose an unsupervised approach for choosing an l_p norm minimizing hubs while simultaneously maximizing the classification performance.

Frequently, data is not embedded in a conventional vector space, as for instance in the cases of graphs, sequences or other data structures. Specific approaches for the analysis of dissimilarities or proximities are discussed in the following three contributions:

Silvestre et *al.* investigate the use of *Dataset structure as prior information for parameter-free regularization of extreme learning machines.* The suggested regularization approach employs a parameter-free affinity matrix and is evaluated in terms of example classification problems.

The paper *Median variants of Learning Vector Quantization for learning of dissimilarity data* by Nebel et *al.* presents a prototype-based learning algorithm for situations in which labeled data are given in form of a dissimilarity matrix only, without an explicit vector space representation. This is particularly interesting for the analysis of structured data like sequence data or graphs.

The paper entitled *Metric learning for sequences in relational LVQ* by Mokbel et *al.* deals with the integration of low-rank metric adaptation into modern variants of Learning Vector Quantization. The proposed approach scales with both data dimensionality and the cardinality of the considered set, and is sound from a theoretical point of view as well.

Sequential data are also in the center of the contribution *Neural Networks for sequential data: a pre-training approach based on Hidden Markov Models* by Pasa et *al.* The suggested pre-training scheme is applicable in the presence of non-static data. The authors demonstrate the flexibility of the approach by applying it to two

different recurrent network architectures used for example prediction tasks in the context of polyphonic music.

The last three articles further exemplify the wide range of problems and methodological frameworks addressed at the 2014 ESANN conference:

When on-line learning and model interpretability are key issues to cope with, prototype-based models play a central role. In this context, Fischer et *al.* propose *Efficient rejection strategies for prototype-based classification*, allowing to decide whether a reliable decision is possible or, alternatively, a reject is a preferable.

In their contribution *Selective Neural Network ensembles in reinforcement learning: taking the advantage of many agents,* Faußer and Schwenker present an extensive empirical evaluation of a method for selecting subsets of agents from large ensembles. Experimental results with respect to benchmark problems confirm that the method bears the promise to improve performance significantly in comparison with single agents or full ensembles.

Exploiting similarity in system identification tasks with recurrent neural networks is the topic of a contribution by Spieckermann et *al.* The authors introduce a dual-task learning approach based on recurrent networks with factored tensor components. In comparison with competing models, the method displays superior performance with respect to efficiency and accuracy in benchmark problems.

As guest editors we would like to thank all authors for their submissions and the reviewers for their excellent work. The compilation of the special issue imposes a very tight schedule and we thank reviewers and authors for making its publication possible less than a year after ESANN 2014 took place. We also thank the Neurocomputing Editorial Board for giving us the opportunity to compile this issue. The support by Elsevier as a publisher is greatly acknowledged, in particular we would like to express thanks to Chandini Emmanuel and Vera Kamphuis from the Elsevier editorial office.

We thank Prof. M. Verleysen for his support with respect to this special issue and, most importantly, for organizing the outstanding series of ESANN conferences, which enjoys an ever-increasing recognition under his leadership. It is a pleasure for us to invite all interested readers of this special issue to attend future ESANN conferences, which are regularly announced at http://www.elen.ucl.ac.be/esann.

Michael Biehl
University of Groningen
Johann Bernoulli Institute for
Mathematics and Computer Science
P.O. Box 407, 9700 AK Groningen,
The Netherlands
E-mail address: m.biehl@rug.nl

Alessandre Ghio
University of Genoa
Department of Computer Science,

Bioengineering, Robotics and
Systems Engineering (DIBRIS)
Via Opera Pia 13, I-16145 Genoa,
Italy
E-mail address:
alessandro.ghio@unige.it

Frank-Michael Schleif
The University of Birmingham
School of Computer Science
Edgbaston, Birmingham B15 2TT,
United Kingdom
E-mail address:
schleify@cs.bham.ac.uk