ESANN 2017 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium), 26-28 April 2017, i6doc.com publ., ISBN 978-287587039-1.
Available from http://www.i6doc.com/en/.

# Biomedical data analysis in translational research: Integration of expert knowledge and interpretable models

G. Bhanot[1] and M. Biehl[2] and T. Villmann[3] and D. Zühlke[4] *

1- Rutgers University, Department of Molecular Biology and Biochemistry
Piscataway, NJ 08854, USA

2- University of Groningen, Johann Bernoulli Institute for Mathematics and Computer Science, Nijenborgh 9, 9747AG Groningen, The Netherlands

3- University of Applied Sciences, Computational Intelligence Group
Technikumplatz 17, D-09648 Mittweida, Germany

4- Seven Principles AG, Erna-Scheffler-Str. 1A, D-51103 Köln, Germany

**Abstract**. In various fields of biomedical research, the availability of electronic data has increased tremendously. Not only is the amount of disease specific data increasing, but so is its structural complexity in terms of dimensionality, multi-modality and inhomogeneity. Consequently, there is an urgent need for better coordination between bio-medical and computational researchers in order to make an impact on patient care. In any such effort, the integration of expert knowledge is essential. A careful synthesis of good analytical techniques applied to relevant medical questions would make the analysis both accurate and interpretable and facilitate transdisciplinary collaboration. This article summarizes recent challenges and introduces the contributions to this ESANN special session.

## 1 Introduction

New technologies in various fields of biomedical research have led to a dramatic increase in the amount and quality of electronic clinical/medical/biological data. This data is increasingly more accurate, more complex (high-dimensional), multi-modal and inhomogeneous. Modern electronic genomics health records consist of high-dimensional vectors and matrices, as well as questionnaire data, images, text annotations etc. in a variety of formats. The health status and medical history of the patient is reflected in the complex structure of these data components and their interrelation, see [1–7] for example resources in the context of genomics data.

More recently [8, 9], there is a shift towards *personalized medicine.* This is due to the recognition that complex diseases such as diabetes, heart disease or cancer, although they have the same phenotype and look symptomatically the same, have a heterogeneous, patient specific basis [10]. This patient specificity may be due to the patient's genetic/ethnic background, the presence of specific

*The authors thank the *Leibniz Center for Informatics – Schloss Dagstuhl* for the organization and support of Seminar 16261 [14], where the idea for this ESANN 2017 special session was initiated.

ESANN 2017 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium), 26-28 April 2017, i6doc.com publ., ISBN 978-287587039-1. Available from http://www.i6doc.com/en/.

mutations as well as the patient's lifestyle, diet, previous exposure to pathogens, levels of stress etc. There is now a sense that in some cases, rather than treating the patient using information gained from statistical analysis of many patients with the same disease, there is a need for patient specific strategies. The mantra is: *Do not just treat the disease as a single entity, treat the patient specific version of the disease.* This has led to the need for more refined methods to understand the causes of each disease.

Machine learning methods are at the center of these efforts [11, 12]. Their promise is that they will allow a more stratified analysis of disease and a more detailed understanding of the reduced set of variables relevant to treatment. For example, in many cancer centers around the world, there are *Personalized Medicine Initiatives* devoted to the treatment of patients who have failed first and second lines of cancer therapy. In such centers, teams of oncologists, surgeons, radiologists, geneticists, clinical trial experts, and bioinformaticians meet and study, e.g., the mutations in recurrent tumors that have caused relapse, using information from sequencing of patient specific tumors. Such analyses often suggest novel therapies and novel methods of treatment, in some cases resulting in dramatic responses in patients who have failed standard therapies [13].

In this direction, numerous successful machine learning and statistical approaches have been developed to address specific tasks using pattern recognition techniques applied to patient records as well as patient specific SNP (single nucleotide polymorphism), gene and protein expression data. However, many of these approaches neither integrate available expert knowledge systematically into the underlying mathematical and statistical models, nor do they make use of expert (prior) knowledge to improve performance and interpretability.

There is now an urgent and well recognized need to develop, investigate, and apply machine learning and statistical methods for structured biomedical data analysis, in collaborations which include both experts from knowledge representation and integration as well as bio-medical and clinical experts who have a strong interest in developing interpretable models. In other words, the community needs to create forums for the exchange of ideas among practitioners in these disciplines, who in their normal life have little or no opportunity for such dialogues. It is necessary to develop and optimize methods and processing pipelines which offer efficient solutions for a wide range of bio-medical applications; i.e. to develop methodologies which allow the systematic incorporation of domain knowledge into the mathematical approaches used to analyse data. This would provide an incentive for researchers from both medical and analytical backgrounds to collaborate and develop interpretable models which may suggest novel methods of treatment. With these aims in mind, an increasing number of seminars, conferences, workshops, and dedicated sessions have been organized in recent years, including [14] and several special sessions at previous ESANN conferences, e.g. [15–18].

Before summarizing the contributions to the ESANN 2017 special session on *Biomedical data analysis in translational research*, we highlight recent developments and discuss current challenges in the following sections. Obviously, we

ESANN 2017 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium), 26-28 April 2017, i6doc.com publ., ISBN 978-287587039-1.
Available from http://www.i6doc.com/en/.

cannot provide a complete overview or review all relevant literature in this field of ever-growing relevance.

## 2 Important concepts and approaches

In biomedical research and practice, structured data sets play a role of increasing importance, in the form of very complex and inhomogeneous electronic health records. Their analysis, for instance in the context of decision support systems in diagnosis and patient monitoring, requires specific tools for separate modalities of data as well as their integrative treatment in more advanced, unified approaches [19–22]. Modern technological platforms for bio-medical data acquisition and measurement generate data with very high spatial or temporal resolution. Some of these comprise quantities of equal or similar nature in the different dimensions, while in many cases qualitatively disparate features are collected.

Most modern data sets display high structural complexity or dimensionality, which in some cases results in thousands to billions of variables. Very frequently, the number of data dimensions far exceeds the number of samples. These properties result in fundamentally new challenges for automated data processing. Expert knowledge about the origin of the data or the relevant biochemical and bio-medical processes has to be incorporated systematically in order to reduce complexity, yielding regularization restrictions and techniques specifically reflecting the medical context and/or underlying biological processes [14].

In recent years several approaches have been proposed in the context of biomedical data analysis. We provide below some examples of the issues and approaches that seem to be promising, with a few example references as a starting point for further exploration.

- Biologically motivated sparse coding [30] as present in the visual system allows for a compact representation of complex data by adaptive, possibly over-complete basis function systems. This leads to low dimensional representations in functional space, which in turn results in interpretable models. Adaptation of the models can be guided by general information theoretical principles, such that natural regularization conditions arise. For a brief overview of sparse representation techniques in the context of machine learning and further references see [31, 32].

- Relational and median approaches for data classification and clustering provide valid generalization of vector data analysis in the presence of very high-dimensional or non-metric data. Like vector based methods, these approaches offer the possibility to integrate model specific dissimilarity concepts in an intuitive way. Examples of specific methods for relational or non-vectorial data in various scenarios can be found in [32–36].

- Functional representations of data take into account the functional nature of time series or spectral data and make use of their characteristic

ESANN 2017 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium), 26-28 April 2017, i6doc.com publ., ISBN 978-287587039-1.
Available from http://www.i6doc.com/en/.

properties such as continuity and differentiability [37, 38]. In this way, an inherent regularization is achieved which drastically reduces the number of free model parameters for a given data set. Corresponding algorithms and applications are presented in, for instance, [39–41].

- Learning of imbalanced data deals with the complications posed by rare events, which influences the sensitivity and accuracy of models. Standard approaches include over- and undersampling techniques or the use of class-specific cost functions [46, 47]. Specific restrictions due to expert requirements can be incorporated to meet demand specifications, e.g. in terms of the so-called Receiver Operator Characteristics (ROC) [42–44]. In this context, outlier detection, rejection mechanisms and the evaluation of model confidence and certainty play an important role, see [32,42,45–48] for examples.

- Interactive data visualization and analysis are addressed in visual analytics and concern the possibility of employing advanced algorithms for dimensional reduction and the detection of essential information hidden in complex data [49]. Low-dimensional projections facilitate direct human inspection and allow integrating the astonishing cognitive capabilities of human vision into the processing loop, see [49, 50] for a overviews and further references.

- Information optimization based methods use the fundamental entropic principle in nature as the general paradigm for data and information processing, taking into account many orders of correlations. Information theoretic learning techniques are reviewed in [51], which serves as a source for further references.

These approaches and paradigms share a number of fundamental principles of modeling and data analysis. Standard algorithms frequently rely on Euclidean metrics, which is often either not justified or costly to compute for complex high-dimensional data. Alternative dissimilarities, specifically designed for the data domain, need to be addressed, such as correlations, divergences for probabilistic data, structured dissimilarities or pairwise differences [29,54,55]. Other modeling methods often rely on a probabilistic treatment and generative models, e.g. [52,53]. However, density and probability estimation constitute a great challenge in high dimensional spaces and even more so for non-metric data [11, 49, 51].

Frequently, only a combination of technologies leads to success. Furthermore, it is still an open question how to systematically combine different aspects of the data and which methods to use. This question requires collaboration among those with expert knowledge and those with knowledge of algorithms and methods, in each application domain. This knowledge may be integrated as explicit prior information, but frequently requires human interaction during the analysis process. The relevant expert knowledge and experience is usually interdisciplinary and implicit, which requires dedicated processing of the information.

ESANN 2017 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium), 26-28 April 2017, i6doc.com publ., ISBN 978-287587039-1. Available from http://www.i6doc.com/en/.

Here, recommendation schemes and fuzzy decision systems constitute first attempts to tackle these problems. The goal of obtaining interpretable models is closely related to the above mentioned problems [16, 17]. For example, popular systems like *deep* multi-layered neural networks [23, 24] or support vector machines (SVM) [25, 26] do not provide easily accessible information on how the decision is obtained. Furthermore, the model complexity of the resulting systems is often huge compared to sparse systems with similar performance. Thus, conceptual extensions and modifications of *black-box* systems are necessary in order to acquire insight into the underlying decision criteria and to obtain problem adequate complexity [16, 17]. Other systems, for instance Learning Vector Quantization and other prototype-based models [27–29, 55], allow for an intuitive interpretation of the resulting systems and make use of predefined model complexity.

## 3 Relevant research questions and concrete problems

Any systematic attempt to address the problems outlined above requires the characterization of promising paradigms and methods for expert knowledge integration in biomedical data information systems. Efficient and reliable processing pipelines have to be developed in close dialogue with practitioners and domain experts.

In the following we provide a - certainly incomplete - list of relevant questions that need to be answered. These questions can be broken down into five main areas:

**I. Analysis of structured, inhomogeneous and multimodal data**

(a) Which general principles, such as information theory, preservation of inherent data structures or topological properties offer suitable frameworks in which to achieve a compact representation of high dimensional data? How can expert knowledge be integrated systematically in this context?

(b) How can an appropriate level of interpretability of a model be obtained and objectively assessed in the context of the application domain? How can the model complexity be chosen adequately to allow for robust and interpretable results?

(c) Which models are suitable to represent specific biomedical data, such as functional or multimodal data adequately? How should models be evaluated and compared?

(d) What are suitable similarity or dissimilarity measures for structured data? Which families of parameterized measures are available and can be adapted easily? Can the domain experts' implicit concepts of similarity be inferred by interactive systems and incorporated into the models?

ESANN 2017 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium), 26-28 April 2017, i6doc.com publ., ISBN 978-287587039-1. Available from http://www.i6doc.com/en/.

**II. Feature selection and biomarker detection**

(a) How can the inherently non-Euclidean structure of biomedical data be inferred? Which aspects of particular relevance for the application should be emphasized? How can these aspects be integrated into a mathematically consistent description?

(b) How can relevant features or combinations thereof be identified, which are specific to a certain class or cluster of samples, for a particular disease or behavior? To what extent can domain knowledge be formulated and integrated into the analysis, so as to achieve a *semi-guided* feature selection?

(c) How can we identify, select, and verify relevant biomarkers in the exploration of high-dimensional and complex data structures? Can systematic methods for an interactive user-driven selection process be devised?

(d) Are features identified in different model frameworks comparable so that it is possible to identify their true relevance for the task at hand?

**III. Diagnosis and classification**

(a) How can we design learning algorithms and classifiers, which are robust with respect to noise, uncertain labels, missing values, or very imbalanced classes?

(b) How can confidence and certainty of a classifier be estimated reliably? Can we make use of expert knowledge in the design of reject options or alert mechanisms?

(c) How can we improve the sensitivity, specificity of classifiers for reliable diagnosis systems? How can prior knowledge be incorporated to achieve this goal?

(d) How can several classifiers be combined for recommendation or diagnostic systems? Is it possible to employ redundant information for improvement of certainty and/or confidence? How can one deal with contradictory or inconsistent information and classifier decisions?

(e) How can we model temporal aspects of diseases, like stages and disease progression/regression under treatment?

**IV. Generative models of biomedical processes**

(a) Can we formulate appropriate evaluation criteria for the reliability of generative models? Which are stable indicators for model adequacy and accuracy?

(b) How does the interpretability of models interact with their faithfulness and accuracy in relation to expert knowledge?

(c) How can the modelled processes be presented such that humans can judge the reliability and quality of the model? How can, in turn, user feedback be integrated into the models systematically? Can these approaches be used to reduce the level of uncertainty?

ESANN 2017 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium), 26-28 April 2017, i6doc.com publ., ISBN 978-287587039-1. Available from http://www.i6doc.com/en/.

(d) Can generative models contribute to a better understanding of diseases, disease progression and more general biomedical processes? How does one simulate adequate scenarios?

**V. Visualization and visual analytics**

(a) Which visualization techniques are most suitable for heterogeneous and structured data in the biomedical context? How can we visualize relevant features in order to facilitate their interpretation by human experts?

(c) How can one design feedback loops that can improve visualization models by integrating requirements/limitations of domain experts with the model?

(d) How can we effectively visualize patient monitoring and critical events in a longitudinal assessment of patients?

(e) How can we usefully display the errors/uncertainties resulting e.g. from embedding high-dimensional data into low dimensional spaces? Can these be visualized in an intuitive way?

## 4 Contributions to the ESANN 2017 special session on Biomedical data analysis in translational research

The three accepted contributions to the special session present a nice cross section through the relevant research questions discussed in section 3. The authors present a prognostic classification model for a specific clinical application, a novel concept for the analysis of relevance bounds in feature selection, and strategies for the detection of rare diseases from imbalanced data.

Using data in routine clinical work is not just fallow land waiting for data scientists and computational researchers to come and farm it. A plethora of clinical scores based on aggregated biomedical measurements and diagnostic interpretations exist that can inform and compete with novel data analysis methods which aim to integrate into the clinical routine. The authors of [56] challenge two standard clinical prognostic scores which are currently in vogue for the prediction of preterm infant mortality. Not only do they identify additional data supporting the performance of standard clinical scores, they also show how to achieve similar performance using measurement data only. This is important in the given clinical application as the conventional scores can only be determined by a labor-intensive procedure which cannot be performed by all caregiving units.

When biomedical data is analysed, identifying a high performance predictive model for outcome or prognosis is often just the first step. Frequently, it is highly desirable to identify the smallest relevant set of biomarkers that can be used for diagnosis or prognosis in practice. Sometimes, the *relevance values* resulting from this analysis also trigger novel clinical research for targeted therapies. Thus, the precise identification of relevances is crucial. While strongly relevant features, which reduce the discriminative power of the classifier when removed one at a time, are reliably detected by most of the common feature selection methods, the identification of weakly relevant features is challenging.

ESANN 2017 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium), 26-28 April 2017, i6doc.com publ., ISBN 978-287587039-1.
Available from http://www.i6doc.com/en/.

Weak features often are not relevant on their own but become important when combined with other features. They also often lack direct indicators of relevance, e.g. high weights in linear classifiers. This is often the case for highly correlated features. The authors of [57] introduce the concepts of the *minimum and maximum linear relevance bounds* which provide more information for the identification of strongly and weakly relevant features.

Another important challenge for computational researchers in the biomedical domain are imbalanced classes. Typically, imbalanced data sets relate to rare diseases, where only very few patients are known, while healthy controls are available in abundance. In [58], a number of strategies to circumvent the shortcomings resulting from very imbalanced data sets is studied. Inborn steroidogenic disorders serve as an example for a relevant diagnostic challenge. Undersampling, probably the most commonly used strategy, performed worse than a baseline given by classifiers specifically designed for missing data but not tailored to imbalanced classes. The considered oversampling strategies outperform undersampling, but their performance is similar to the baseline results. The introduction of suitable, specific costs for the different types of misclassification appears to be a most promising strategy. This supports the general claim that incorporating expert knowledge at the right point can enhance computational models significantly.

## References

[1] A. Kundaje, *Datasets,* url: http://sites.google.com/site/anshulkundaje/idatasets (last accessed: 02/16/2017)
[2] National Center for Biotechnology Information *Gene Expression Omnibus,* url: http://www.ncbi.nlm.nih.gov/geo (last accessed: 02/16/2017)
[3] European Bioinformatics Institue (EMBL-EBI), *Expression Atlas,* url: http://www.ebi.ac.uk/gxa/hme (last accessed: 02/16/2017)
[4] National Cancer Insitute, *The Cancer Genome Atlas (TCGA),* url: http://cancergenome.nih.gov and http://gdac.broadinstitute.org (last accessed: 02/16/2017)
[5] The GTEx consortium, *GTEx portal,* url: http://www.gtexportal.org/home (last accessed: 02/16/2017)
[6] European Bioinformatics Institute, *The International Genome Sample Resource,* url: http://www.internationalgenome.org (last accessed: 02/16/2017)
[7] NIH U.S. National Library of Medicine, *National Information Center on Health Services Research and Health Care,* url: https://www.nlm.nih.gov/hsrinfo/datasites.html and https://www.nlm.nih.gov/services/databases_abc.html (last accessed: 02/16/2017)
[8] J. Mendelsohn, T. Tursz, R.L. Schilsky, V. Lazar, WIN Consortium - challenges and advances, *Nature Reviews Clinical Oncology,* 8: 133-134, 2011.
[9] I.I. Wistuba, J.G. Gelovani, J.J. Jacoby, S.E. Davis, R.S. Herbst, Methodological and practical challenges for personalized cancer therapies, *Nature Reviews Clinical Oncology,* 8: 135-141, 2011.
[10] R. Hodson, Precision medicine, *Nature,* 537: S49, 2016
[11] T. Hastie, R. Tibshirani, J. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction,* Springer, 2009.
[12] C. Bishop, *Pattern Recognition and Machine Learning,* Cambridge University Press, Cambridge, 2007.
[13] J.M. Mehnert, A. Panda, H. Zhong, K. Hirshfield, S. Damare, K. Lane, L. Sokol, M.N. Stein, L. Rodriguez-Rodriquez, H.L. Kaufman, S. Ali, J.S. Ross, D.C. Pavlick, G. Bhanot, E.P. White, R.S. DiPaola, A. Lovell, J. Cheng, S. Ganesan, Immune activation and re-

sponse to pemprolizumab in POLE-mutant endometrial cancer, *J Clin Invest.,* 126(6): 2334-2340, 2016.

[14] G. Bhanot, M. Biehl, T. Villmann, D. Zühlke, editors, Integration of expert knowledge for interpretable models in biomedical data analysis (Seminar 16261), Schloss Dagstuhl – Leibniz-Zentrum für Informatik, *Dagstuhl Reports,* 6: 88-110, 2016.

[15] V. Bolón-Canedo, B. Remeseiro, A. Alonso-Betanzos, A. Campilho, Machine learning for medical applications. In: M. Verleysen, editor, *Proceedings of the European Symposium on Artificial Neural Networks* (ESANN 2016), d-side pub., pages 225-234, 2016.

[16] V. Van Belle, P. Lisboa, Research directions in interpretable machine learning models. In: M. Verleysen, editor, *Proceedings of the European Symposium on Artificial Neural Networks* (ESANN 2013), d-side pub., pages 533-541, 2013.

[17] A. Vellido, J.D. Martín-Guerro, P. Lisboa, Making machine learning models interpretable. In: M. Verleysen, editor, *Proceedings of the European Symposium on Artificial Neural Networks* (ESANN 2010), d-side pub., pages 163-172, 2012.

[18] P. Lisboa, A. Vellido, J.D. Martín, Computational Intelligence in biomedicine: Some contributions. In: M. Verleysen, editor, *Proceedings of the European Symposium on Artificial Neural Networks* (ESANN 2010), d-side pub., pages 429-438, 2010.

[19] S.J. Pan, Q. Yang, A survey on transfer learning, *IEEE Trans. Knowl. Data Eng.,* 22(10): 1345-1359, 2010.

[20] V. Vapnik, A. Vashist, A new learning paradigm: Learning using privileged information, *Neural Networks,* 22(5-6): 544-557, 2009.

[21] J. Feyereisi, U. Aickelin, Privileged information for data clustering, *Information Sciences,* 194: 4-23, 2012.

[22] E. Mwebaze, G. Bearda, M. Biehl, D. Zühlke, Combining dissimilarity measures for prototype-based classification, In: M. Verleysen, editor, *Proceedings of the European Symposium on Artificial Neural Networks* (ESANN 2015), d-side pub., pages 31-36, 2015.

[23] I. Goodfellow, Y. Bengio, A. Courville, *Deep Learning,* MIT Press, 2016.

[24] Y. LeCun, Y. Bengio, G. Hinton, Deep Learning, *Nature,* 521: 436-444, 2015.

[25] C. Cortes, V. Vapnik, Support vector network, *Machine Learning,* 20: 1-20, 1995.

[26] N. Cristianini, J. Shawe-Taylor, *An introduction to support vector machines and other kernel-based learning methods,* Cambridge University Press, 2000.

[27] T. Kohonen, *Self-Organizing Maps,* Springer, 1997.

[28] D. Nova, P.A. Estevez, A review of learning vector quantization classifiers, *Neural Computing and Applications,* 25(3-4): 511-524, 2014.

[29] M. Biehl, B. Hammer, T. Villmann, Prototype-based models in machine learning, *Wileys Interdisciplinary Reviews (WIREs) Cognitive Science,* 7(2): 92-111, 2016.

[30] B. Olshausen, D. Field, Emergence of simple-cell receptive field properties by learning a sparse code for natural images. *Nature,* 381: 606-609, 1996.

[31] T. Villmann, F.-M. Schleif, B. Hammer, Sparse representation of data. In: M. Verleysen, editor, *Proceedings of the European Symposium on Artificial Neural Networks* (ESANN 2010), d-side pub., pages 225-243, 2010.

[32] F.-M. Schleif, X. Zhu, B. Hammer, Sparse conformal prediction for dissimilarity data. *Annals of Mathematics and Artificial Intelligence,* 74(1-2): 95-116, 2015.

[33] X. Zhu, F.-M. Schleif, B. Hammer, Semi-Supervised Vector Quantization for proximity data. In: M. Verleysen, editor, *Proceedings of the European Symposium on Artificial Neural Networks* (ESANN 2013), d-side pub., pages 89-94, 2013.

[34] D. Nebel, B. Hammer, T. Villmann, A Median Variant of Generalized Learning Vector Quantization. In: M. Lee, A. Hirose, Z.-G. Hou, R.M. Kil (eds.), *Neural Information Processing: 20th International Conference ICONIP 2013,* part II, 19-26, Springer, 2013.

[35] E. Pekalska, R.P.W. Duin, *The Dissimilarity Representation for Pattern Recognition: Foundations and Applications,* World Scientific, 2006.

[36] R.J. Hathaway, J.W. Davenport, J.C. Bezdek, Relational duals of the c-means clustering algorithms, *Pattern Recognition,* 22(2): 205-212, 1989.

[37] P. Geurts, *Pattern recognition for time series classification,* In: *Europ. conference on principles of data mining and knowledge discovery,* Springer, pp. 115-127, 2001.

[38] J. Ramsay, B. Silverman, *Functional Data Analysis*, Springer, 2006.

[39] M. Kästner, B. Hammer, M. Biehl, T. Villmann, Generalized functional relevance learning