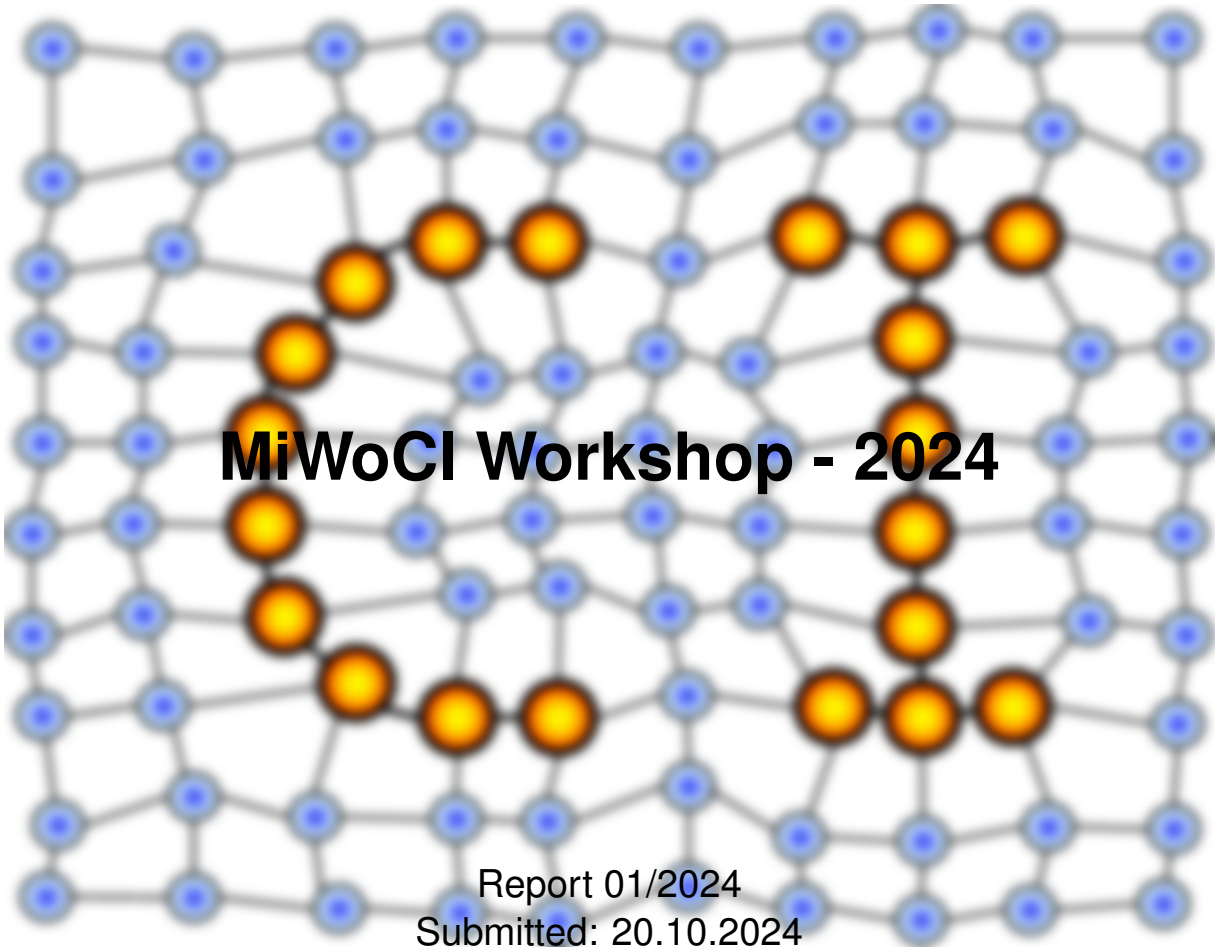
MACHINE LEARNING REPORTS



MiWoCI Workshop - 2024

Report 01/2024
Submitted: 20.10.2024
Published: 20.10.2024

Frank-Michael Schleif[1,2*,3*], Sven Hellbach[4], Marika Kaden[2], Thomas Villmann[2] (Eds.)
(1) Technical University of Applied Sciences Wuerzburg-Schweinfurt,
Sanderheinrichsleitenweg 20, 97074 Wuerzburg, Germany (2) University of Applied Sciences
Mittweida, Technikumplatz 17, 09648 Mittweida, Germany (3) University of Birmingham,
School of Computer Science,
Edgbaston, B15 2TT Birmingham, UK
(4) University of Applied Sciences Zwickau, Kornmarkt 1, 08012 Zwickau, Germany

Impressum

Publisher:	University of Applied Sciences Mittweida Technikumplatz 17, 09648 Mittweida, Germany
Editor:	Prof. Dr. Thomas Villmann Prof. Dr. Frank-Michael Schleif
Technical-Editor:	Prof. Dr. Frank-Michael Schleif
Contact:	frank-michael.schleif@fhws.de
URL:	http://techfak.uni-bielefeld.de/∼ fschleif/mlr/mlr.html
ISSN:	1865-3960

Abstracts of the 16^{th} Mittweida Workshop on Computational Intelligence - MiWoCI 2024 -

Frank-Michael Schleif, Sven Hellbach, Marika Kaden, and Thomas Villmann

Machine Learning Report 01/2024

Preface

From 10. to 12 July 2024 we had the pleasure to organize and attend the 16 th Mittweida Workshop on Computational Intelligence (MiWoCi 2024) as a satellite event of 15 th Workshop on Self Organizing Maps (WSOM+ '24). Thus the tradition of scientific presentations, vivid discussions, and exchange of novel ideas at the cutting edge of research was continued. They were connected to diverse topics in computer science, automotive industry, and machine learning.

This volume contains a collection of extended abstracts and short papers which accompany some of the discussions and presented posters of the MiWoCi Workshop, which cover theoretical aspects, applications, as well as strategic developments in the fields.

Apart from the scientific merrits, this year?s seminar came up with the great chance to attend the 15 th Workshop on Self Organizing Maps and Beyond (WSOM+ '24). WSOM+ is the major anchor conference focusing on Self Organizing Maps and is not only a perfect chance to met high renowned researchers in the field but also to attend the three invited plenaray talks given during WSOM 2024:

- John Aldo Lee - Université catholique de Louvain, Belgium
- Peter Tino - University of Birmingham, United Kingdom
- Barbara Hammer - University Bielefeld, Germany

This year the MiWoCi Workshop was also accompanied by a poster spotlight at the WSOM 2024 for each poster contribution. Our particular thanks for a perfect local organization of the workshop go to Thomas Villmann and his team as spiritus movens of the seminar and his PhD and Master students.

Contents

Iterated Relevance Matrix Analysis (IRMA): Improved Robustness and Interpretability of Feature Relevances

Elina L. van den Brandhof[1,2a], Sofie Lövdal[1,2b],James M. Hawley[3,4], Lorna C. Gilligan [5], Angela E. Taylor[6], Wiebke Arlt[3,4], and Michael Biehl[1,5]

1 - Univ. of Groningen - Bernoulli Institute for Mathematics, Computer Science and Artificial Intelligence, Groningen, NL
2- Univ. Medical Center Groningen (UMCG), NL, (a) Dept. of Neurology, Groningen, (b) Dept. of Nuclear Medicine and Molecular Imaging
3 - Medical Research Council Laboratory of Medical Sciences, London, UK
4 - Imperial College, Inst. of Clinical Sciences, Fac. of Medicine, London, UK
5 - Univ. of Birmingham, Inst. of Metabolism and Systems Research, UK

Abstract

Iterated Relevance Matrix Analysis (IRMA) [1] involves the iterative re-training of Generalized Matrix Learning Vector Quantization (GMLVQ) systems [2]. In each step, previously identified discriminative directions are projected out and the iteration proceeds until no further class-specific information can be found. As an illustrative example problem, we apply IRMA to the classification of adrenocortical tumors based on steroid metabolomics data [3, 4]. In order to demonstrate the usefulness of the IRMA approach, we consider training from very small and imbalanced data sets. Extending the basic idea of [1], we suggest the construction of a combined relevance profile from all discriminative IRMA iterations. The combined relevances are shown to be more robust against variations in the training set and class imbalance than original GMLVQ. In addition, we design a distance based classifier using the combined relevance matrix and compare its performance with that of unmodified GMLVQ when applied to previously unseen, prospective data.

References

[1] S. Lövdal and M. Biehl. Iterated Relevance Matrix Analysis (IRMA) for the identification of class-discriminative subspaces. Neurocomputing, Art. No. 127367, 2024.

[2] P. Schneider, M. M. Biehl and B. Hammer. Adaptive relevance matrices in Learning Vector Quantization. Neural Computation, 21(12):3532-3561, 2009.

[3] W. Arlt, M. Biehl, A. Taylor et al. Urine steroid metabolomics as a biomarker tool for detecting malignancy in adrenal tumors. J. Clin. Endocrinol. Metab. 96(12):3775-84, 2011.

[4] I. Bancos, A.E. Taylor, V. Chortis et al. Urine steroid metabolomics for the differential diagnosis of adrenal incidentalomas in the EURINE-ACT study: a prospective test validation study. The Lancet Diabetes & Endocrinology, 8(9):773-781, 2020.

Rediscovering Chaos? Analysis of GPU Computing Effects in Graph-coupled NeuralODEs

Simon Heilig

Friedrich-Alexander University Erlangen-Nürnberg, M.Sc. Data Science
simon99.heilig@gmail.com

1 Motivation

In physical systems that can be described in terms of (non-)linear differential equations, future states can be deterministically predicted just by knowing the past states. If the equations can not be solved analytically by means of the Fourier transformation, numerical schemes integrate the systems with finite precision. Among the first known systems that exhibit high sensitivity to the initial state, are the 3-body problem or Lorenz' system of three first order nonlinear coupled equations [8]. This means, that different initial states within a small neighborhood can posses divergent trajectories. These perturbations are inevitable with numerical schemes that rely on finite floating point precision.

Recent works on the conjoining of dynamical systems and neural networks focus on so-called neuralODEs [1], i.e., dynamical systems parameterized with neural networks. It can be shown that nonlinear time-dependent systems generalize the forward pass in neural networks with residual connections, e.g., ResNets derived from a continuous ODE [5]:

$$\frac{\partial \mathbf{x}(t)}{\partial t} = \sigma(\mathbf{W}(t)\mathbf{x}(t) + \mathbf{b}(t)), \tag{1}$$

that can be discretized by the forward Euler scheme ($\frac{\partial \mathbf{x}(t)}{\partial t} \approx \frac{\mathbf{x}(t+\epsilon)-\mathbf{x}(t)}{\epsilon}$) to result in $\mathbf{x}^{(n+1)} = \mathbf{x}^{(n)} + \epsilon\sigma(\mathbf{W}^{(n+1)}\mathbf{x}^{(n)} + \mathbf{b}^{(n+1)})$. Here, $\mathbf{W} \in \mathbb{R}^{d\times d}, \mathbf{b} \in \mathbb{R}^d$ are the parameters of the system with state $\mathbf{x} \in \mathbb{R}^d$ at time $t = \epsilon n \geq 0$ and σ is a nonlinear activation function, e.g., tanh or ReLU.

Deterministic chaos can play a central role in graph-coupled neuralODEs [7] since they represent nonlinear coupled systems, a necessary condition for chaos [8]. When integrating such systems, which can be seen as the forward pass of a layer-wise graph neural network, numerical errors of the integration method, finite floating point precision and non-deterministic GPU operations influence the possibly chaotic trajectory of the state. Hence, this work is devoted to experimentally investigate these compounding effects and to assess the reproducibility of results when computing with GPUs.

2 Graph-coupled NeuralODEs

The solution of the initial value problem (IVP) formulated in [7]:

$$\begin{aligned} \frac{\partial \mathbf{X}(t)}{\partial t} &= \sigma(\mathbf{X}(t)\mathbf{W}(t) + \mathbf{A}\mathbf{X}(t)\mathbf{V}(t) + \mathbf{b}(t)) \\ \mathbf{X}(0) &= \mathbf{X}^0 \end{aligned} \quad , t \in [0, T], \tag{2}$$

is a continuous perspective on the layer-wise message-passing framework [3] for graph neural networks. Here, $\mathbf{X} \in \mathbb{R}^{n \times d}$ is the feature matrix of n nodes in an undirected graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ with node set $\mathcal{V}$ and edge set $\mathcal{E}$ summarized in the adjacency matrix $\mathbf{A} \in \{0, 1\}^{n \times n}$. The output $\mathbf{X}(T)$ at the terminal time $T > 0$ is taken as embedding of the nodes for computing downstream tasks like link prediction, node classification, or even global graph regression [6]. Hence, the dynamics in Eq. (2) are seen as an information processing system on a graph.

Applying the forward Euler scheme to Eq. (2), a residual message-passing neural network is obtained:

$$\mathbf{x}_u^{(n+1)} = \mathbf{x}_u^{(n)} + \epsilon\sigma(\mathbf{W}^{(n+1)}\mathbf{x}_u^{(n)} + \sum_{v \in \mathcal{N}_u} \mathbf{V}^{(n+1)}\mathbf{x}_v^{(n)} + \mathbf{b}^{(n+1)}), \quad \forall u \in \mathcal{V}. \tag{3}$$

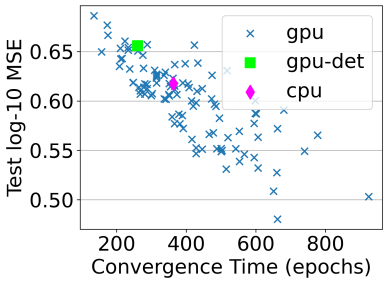
Looking at the single node perspective, the explicit neighborhood coupling becomes visible in Eq. (3). The update function with the central neighborhood-aggregation can be implemented in Pytorch-Geometric [2] to parallelize the computation per node and obtain a fast neighborhood aggregation by using the atomic scatter[1] operation with non-deterministic ordering of the neighbors in Eq. (3).

3 Experiments

Setup. Employing the graph property prediction from [4], this experiment compares GPU and CPU computing effects while training the discrete graph neural network in Eq. (3). The node-regression task is comprised of predicting the eccentricity of each node (the maximum length of all shortest paths rooted at the node). Via random seeds, the batches are guaranteed to be the same in each training process. Each model is trained 100 times, each with a total of 1500 epochs and a patience of 100 epochs for early-stopping when the validation mean-squared error (MSE) stops to decrease. Adam optimization with learning rate 0.003 and weight decay of 10^{-6} is used to repeatedly perform a model selection of the graph neural network Eq. (3) on the number of layers $L \in \{1, 5, 10, 15, 20\}$ with integration step size $\epsilon = 1.0$.

Results. Figure 1 shows the $\log_{10}$ MSE score on the test set obtained after 100 independent training runs for $L = 10$ on either pure GPU, GPU with deterministic operations enforced or pure CPU. Next, the repeated model selection is summarized by keeping track of the selected best-performing model for GPU-based training compared to the CPU-based ranking of evaluation set performance.

[1] https://pytorch-scatter.readthedocs.io/en/latest/functions/scatter.html



L	**Ranked #1**	**CPU Rank**
1	0	5
5	0	4
10	74	1
15	12	3
20	14	2

Fig. 1. Left) Test performance on Eccentricity for $L = 10$ alongside the early-stopped epoch for all 100 independent training runs. Right) Summary of the repeated model selection. It shows how often the GPU-trained models were selected as best performing on the evaluation set as compared to CPU-trained ranking.

4 Discussion

From the experiment conducted herein, it becomes evident that the exponential receptive field of message-passing neural networks leads to significant numerical differences in the downstream task behavior when computing with GPU. Deterministic scatter operations help to mitigate this problem at the cost of increased computation time. Chaotic behavior in terms of downstream task differences could not be observed when using finite precision CPU integration. Future work should study GPU-accelerated neighborhood computations that do not suffer from the exponential propagation of numerical errors but maintain the efficiency.

References

1. Chen, R.T.Q., Rubanova, Y., Bettencourt, J., Duvenaud, D.K.: Neural Ordinary Differential Equations. In: Bengio, S., Wallach, H., Larochelle, H., Grauman, K., Cesa-Bianchi, N., Garnett, R. (eds.) Advances in Neural Information Processing Systems. vol. 31. Curran Associates, Inc. (2018)
2. Fey, M., Lenssen, J.E.: Fast graph representation learning with PyTorch Geometric. In: ICLR Workshop on Representation Learning on Graphs and Manifolds (2019)
3. Gilmer, J., Schoenholz, S.S., Riley, P.F., Vinyals, O., Dahl, G.E.: Neural message passing for Quantum chemistry. In: Proceedings of the 34th International Conference on Machine Learning. ICML'17, vol. 70, p. 1263–1272. JMLR.org (2017)
4. Gravina, A., Bacciu, D., Gallicchio, C.: Anti-Symmetric DGN: a stable architecture for Deep Graph Networks. In: The Eleventh International Conference on Learning Representations (2023), https://openreview.net/forum?id=J3Y7cgZOOS
5. Haber, E., Ruthotto, L.: Stable architectures for deep neural networks. Inverse Problems **34**(1) (2017)
6. Hamilton, W.L., Ying, R., Leskovec, J.: Inductive representation learning on large graphs. In: Proceedings of the 31st International Conference on Neural Information Processing Systems. p. 1025–1035. NIPS'17, Curran Associates Inc. (2017)

7. Poli, M., Massaroli, S., Park, J., Yamashita, A., Asama, H., Park, J.: Graph neural ordinary differential equations. arXiv preprint arXiv:1911.07532 (2019)
8. Schuster, H.G., Just, W.: Deterministic chaos: an introduction. John Wiley & Sons (2006)

Coping with Drift in Hyperspectral Sensor Data

Valerie Vaquet, Barbara Hammer

Bielefeld University, Germany

Abstract

Hyperspectral imaging (HSI) constitutes an advanced sensor technology with applications in many fields ranging from earth observation to quality control. It can play a crucial role in investigating climate change, for instance, by remote sensing [7, 16, 14], and support coping with the associated consequences, for example as an assisting tool in agriculture [9, 1, 12]. Especially the application of HSI will further expand with the increasing availability of low-cost multispectral gadgets, which combine the advantage of a flexible application with being more affordable than complex hyperspectral cameras [1, 13].

Many tasks can be automated by coupling HSI sensors with machine leanring (ML) algorithms. Next to application in earth observation, e.g., estimating drought stress in forests [14, 2], and agriculture, e.g., detecting different kinds of sickness in plants in an early stage [9, 17], there are many more case studies of successful applications. Some examples are quality control in food production [4] and pharmaceutical applications [11], water resource management [6], medical diagnosis [10], artwork [8], and forensic document analysis [3]. However, when applying ML to sensor data, there is the risk of so-called sensor shifts harming the performance of the ML model. Minor changes in a sensor or across different instruments can induce differences in the data distributions [5, 15]. These can make models inaccurate or, in severe cases, render them useless. Retraining the models is expensive due to the associated labeling effort. Thus, this is not an option for low-cost HSI sensors since the cost for labeling and retraining would exceed the price for the sensor. In this work, will focus on this issue by investigating the following research questions:

- How can HSI sensor shifts be characterized?
- What are suitable methodologies to mitigate the effects of sensor shifts on machine learning models?

References

[1] T. Adão, J. Hruška, L. Pádua, J. Bessa, E. Peres, R. Morais, and J. Sousa. Hyperspectral Imaging: A Review on UAV-Based Sensors, Data Processing and Applications for Agriculture and Forestry. *Remote Sensing*, 9(11):1110, Oct. 2017.

[2] J. Behmann, J. Steinrücken, and L. Plümer. Detection of early plant stress responses in hyperspectral images. *ISPRS Journal of Photogrammetry and Remote Sensing*, 93:98–111, July 2014.

[3] G. Edelman, E. Gaston, T. Van Leeuwen, P. Cullen, and M. Aalders. Hyperspectral imaging for non-contact analysis of forensic traces. *Forensic Science International*, 223(1-3):28–39, Nov. 2012.

[4] G. Elmasry, M. Kamruzzaman, D.-W. Sun, and P. Allen. Principles and Applications of Hyperspectral Imaging in Quality Evaluation of Agro-Food Products: A Review. *Critical Reviews in Food Science and Nutrition*, 52(11):999–1023, Nov. 2012.

[5] R. N. Feudale, N. A. Woody, H. Tan, A. J. Myles, S. D. Brown, and J. Ferré. Transfer of multivariate calibration models: a review. *Chemometrics and Intelligent Laboratory Systems*, 64(2):181–192, Nov. 2002.

[6] M. Govender, K. Chetty, and H. Bulcock. A review of hyperspectral remote sensing and its application in vegetation and water resource studies. *Water SA*, 33(2), Dec. 2009.

[7] F. Hilton, R. Armante, T. August, C. Barnet, A. Bouchard, C. Camy-Peyret, V. Capelle, L. Clarisse, C. Clerbaux, P.-F. Coheur, A. Collard, C. Crevoisier, G. Dufour, D. Edwards, F. Faijan, N. Fourrié, A. Gambacorta, M. Goldberg, V. Guidard, D. Hurtmans, S. Illingworth, N. Jacquinet-Husson, T. Kerzenmacher, D. Klaes, L. Lavanant, G. Masiello, M. Matricardi, A. McNally, S. Newman, E. Pavelin, S. Payan, E. Péquignot, S. Peyridieu, T. Phulpin, J. Remedios, P. Schlüssel, C. Serio, L. Strow, C. Stubenrauch, J. Taylor, D. Tobin, W. Wolf, and D. Zhou. Hyperspectral Earth Observation from IASI: Five Years of Accomplishments. *Bulletin of the American Meteorological Society*, 93(3):347–370, Mar. 2012.

[8] M. Kubik. Chapter 5 Hyperspectral Imaging: A New Technique for the Non-Invasive Study of Artworks. In *Physical Techniques in the Study of Art, Archaeology and Cultural Heritage*, volume 2, pages 199–259. Elsevier, 2007.

[9] B. Lu, P. Dao, J. Liu, Y. He, and J. Shang. Recent Advances of Hyperspectral Imaging Technology and Applications in Agriculture. *Remote Sensing*, 12(16):2659, Aug. 2020.

[10] G. Lu and B. Fei. Medical hyperspectral imaging: a review. *Journal of Biomedical Optics*, 19(1):010901, Jan. 2014.

[11] Y. Roggo, A. Edmond, P. Chalus, and M. Ulmschneider. Infrared hyperspectral imaging for qualitative analysis of pharmaceutical solid forms. *Analytica Chimica Acta*, 535(1-2):79–87, Apr. 2005.

[12] R. N. Sahoo, S. S. Ray, and K. R. Manjunath. Hyperspectral remote sensing of agriculture. *Current Science*, 108(5):848–859, 2015.

[13] Stuart, McGonigle, and Willmott. Hyperspectral Imaging in Environmental Monitoring: A Review of Recent Developments and Technological Advances in Compact Field Deployable Systems. *Sensors*, 19(14):3071, July 2019.

[14] J. Transon, R. d'Andrimont, A. Maugnard, and P. Defourny. Survey of Hyperspectral Earth Observation Applications from Space in the Sentinel-2 Context. *Remote Sensing*, 10(3):157, Jan. 2018.

[15] D. Tuia, C. Persello, and L. Bruzzone. Domain Adaptation for the Classification of Remote Sensing Data: An Overview of Recent Advances. *IEEE Geoscience and Remote Sensing Magazine*, 4(2):41–57, June 2016.

[16] A. Vali, S. Comai, and M. Matteucci. Deep Learning for Land Use and Land Cover Classification Based on Hyperspectral and Multispectral Earth Observation Data: A Review. *Remote Sensing*, 12(15):2495, Aug. 2020.

[17] J.-C. Zhang, R.-l. Pu, J.-h. Wang, W.-j. Huang, L. Yuan, and J.-h. Luo. Detecting powdery mildew of winter wheat using leaf level hyperspectral measurements. *Computers and Electronics in Agriculture*, 85:13–23, July 2012.

A Measure Theoretic Approach to Concept Drift in Infinite Data Streams

Fabian Hinder and Barbara Hammer

Bielefeld University, Germany

Abstract

In the context of batch learning, it is commonly assumed that a finite amount of data is generated from the same distribution during training, testing, and application, which remains unchanged over time. However, this assumption is not realistic. Most data sources spontaneously generate new data over time, resulting in an ultimately infinite amount of data – a setup considered in stream learning. In addition, the distribution underlying the data is subject to change due to constant changes in the environment – a phenomenon known as concept drift. Currently, there is no general theory on how to extend the concept of probability distributions, which is crucial in all of probability theory, statistics, and machine learning theory, to this setup. Approaches such as considering the data as a stochastic process [3, 1], sample-based approaches [2], or distribution processes [4, 5] are relevant, but lack certain aspects. In this work, we suggest a generalization of the concept of probability distributions in a way that allows us to capture the properties of data streams with concept drift. We start by considering distributions on time windows, which is an important concept in many stream learning algorithms. We then show that, under mild assumptions on the relationship between window distributions, all data in a stream can be described by a unique, potentially infinite measure on data and time. This construction is analogous to the notion of holistic distributions considered in [4, 5] and agrees with it in the case of a finite time horizons and data from standard Borel spaces.

References

[1] A. Agarwal and J. C. Duchi. The generalization ability of online algorithms for dependent data. *IEEE Transactions on Information Theory*, 59(1):573–587, 2012.

[2] J. Gama, I. Žliobaitė, A. Bifet, M. Pechenizkiy, and A. Bouchachia. A survey on concept drift adaptation. *ACM computing surveys (CSUR)*, 46(4):1–37, 2014.

[3] S. Hanneke and L. Yang. Statistical learning under nonstationary mixing processes. In *The 22nd International Conference on Artificial Intelligence and Statistics*, pages 1678–1686. PMLR, 2019.

[4] F. Hinder, A. Artelt, and B. Hammer. Towards non-parametric drift detection via dynamic adapting window independence drift detection (dawidd). In *International Conference on Machine Learning*, pages 4249–4259. PMLR, 2020.

[5] F. Hinder, V. Vaquet, and B. Hammer. One or two things we know about concept drift – a survey on monitoring evolving environments, 2023.

Online Learning Dynamics in Layered Neural Networks with Arbitrary Activation Functions

Otavio Citton, Frederieke Richert, Michael Biehl

Institute for Mathematics, Computer Science and Artificial Intelligence, University of Groningen, The Netherlands

Abstract

Activation functions are a fundamental piece of a neural network architecture. Its non-linearity is what allow the network to have the ability of approximate more complicated rules. Understanding the role these functions play in the learning dynamics is of fundamental importance and might permit us to design more efficient and more interpretable models, two problems that the field of machine learning having been facing since the introduction of deep learning and the large amount of data necessary to train the models.

We revisit and extend the statistical physics based analysis of layered neural networks trained by online gradient descent [1, 2]. We focus on the influence of the hidden unit activation functions on the typical learning behavior in model scenarios. Expanding activation functions in terms of Hermite polynomials enables us to extend the formalism to the analysis of soft committee machines with arbitrary activation in student-teacher scenarios. This approach requires much lower computational effort than naive numerical integration, which is practically infeasible. Moreover, it now becomes possible to treat mismatched scenarios in which the student activation function differs from the one used in the target rule definition. This makes it possible to study realistic models of machine learning.

This presentation is based in our results from [3] that was accepted for publication in the ESANN 2024 proceedings.

References

[1] D. Saad and S. A. Solla. On-line learning in soft committee machines. Phys. Rev. E, 52:4225-4243, Oct 1995.

[2] M. Straat and M. Biehl. On-line learning dynamics of ReLU neural networks using statistical physics techniques. In M. Verleysen, editor, Proc. European Symposium on Artificial Neural Networks (ESANN), pages 517-522, 2019.

[3] O. Citton, F. Richert, and M. Biehl. On-line Learning Dynamics in Layered Neural Networks with Arbitrary Activation Functions. In M. Verleysen, editor, Proc. European Symposium on Artificial Neural Networks (ESANN), (To be published in October, 2024).

GMLVQ for fMRI Analysis in the Context of Movement Disorders

Mariya Shumska[1,2,3], Jelle Dalenberg[1,2], Remco Renken[4], Marina de Koning-Tijssen[1,2], and Michael Biehl[3]

[1]Expertise Center Movement Disorders Groningen, University Medical Center Groningen, The Netherlands
[2]Department of Neurology, University Medical Center Groningen, University of Groningen, the Netherlands
[3]Bernoulli Institute for Mathematics, Computer Science and Artificial Intelligence, University of Groningen, The Netherlands
[4]Department of Neuroscience, University Medical Center Groningen, University of Groningen, the Netherlands

Abstract

Hyperkinetic movement disorders (HMDs) are characterized by excessive, involuntary movements that significantly affect patients' quality of life. Cortical myoclonus, a specific type of HMD, manifests as brief, involuntary muscle jerks, and is caused by abnormal activity in the central nervous system [1, 3]. However, its underlying mechanisms remain not fully clear. Analyzing fMRI data, which provides critical insights into brain activity, can be instrumental in filling the gaps in our understanding of myoclonus.

In this project, we use a prototype-based algorithm called Generalized Matrix Learning Vector Quantization (GMLVQ) [2]. It is robust in handling high-dimensional input, making it well-suited for complex datasets like those derived from fMRI, and it offers interpretability of results, which is crucial for medical applications.

Our study demonstrates that GMLVQ can serve as an effective feature selection tool and a classifier for small datasets, providing both explainability and valuable insights into hyperkinetic movement disorders such as cortical myoclonus.

References

[1] C. D. Marsden, M. Hallett, and Stanley Fahn. 13 - The nosology and pathophysiology of myoclonus. In C. David Marsden and Stanley Fahn, editors, *Movement Disorders*, pages 196–248. Butterworth-Heinemann, January 1981.

[2] Petra Schneider, Michael Biehl, and Barbara Hammer. Adaptive relevance matrices in learning vector quantization. *Neural Computation*, 21(12):3532–3561, December 2009.

[3] Hiroshi Shibasaki and Mark Hallett. Electrophysiological studies of myoclonus. *Muscle & Nerve*, 31(2):157–174, February 2005.

Towards Explainable Rejects for prototype-Based Classifiers?

Johannes Brinkrolf, Valerie Vaquet, Fabian Hinder, and Barbara Hammer

Bielefeld University, Faculty of Technology, Germany

Abstract

Prototype-based methods constitute a robust and transparent family of machine-learning models. To increase robustness in real-world applications, they are frequently coupled with reject options. While the state-of-the-art method, *relative similarity* [2], couples the rejection of samples with high aleatoric and epistemic uncertainty, the technique lacks transparency, i.e., an explanation of why a sample has been rejected.

In this work, we derive an explanation scheme for reject options in prototype-based classification from the analytical study of relative similarity in [1]. Further, we demonstrate the potential of this explanation scheme in a classification scenario with more than two classes.

References

[1] Johannes Brinkrolf. "Learning Vector Quantization for the Real-World: Privacy, Robustness, and Sparsity". PhD thesis. 2023, p. 139. URL: `https://pub.uni-bielefeld.de/record/2985339`.

[2] Lydia Fischer, Barbara Hammer, and Heiko Wersing. "Optimal local rejection for classifiers". In: *Neurocomputing* 214 (2016), pp. 445–457.

Runtime-processing for microgravity investigations directly attached to the experiment

Jan Auth, Mohammad Babar Jafree, Wael Aly Mansour, Florian Zaussinger

University of Applied Sciences Mittweida, Germany

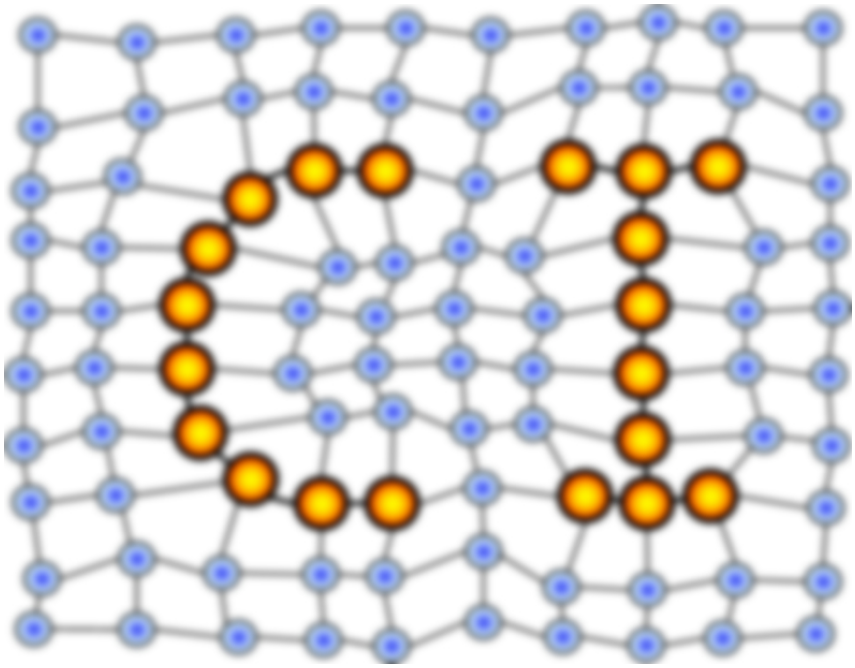
Abstract

Physical experiments in microgravity, aiming for setup on the ISS, getting more and more remote as they evolve. And so questions regarding data processing arise: How to use results from a previous experiment in the following one, e.g. 48h later? How to make reasoning for the next experimental setup? We give aid to formulate such questions and answer them by providing dedicated pipelines and visualizations.

We present how we applied such a pipeline on the LAPLACE Experiment for the ISS and present the status on the adaption to the Compact experiment heading towards a parabolic flight campaign in September 2024.

MACHINE LEARNING REPORTS

Report 01/2024



Impressum

Machine Learning Reports ISSN: 1865-3960

▽ Publisher/Editors

Prof. Dr. rer. nat. Thomas Villmann
University of Applied Sciences Mittweida
Technikumplatz 17, 09648 Mittweida, Germany
• http://www.mni.hs-mittweida.de/

Prof. Dr. rer. nat. Frank-Michael Schleif
Technical University of Applied Sciences Wuerzburg-Schweinfurt
Sanderheinrichsleitenweg 20, 97074 Wuerzburg, Germany
• https://fiw.thws.de/

▽ Acknowledgments

We would like to thank the reviewers for their time and patience.