

Seminar im Wintersemester 2026/2027

22.07.2026

Mathematische Aspekte des Maschinellen Lernens

In vielen Bereichen der Wissenschaft und Industrie wurden durch Anwendung von Methoden des Maschinellen Lernens in jüngster Vergangenheit Durchbrüche erreicht. So sind beispielsweise große Fortschritte im Bereich des autonomen Fahrens erst durch den Einsatz dieser Methoden möglich geworden.

In dem Seminar möchten wir uns mit den mathematischen Aspekten von ausgewählten Themen des Maschinellen Lernens beschäftigen. Dabei geht es einerseits um die Konstruktion der Verfahren, aber auch Konvergenz- sowie Implementierungsaspekte. Im Fokus des Seminars liegen Regressionsverfahren, neuronale Netze, Trainingsmethoden und die Anwendung auf Differentialgleichungen.

Das Seminar richtet sich an Studierende im Masterstudiengang. Es werden Kenntnisse in numerischer Mathematik und Wahrscheinlichkeitstheorie vorausgesetzt, wie sie im Bachelorstudiengang Mathematik vermittelt werden.

Vorbesprechung: 22.07.2026, 13:10 Uhr im SR 3.061 (Geb. 20.30).

Bei Interesse bitten wir um vorherige Anmeldung per Mail bei Herrn Schneiderhan unter david.schneiderhan@kit.edu.

Eine Teilnahme an der Vorbesprechung ohne vorherige Anmeldung ist ebenfalls möglich.

Das Seminar findet mittwochs um 9:45 Uhr im SR 2.066 (Geb. 20.30) statt.

Themenvorschläge für Seminarvorträge

1. Einführung

- (a) Einführung in das maschinelle Lernen (überwachtes, unüberwachtes Lernen, Klassifikation und Regression) [Bac, 2. Kap.] [Her23, 1. Kap.]
- (b) VC-Dimension [Gro22]
- (c) SVM, Soft-Margin SVM, Perzeptron, lineare Separabilität, XOR-Problem, Übergang SVM zu NN [MRT18, Kap. 5][MRT18, Kap. 6][Jo21, Kap. 8]

2. Neuronale Netze

- (a) Approximationstheorie von Neuronalen Netzen [EPGB19]

3. Trainingsmethoden

- (a) Automatisches Differenzieren [A⁺18, Kap. 3], [Agg20, Kap. 11] evtl. [BPRS18], [GW08]
- (b) Abstiegsverfahren: Momentum, SGD, RMSProp, ... [A⁺18, Kap. 3]
- (c) Abstiegsverfahren: ADAM [KB14], [RKK19]
- (d) Abstiegsverfahren: Methoden Vergleich: Recent Advancements [LQJ⁺25], [BP24], [WLS⁺19] evtl. [ARP24]

4. Neural Tangent Kernel + Deep Gaussian Processes
 - (a) (Kernel-)Regression [Rad22]
 - (b) NN als Gauß-Prozess [Rad22]
 - (c) Deep Gaussian Processes [DGST18], [DL13]
 - (d) Neural Tangent Kernel und 1-Layer NN mit unendlicher Layergröße [Rad22]
5. Monotone und Min-Max neuronale Netze [Ige23]
6. Ordinary differential equations
 - (a) Neural Ordinary Differential Equation (NODE) [CRBD18]
 - (b) ODE: Framework to Solve Ordinary Differential Equations using Artificial Neural Networks [LW20]
7. Partial differential equations
 - (a) Operator Learning [LKA⁺21]
 - (b) Derivative-informed neural operator [ORCVG24], [ORVCG22]
 - (c) Physics informed neural networks [KKL⁺21], [BSK25]

Literatur

- [A⁺18] Charu C Aggarwal et al. Neural networks and deep learning. *Springer*, 10:978–3, 2018.
- [Agg20] Charu C Aggarwal. *Linear Algebra and Optimization for Machine Learning: A Textbook*. Springer, 2020.
- [ARP24] Beatriz Flamia Azevedo, Ana Maria A. C. Rocha, and Ana I. Pereira. Hybrid approaches to optimization and machine learning methods: A systematic literature review. *Machine Learning*, 113:4055–4097, 2024.
- [Bac] Francis Bach. *Learning Theory from First Principles*.
- [BP24] Kewei Bian and Rahul Priyadarshi. Machine learning optimization techniques: A survey, classification, challenges, and future research issues. *Archives of Computational Methods in Engineering*, 31(7):4209–4233, 2024.
- [BPRS18] Atilim Gunes Baydin, Barak A Pearlmutter, Alexey Andreyevich Radul, and Jeffrey Mark Siskind. Automatic differentiation in machine learning: a survey. *Journal of machine learning research*, 18(153), 2018.
- [BSK25] Jan-Hendrik Bastek, WaiChing Sun, and Dennis Kochmann. Physics-informed diffusion models. In Y. Yue, A. Garg, N. Peng, F. Sha, and R. Yu, editors, *International Conference on Learning Representations*, volume 2025, pages 3360–3385, 2025.
- [CRBD18] Ricky TQ Chen, Yulia Rubanova, Jesse Bettencourt, and David K Duvenaud. Neural ordinary differential equations. *Advances in neural information processing systems*, 31, 2018.
- [DGST18] Matthew M. Dunlop, Mark A. Girolami, Andrew M. Stuart, and Aretha L. Teckentrup. How deep are deep gaussian processes? *Journal of Machine Learning Research*, 19(54):1–46, 2018.
- [DL13] Andreas Damianou and Neil D. Lawrence. Deep gaussian processes. In Carlos M. Carvalho and Pradeep Ravikumar, editors, *Proceedings of the Sixteenth International Conference on Artificial Intelligence and Statistics (AISTATS 2013)*, volume 31 of *Proceedings of Machine Learning Research*, pages 207–215. PMLR, 2013.
- [EPGB19] Dennis Elbrächter, Dmytro Perekrestenko, Philipp Grohs, and Helmut Bölcskei. Deep neural network approximation theory. *arXiv preprint arXiv:1901.02220*, 2019.
- [Gro22] *Mathematical Aspects of Deep Learning*. Cambridge University Press, 2022.
- [GW08] Andreas Griewank and Andrea Walther. *Evaluating derivatives: principles and techniques of algorithmic differentiation*, volume 105. Siam, 2008.
- [Her23] Evelyn Herberg. *Lecture Notes: neural network architectures*. 2023.

- [Ige23] Christian Igel. Smooth monotonic networks. *arXiv preprint arXiv:2306.01147*, 2023.
- [Jo21] Taeho Jo. *Machine Learning Foundations: Supervised, Unsupervised, and Advanced Learning*. Springer Nature, 2021.
- [KB14] Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.
- [KKL⁺21] George Em Karniadakis, Ioannis G. Kevrekidis, Lu Lu, Paris Perdikaris, Sifan Wang, and Liu Yang. Physics-informed machine learning. *Nature Reviews Physics*, 3:422–440, 2021.
- [LKA⁺21] Zongyi Li, Nikola Kovachki, Kamyar Azizzadenesheli, Burigede Liu, Kaushik Bhattacharya, Andrew Stuart, and Anima Anandkumar. Fourier neural operator for parametric partial differential equations. In *International Conference on Learning Representations (ICLR)*, 2021.
- [LQJ⁺25] Xiaodong Liu, Huaizhou Qi, Suisui Jia, Yongjing Guo, and Yang Liu. Recent advances in optimization methods for machine learning: A systematic review. *Mathematics*, 13(13):2210, 2025.
- [LW20] Liam LH Lau and Denis Werth. Oden: A framework to solve ordinary differential equations using artificial neural networks. *arXiv preprint arXiv:2005.14090*, 2020.
- [MRT18] Mehryar Mohri, Afshin Rostamizadeh, and Ameet Talwalkar. *Foundations of machine learning*. MIT press, 2018.
- [ORCVG24] Thomas O'Leary-Roseberry, Peng Chen, Umberto Villa, and Omar Ghattas. Derivative-informed neural operator: An efficient framework for high-dimensional parametric derivative learning. *Journal of Computational Physics*, 496:112555, 2024.
- [ORVCG22] Thomas O'Leary-Roseberry, Umberto Villa, Peng Chen, and Omar Ghattas. Derivative-informed projected neural networks for high-dimensional parametric maps governed by pdes. *Computer Methods in Applied Mechanics and Engineering*, 388:114199, 2022.
- [Rad22] Adityanarayanan Radhakrishnan. Lecture notes, modern machine learning: Simple methods that work, 2022.
- [RKK19] Sashank J Reddi, Satyen Kale, and Sanjiv Kumar. On the convergence of adam and beyond. *arXiv preprint arXiv:1904.09237*, 2019.
- [WLS⁺19] Dorina Weichert, Patrick Link, Anke Stoll, Stefan Rüping, Steffen Ihlenfeldt, and Stefan Wrobel. A review of machine learning for the optimization of production processes. *The International Journal of Advanced Manufacturing Technology*, 104(5–8):1889–1902, 2019.