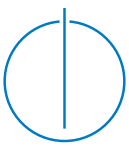
DEPARTMENT OF INFORMATICS

TECHNISCHE UNIVERSITÄT MÜNCHEN

Master's Thesis in Informatics

Neural Network Hyperparameter Optimization with Sparse Grids

Maximilian Michallik



DEPARTMENT OF INFORMATICS

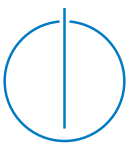
TECHNISCHE UNIVERSITÄT MÜNCHEN

Master's Thesis in Informatics

Neural Network Hyperparameter Optimization with Sparse Grids

Parameteroptimierung von neuronalen Netzen mit dünnen Gittern

Author:	Maximilian Michallik
Supervisor:	Dr. Felix Dietrich
Advisor:	Dr. Michael Obersteiner, Dr. Felix Dietrich
Submission Date:	



I confirm that this master's thesis in informatics is my own work and I have documented all sources and material used.

Munich, Maximilian Michallik

Acknowledgments

Abstract

Zusammenfassung

Contents

1 Introduction

2 Theoretical Background

2.1 Sparse Grids

2.1.1 Numerical Approximation of Functions

2.1.2 Adaptive Sparse Grids

2.2 Hyperparameter Optimization

2.2.1 Random Search

2.2.2 Bayesian Optimization

2.2.3 Grid Search

2.3 Adaptive Grid Search with Sparse Grids

3 Implementation

4 Results

5 Conclusion and Outlook

List of Figures

List of Tables

Bibliography

[1] H.-J. Bungartz and M. Griebel, "Sparse grids," *Acta Numerica*, vol. 13, pp. 147–269, 2004. DOI: 10.1017/S0962492904000182.

[2] D. Pflüger, "Spatially adaptive sparse grids for high dimensional problems," vol. 13, 2010.

[3] J. Garcke, "Sparse grids in a nutshell," in *Sparse Grids and Applications*, J. Garcke and M. Griebel, Eds., Berlin, Heidelberg: Springer Berlin Heidelberg, 2013, pp. 57–80, ISBN: 978-3-642-31703-3.

[4] J. Garcke, M. Griebel, and M. Thess, "Data mining with sparse grids," *Computing*, vol. 67, 2001. DOI: 10.1007/s006070170007.

[5] M. Griebel, M. Schneider, and C. Zenger, "A combination technique for the solution of sparse grid problems," 1990.

[6] M. Obersteiner and H.-J. Bungartz, "A generalized spatially adaptive sparse grid combination technique with dimension-wise refinement," *SIAM Journal on Scientific Computing*, vol. 43, no. 4, A2381–A2403, 2021. DOI: 10.1137/20M1325885. eprint: https://doi.org/10.1137/20M1325885.

[7] M. Hegland, J. Garcke, and V. Challis, "The combination technique and some generalisations," *Linear Algebra and its Applications*, vol. 420, no. 2, pp. 249–275, 2007, ISSN: 0024-3795. DOI: https://doi.org/10.1016/j.laa.2006.07.014.

[8] J. Garcke, "Regression with the optimised combination technique," in *Proceedings of the 23rd International Conference on Machine Learning*, ser. ICML '06, Pittsburgh, Pennsylvania, USA: Association for Computing Machinery, 2006, pp. 321–328, ISBN: 1595933832. DOI: 10.1145/1143844.1143885.

[9] M. Fabry, "Spatially adaptive density estimation with the sparse grid combination technique," Masterarbeit, Technical University of Munich, Sep. 2020.

[10] C. C. Moser, "Machine learning with the sparse grid density estimation using the combination technique," Bachelorarbeit, Technical University of Munich, Sep. 2020.

[11] C. R. Harris, K. J. Millman, S. J. van der Walt, R. Gommers, P. Virtanen, D. Cournapeau, E. Wieser, J. Taylor, S. Berg, N. J. Smith, R. Kern, M. Picus, S. Hoyer, M. H. van Kerkwijk, M. Brett, A. Haldane, J. F. del Río, M. Wiebe, P. Peterson, P. Gérard-Marchant, K. Sheppard, T. Reddy, W. Weckesser, H. Abbasi, C. Gohlke, and T. E. Oliphant, "Array programming with NumPy," *Nature*, vol. 585, no. 7825, pp. 357–362, Sep. 2020. DOI: 10.1038/s41586-020-2649-2.

[12] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, *et al.*, "Scikit-learn: Machine learning in python," *the Journal of machine Learning research*, vol. 12, pp. 2825–2830, 2011.

[13] B. Peherstorfer, D. Pflüge, and H.-J. Bungartz, "Density estimation with adaptive sparse grids for large data sets," in *Proceedings of the 2014 SIAM International Conference on Data Mining (SDM)*, pp. 443–451. DOI: 10.1137/1.9781611973440.51. eprint: https://epubs.siam.org/doi/pdf/10.1137/1.9781611973440.51.