

Lavinia Schmid

Data Scientist - Machine Learning

✉ lavinia.schmid@example.de • 📞 +49 30 8901 2345 • 📍 Berlin, Deutschland
🌐 laviniaschmid.dev • 🌐 linkedin.com/in/laviniaschmid • 📄 github.com/laviniaschmid

Profil

Data Scientist mit Schwerpunkt Deep Learning bei AUTO1 Group und Otto Group. 6 Jahre Erfahrung in Ranking-Transformern, Modell-Distillation und MLOps. Training auf 48-GPU-Cluster (Ray Train) mit 1,9x Throughput vs DDP-Baseline. Migration von 7 Tree-Modellen auf destillierte neuronale Äquivalente.

Berufserfahrung

04/2022 - heute

Hamburg, Deutschland

Data Scientist - Machine Learning

Otto Group

Ranking-ML im Search-Personalization-Team

- Training eines Ranking-Transformers auf 48 GPUs (Ray Train) mit 1,9x Throughput gegenüber bisherigem DDP-Baseline
- Migration von 7 Tree-Modellen auf destillierte neuronale Äquivalente, p99-Online-Latenz von 38 ms auf 14 ms gesenkt
- Co-Eigner der Kubeflow-Pipelines und MLflow-Registry mit 34 Produktions-Endpoints
- Implementierung eines NDCG-fokussierten Evaluation-Frameworks für Search-Ranking
- Mentoring von 3 Junior-Engineers im PyTorch- und Hugging-Face-Stack

09/2019 - 03/2022

Berlin, Deutschland

Junior Data Scientist

AUTO1 Group

Pricing- und Demand-ML im Wholesale-Bereich

- Aufbau eines XGBoost-Pricing-Modells (R^2 0,86) auf 1,8 Mio. Fahrzeug-Transaktionen, 4,2 Mio. EUR ROI
- Migration der Training-Pipeline auf Ray + SageMaker, Trainingszeit von 14h auf 1,6h reduziert
- Co-Autor eines KDD 2022 Workshop Papers zu Restwert-Schätzung mit Quantilregression

Ausbildung

10/2017 - 09/2019

Karlsruhe, Deutschland

M.Sc. Informatik

Karlsruher Institut für Technologie (KIT)

Machine Learning • 1,2

10/2014 - 09/2017

Aachen, Deutschland

B.Sc. Informatik

RWTH Aachen

Informatik • GPA: 1,4

Fähigkeiten

- Python (PyTorch, JAX)
- Hugging Face & Transformers
- Ray Train & Ray Serve
- MLflow 2.x & Kubeflow
- ONNX & TensorRT
- SageMaker & Vertex AI
- Snowflake & Databricks
- Distillation & Quantization

Projekte

- 05/2024 - 09/2024 **Transformer-Ranking-Migration**
- Migration von Random Forest zu Transformer-Ranker, NDCG@10 +6,8%, p99-Latenz 14 ms
- 02/2023 - 06/2023 **Distillation-Toolkit (Open Source)**
- Open-Source-PyTorch-Library zur Knowledge-Distillation, 920 GitHub-Stars

Zertifikate

- 08/2025 **NVIDIA DLI - Fundamentals of Deep Learning**
- 01/2025 **Databricks Certified Machine Learning Professional**
- 04/2023 **TensorFlow Developer Certificate**

Sprachen

Deutsch	Englisch	Italienisch
Muttersprache	C1	B2

Stärken

GPU-Throughput

Optimiere Trainingspipelines mit Ray Train, Mixed Precision und FSDP für 1,5-2x Speedup

Latenz-Disziplin

Distilliere und quantisiere Modelle bewusst auf p99-Latenz-Budgets unter 25 ms

Wissens-Sharing

Halte hausinternes PyTorch-Bootcamp für 60+ Engineers in 4 Cohorts