Johannes Roth

Data Scientist / Machine Learning Engineer

✉ johannes@roth24.de ☎ +49 176 340 123 80 Kurt-Eisner-Str. 66, 04275 Leipzig, Germany

in LinkedIn ○ GitHub 🌐 jroth.space

Summary

Detail-oriented Data Scientist and PhD Candidate with extensive experience in developing, implementing, and evaluating machine learning models, particularly in computer vision and medical imaging domains. Proven ability to manage the full project lifecycle, from data acquisition and preprocessing to model training, optimization (e.g., Bayesian Optimization), and deployment (Python backends, microservices). Seeking to leverage expertise in Python, TensorFlow/PyTorch, data analysis, and statistical modeling to solve challenging problems in an industry setting.

Work Experience

PhD Candidate (Computational Neuroscience)

Max Planck Institute for Human Cognitive and Brain Sciences (Leipzig) University of Gießen *May 2022 — Present*

- Developed computational models and applied machine learning techniques (encoding/decoding models, RSA) to analyze fMRI data and study the human visual system.
- Designed and analyzed fMRI experiments, implementing analysis pipelines using Python (Nilearn, Scikit-learn) and MATLAB.
- Investigated methods to enhance data acquisition efficiency using high-frequency stimulus presentation and adaptive selection strategies.

Research Assistant (ML in Medicine)

ScaDS.AI Dresden/Leipzig *Jun 2021 — May 2022*

- Developed and improved deep learning models (Python, PyTorch/TensorFlow) for medical applications.
- Enhanced uncertainty-aware prostate cancer mortality prediction models, achieving state-of-the-art results.
- Built and evaluated models for automated brain tumor segmentation using UNet-based architectures on MRI data.

Data Scientist (Working Student)

CHECK24 Vergleichsportal *Oct 2019 — May 2021*

- Led the development and deployment of an image processing micro-service using Python (Flask/Gunicorn, Redis).
- Trained and deployed ML models (TensorFlow, PyTorch) for image deduplication, retrieval, classification, upscaling, and quality assessment.
- Improved search result relevance by applying Bayesian optimization to recommender systems and developing ranking models.
- Conducted statistical analyses (e.g., outlier detection, time series visualization) and presented results to stakeholders.

Full-stack Developer (Freelance)

Kimetric UG   *Oct 2020 − May 2021*

- Developed two websites for academic clients, managing the full development lifecycle from specification to deployment.
- Configured Linux hosting environments with Nginx and Gunicorn.
- Built backends using Django for site and user management; created frontends using HTML, CSS, and JavaScript.

Data Scientist (Working Student)

Webdata Solutions GmbH (now part of Vistex)   *Oct 2018 − Oct 2019*

- Led development of an image retrieval and deduplication pipeline, significantly outperforming previous methods.
- Crawled, cleaned, and processed large-scale image datasets for model training.
- Trained models (TensorFlow, PyTorch) for image representation learning, segmentation, and text classification.

Data Analyst (Working Student)

Mercateo Services GmbH (now Unite)   *Feb 2018 − Aug 2018*

- Assisted in developing cloud-based data warehousing solutions using AWS and Apache NiFi.
- Developed ETL processes and built integration test infrastructure using Python.
- Contributed to reporting systems using Pentaho, Java, and R.

Education

M.Sc. Computer Science

Leipzig University   *2017 − 2021*
Grade: 1.2 (Sehr Gut / Very Good); Focus: Data Science, Machine Learning, Medical Image Processing.

B.Sc. Business Information Systems

Leipzig University   *2014 − 2017*
Grade: 1.5 (Sehr Gut / Very Good)

Skills

Programming:	Python, SQL, Git, Bash, MATLAB, Java, R, JavaScript, HTML/CSS
ML/Data Science:	Scikit-learn, Pandas, NumPy, SciPy, TensorFlow, PyTorch, Keras, Statsmodels, Classification, Regression, Clustering, Dimensionality Reduction, Deep Learning (CNNs, GANs, Transformers), Computer Vision, Statistical Analysis, Data Visualization, Bayesian Optimization, Recommender Systems, Ranking Models
Tools & Platforms:	Docker, Linux, Nginx, Gunicorn, Redis, Postgres, SQLite, Flask, Django, High-Performance Computing (SLURM), AWS (basic), Apache NiFi
Domain Specific:	Medical Image Analysis, fMRI Data Handling
Languages:	German (Native), English (Fluent)

Publications & Presentations

2025   M. Badwal, J. Bergmann, **J. Roth**, C. Doeller, M. Hebart. *The scope and limits of fine-grained image and category information in the ventral visual pathway*, Journal of Neuroscience, 45(3) (In Press).

2023   **J. Roth**, Y. Miyawaki, M. Hebart. *Assessing high stimulus presentation rates for fMRI studies*, Journal of Vision (VSS Abstract), 23(9), 5070–5070.

2021   **J. Roth**, J. Keller, S. Franke, T. Neumuth, D. Schneider. *Multi-plane UNet++ ensemble for glioblastoma segmentation*, MICCAI BrainLesion Workshop (BrainLes), LNCS 12929, pp. 285–294 (Conference Paper).

2021   **J. Roth**, K. Seeliger, T. Schmid, M. Hebart. *Synthesizing preferred stimuli for voxels in the visual system*, Poster at Computational and Systems Neuroscience (Cosyne) 2021.

Volunteer Work

Internal PhD Representative

MPI CBS   *2023 − 2024*

Represented PhD student interests in institutional policy discussions and administrative matters.