Frederik Pahde

Curriculum Vitae

Samariterstr. 6
10247 Berlin, Germany
(+49) 15784471604
frederikpahde@gmail.com
linkedin.com/in/frederik-pahde-176277a7

Education

2023 - today **Technical University Berlin**, *PhD Student*, Supervisor: Prof. Wojciech Samek.

2017 **University of Oslo**, *Semester abroad*, Master level courses.

2015 – 2018 **Humboldt-Universität zu Berlin**, *Master of Science in Information Systems*, Emphasis on Machine Learning, *Final grade - 1.2 (very good), A (= best 10%).*

2012 – 2015 **DHBW Mannheim (Cooperative State University)**, *Bachelor of Science in Information Systems*, Emphasis on Software Engineering, *Final grade - 1.4 (very good), A (= best 10%).*

Work Experience

Aug 2021 – today **Researcher / PhD Student** , *Fraunhofer Heinrich-Hertz Institute*, Berlin (Germany).
Currently I work in the eXplainable AI group under the supervision of Dr. Sebastian Lapuschkin and Prof. Dr. Wojciech Samek. My research focuses on the development of XAI-based methods for model improvement and data refinement with a focus on medical applications.

Jun 2019 – Jul 2021 **Machine Learning Scientist**, *Amazon.com, Inc.*, Edinburgh (United Kingdom).
I worked as a full-time Machine Learning Scientist at Amazon, focusing on explainability and fairness in ML in the recruiting technology domain.

Oct 2017 – Feb 2019 **Master's Thesis Student / Research Associate**, *SAP AI Research*, Berlin (Germany).
In my master's thesis I have focused on few-shot learning. I have continued my research in that field in several projects as a full-time research associate (6-months contract).

Feb 2017 – Jun 2017 **Machine Learning Developer (Intern)**, *Bakken & Bæck*, Oslo (Norway).
I performed a variety of natural language processing tasks including text embedding learning and language detection. My responsibilities included data retrieval, data preparation and prototype development.

Oct 2015 – Jan 2017 **Working Student**, *SAP SE - Medical Research Insights*, Potsdam (Germany).
Supporting the SAP Medical Research Insights team at the SAP Innovation center, I have worked on several projects including infrastructural topics, product development and machine learning prototypes.

Oct 2012 – Sep 2015 **Bachelor Student (Cooperative Education)**, *SAP SE*, Walldorf (Germany).
During my bachelor's program I was employed by SAP SE and have worked in several development teams, including a 3-months internship in a mobile application development team in Montreal (Canada).

Selected Publications

2024 **Navigating Neural Space: Revisiting Concept Activation Vectors to Overcome Directional Divergence**, *Frederik Pahde, Maximilian Dreyer, Leander Weber, Moritz Weckbecker, Christopher Anders, Thomas Wiegand, Wojciech Samek, Sebastian Lapuschkin* (preprint).

2023 **From Hope to Safety: Unlearning Biases of Deep Models via Gradient Penalization in Latent Space**, *Maximilian Dreyer*, Frederik Pahde*, Christopher Anders, Wojciech Samek, Sebastian Lapuschkin*, AAAI 2024 (*oral*).

2023 **Reveal to Revise: An Explainable AI Life Cycle for Iterative Bias Correction of Deep Models**, *Frederik Pahde*, Maximilian Dreyer*, Wojciech Samek, Sebastian Lapuschkin*, MICCAI 2023.

2023 **Optimizing Explanations by Network Canonization and Hyperparameter Search**, *Frederik Pahde, Galip Yolcu, Alexander Binder, Wojciech Samek, Sebastian Lapuschkin*, CVPRW 2023.

2021 **Multimodal Prototypical Networks for Few-shot Learning**, *Frederik Pahde, Mihai Puscas, Tassilo Klein, Moin Nabi*, WACV 2021.

2019 **Self-paced Adversarial Training for Multimodal Few-shot Learning**, *Frederik Pahde, Oleksiy Ostapenko, Patrick Jähnichen, Tassilo Klein, Moin Nabi*, WACV 2019.

Research Talks

2023 **Reveal to Revise: An Explainable AI Life Cycle for Iterative Bias Correction of Deep Models**, Invited Talk at "Fairness of AI in Medical Imaging" Workshop.

2018 **Cross-Model Deep Generative Models for Discriminative Tasks: A Sample-Efficient Perspective**, Talk (45 min) at NVIDIA GTC Europe.

2018 **Cross-Model Hallucination for Few-Shot Learning**, Invited Talk at Max-Dellbrück Center for Molecular Medicine.

Selected Research Projects

Aug 2023 – Jan 2024 **Revisting Concept Activation Vectors to Overcome Directional Divergence**, Fraunhofer Heinrich-Hertz Institute.

We show that separability-oriented computation of concept activation vectors (CAVs), e.g., with SVMs, leads to solutions, which may diverge from the actual goal of precisely modeling the concept direction and introduce pattern-based CAVs, solely focussing on concept signals, thereby providing more accurate concept directions.

Mar 2023 – Aug 2023 **Right Reason CIArC - Model Correction via Explicit Gradient Penalization in Latent Space**, Fraunhofer Heinrich-Hertz Institute.

We introduced a novel model correction method, named RR-CIArC, ensuring the right reasons for predictions on the concept level by reducing the model's sensitivity towards biases by explicitely penalizing the model for the use of undesired data artifacts.

Aug 2022 – Mar 2023 **XAI Life Cycle for Iterative Model Improvement in Medical Applications**, Fraunhofer Heinrich-Hertz Institute.

We proposed an XAI life cycle in which model misbehavior can effectively be revealed and corrected in an iterative fashion. The framework was successfully evaluated for two medical tasks, namely skin cancer detection (ISIC challenges) and pediatric bone age prediction.

Oct 2017 – Jul 2018 **Multi-modal Few-shot Learning using Fine-grained Image Descriptions (Master's Thesis), Grade: 1.0 (*scale: 1-5, 1 is best*)**, Humboldt-Universität zu Berlin/SAP AI Research.

I investigated how multi-modal information can be used to train accurate classification models in few-shot scenarios. Specifically, I analyzed strategies for class-discriminative cross-modal data generation in order to be able to extend the limited visual training set conditioned on data in different modalities.

Extra-Curricular Activities

2020 **AI for Medicine**, *deeplearning.ai*.

2018 **Deep Learning Specialization**, *deeplearning.ai*.

2006 – 2012 **Mathematical Student Society "Leonhard Euler"**, *Humboldt-Universität zu Berlin*.

Coursework on mathematical problems for mathematically talented high school students.

Computer Skills

Programming	Python, Java, Scala, R, Octave, MySQL, ABAP, JavaScript
Frameworks	Tensorflow, Keras, PyTorch, zennit, scikit-learn, MXNet, NLTK
Other Tools	Git, LaTeX, Microsoft Office

Languages

German	*Native*
English	*Fully Professional*