

Katja Dittrich

Berlin - 01525 6194651 - katja.dittrich@outlook.com

www.dittrichkatja.com - github.com/coztomate - linkedin.com/in/katja-dittrich

SKILLS

Technical Skills:

- Python / SQL / Git / Bash (intensive use for 1 year), R / R Studio (5 years), IBM SPSS (10 years).
- Machine Learning (Trees, Computer Vision, NLP, LLMs, Transfer Learning), Data Visualization.
- Tools: Tensorflow, Numpy, Pandas, Transformers, LangChain, Matplotlib, Plotly, Docker, Streamlit, Huggingface, Optuna, Google Cloud, Qdrant Vector Store, Weights & Biases.

Soft Skills:

Fluent communication in German and English with technical and non-technical audiences, storytelling, project management, independent working, analytical and critical thinking, problem-solving.

PROFESSIONAL EXPERIENCE

Charité – Universitätsmedizin Berlin: *Research Associate: November 2013 to April 2023*

Cross-Project Responsibilities:

- Management of projects throughout their entire lifecycle, from defining project goals and ensuring funding to implementation within the set timeframe and publication.
- Processing and analysis of cross-sectional and longitudinal datasets (with SPSS and R): linear, logistic, Poisson regression, path analysis, structural equation modeling (SEM), growth curves, mediation analysis, discriminant analysis, (M)ANCOVA, factor analysis.
- Publishing results in international journals and presenting at national and international congresses.
- Teaching in the medical degree program (in "Communication, Interaction, Teamwork").

Senior Scientist: March 2020 to April 2023

- Coordination of a multicenter care study (DEPSY) and a subproject of a multicenter intervention study (UBICA II).
- Supervision of master's and doctoral students and supervision of student assistants.
- Data management and quality monitoring.
- Development of web-based database (REDCap), survey tools (SoSci Survey, LimeSurvey), study infrastructure, and teams for the implementation of studies.

Junior Scientist: November 2013 to February 2020

- Collection of psychometric, psychobiological data (interviews, questionnaires, IQ tests, video diagnostics, saliva, heart rate variability, MRI).
- Recruitment of study participants for two multicenter studies (UBICA, BerlinLCS).

Freiraum Psychotherapie, Center for Psychotherapy at HU Berlin (ZPHU) & St. Hedwig Hospital Acute

Psychiatry: *Psychologist, Psychotherapist: July 2017 to March 2023*

- In summary: Application of Cognitive Behavioral Therapy techniques, individual and group therapy, social skills training.

EDUCATION

Data Science Retreat Berlin - *Full-time Data Science Training: October 2023 to December 2023*

Center for Psychotherapy at HU Berlin: *License as a Psychotherapist (Cognitive Behavioral Therapy): April 2021*

Humboldt University of Berlin / Charité Berlin - *Ph.D. in Natural Sciences (Dr. rer. nat.): February 2020*

Humboldt University of Berlin - *Psychology Studies (Diploma): October 2007 to October 2013*

SELECTED PROJECTS (with Links)

Science Query: Keyword-Driven Abstract Search - January 2024 (current)

(Python, LangChain, OpenAI, Streamlit, ChromaDB)

- Utilization of the ArXiv and PubMed APIs for querying articles by search terms.
- Implementation of Retrieval Augmented Generation (RAG) to create a question-answer app using GPT3.5-Turbo.

Image Similarity Search for Brain Tumor Diagnoses - December 2023

(Python, Google Cloud & Colab, Qdrant, Tensorflow, Sklearn)

- Team project applying Convolutional Neural Networks and Transfer Learning for classification of tumor types and embedding extraction.
- Development of a vector database for similarity search of MRI images and automated querying from the cloud.

Web Scraping and Transfer Learning for Cat Breed Classification - December 2023

(Python, Tensorflow, WandB, Sklearn)

- Web scraping to collect relevant images and creation of a classification model using Transfer Learning.

Predicting Earthquake Building Damage (DrivenData Competition) - October 2023

(Python, XGBoost, Optuna, Sklearn, Plotly)

- Achieved top 5% worldwide placement by employing Gradient Boost techniques and automatic hyperparameter optimization.

Study on Intergenerational Transmission of Mental Disorders - May 2012 bis December 2017

(SPSS, R, Process, Lavaan, Ggplot2)

- Application of multiple linear and logistic regression, path analysis, ANOVA.
- Publication of 16 scientific articles, including 6 as first author and 1 as last author (supervisor role), in journals such as the British Journal of Psychiatry, Brain Behavior and Immunity, Psychological Medicine, and 9 conference presentations (national and international).

FURTHER EDUCATION

DeepLearning.AI:

- *October 2023 to January 2024:* Machine Learning Specialization (including Supervised Machine Learning: Regression and Classification, Advanced Learning Algorithms, Unsupervised Learning, Recommenders, Reinforcement Learning).
- *Currently:* Deep Learning Specialization (Neural Networks and Deep Learning, Structuring Machine Learning Projects, Improving Deep Neural Networks: Hyperparameter Tuning, Regularization and Optimization, Convolutional Neural Networks, Sequence Models).

DataCamp

- *December 2023 to January 2024:* Introduction to SQL, Intermediate SQL and Joining Data in SQL, Introduction to Power BI.
- *February 2023:* Data Analyst with R (including data manipulation, exploratory data analysis, hypothesis testing).

Hackerrank Zertifizierung

- *January 2024:* SQL Basic and Intermediate.

Udemy

- *Currently:* ChatGPT and LangChain: The Complete Developer's Masterclass.
-

Berlin, 18 January 2024

