+491601691993 | davidsiriusschulte@gmail.com | davidschulte.github.io | LinkedIn: davidsiriusschulte

David Schulte

WORK EXPERIENCE

Data Scientist   Berlin, DE
Vaayu   9/2022 – present
- Ownership of an NLP model that predicts ~1M product taxonomies per day
- Large-scale analysis of retail shipping behavior and trends
- Simulating the impact of reduction measures on logistics carbon emissions

Data Scientist (Working Student)   Berlin, DE
Jobufo   3/2022 – 8/2022
Development of a sequence labeling model for structured texts

EDUCATION

Statistics (M.Sc)   Berlin, DE
Humboldt University of Berlin   10/2020 – 3/2024
Specialization in Data Science and Machine Learning

Mathematics (B.Sc)   Cologne, DE
University of Cologne   10/2014 – 9/2019
Courses in Probability Theory, Numerical Simulation and Computer Science

Computer Science (Semester abroad)   Tokyo, JP
Waseda University   4/2018 – 8/2018
Courses in Machine Learning, Data Mining and Information Security

PUBLICATIONS

Schulte, D., Hamborg, F., & Akbik, A. (2024). Less is More: Parameter-Efficient   EMNLP 2024
Selection of Intermediate Tasks for Transfer Learning arXiv preprint arXiv:2410.15148.

PROJECTS

hf-dataset-selector
A Python package for efficient selection of datasets for intermediate task transfer learning

Prediction of Customer Returns
Created and evaluated a model that predicts returns to an online retailer

Reinforcement Learning for 3-Player Chinese Checkers
Designed and trained an agent to play a multiplayer board game

SKILLS

Machine Learning, NLP, Model Deployment, Data Analysis, Python, PyTorch, GCP, Kubernetes, Git