Andreas Stephan

Personal data

Birth: 03/08/1993 in Munich

Address: 1020 Vienna, Austria

E-mail: andreas.stephan@univie.ac.at

Website: https://andst.github.io/

Education

University of Vienna
PhD student *01/21 - now*

TU Munich
M.Sc. Mathematics in Data Science *10/17 - 11/19*

University of California, Santa Barbara, USA
Semesters abroad *08/16 - 04/17*

TU Munich
B.Sc. Mathematics *10/14 - 09/17*

TU Munich
B.Sc. Computer Science *10/11 - 09/14*

Experience

Check24 GmbH
Data Scientist (NLP) *03/20 - 12/20*
Experimentation and productionizing of classification and information extraction tasks in the domains insurance and credit scoring.

BMW GROUP / TUM
Master Thesis *04/19 - 10/19*
Comparison of knowledge graph embeddings for entity alignment between graph and vector data. (Grade: 1.0)

BMW GROUP
Internship *10/18 - 03/19*
Development of Natural language processing (NLP) building blocks in Java and Python. Tasks included document classification, document clustering or training data generation with Snorkel.

PuzzlePie KG
Software development *09/17 - 03/18*
Enhancing an event management tool using ReactJS and NodeJS.

Unify Software and Solutions GmbH & Co. KG
Internship *08/15 - 10/15*
Development of a meeting assistant for the enterprise collaboration platform Circuit.

Teaching

WS 23/24: Deep Learning for NLP

SS 23: Scientific Data Management

WS 22/23: Deep Learning for NLP

SS 22: Introduction to Mathematics for Computer Scientists

WS 21/22: Siminar: Weakly Supervised Learning

Invited Talks

02/23: VW Data Labs, Munich

12/21: Workshop: Austrian Meeting on Sentiment Analysis (ÖTSI)

Skills

Languages and Tools:

- Python, JavaScript, some C++
- Linux, Docker, Git
- SQL, Neo4J, MongoDB

Libraries:

- PyTorch, JAX
- ScikitLearn, Pandas, Numpy
- NodeJS, ReactJS

Papers

WTYT: Dawei Zhu, Xiaoyu Shen, Marius Mosbach, Andreas Stephan, and Dietrich Klakow. 2023. Weaker Than You Think: A Critical Look at Weakly Supervised Learning. In Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), pages 14229–14253, Toronto, Canada. Association for Computational Linguistics.

SepLL: Andreas Stephan, Vasiliki Kougia, and Benjamin Roth. 2022. SepLL: Separating Latent Class Labels from Weak Supervision Noise. In Findings of the Association for Computational Linguistics: EMNLP 2022, pages 3918–3929, Abu Dhabi, United Arab Emirates. Association for Computational Linguistics.

WeaNF: Andreas Stephan and Benjamin Roth. 2022. WeaNF":" Weak Supervision with Normalizing Flows. In Proceedings of the 7th Workshop on Representation Learning for NLP, pages 269–279, Dublin, Ireland. Association for Computational Linguistics.

Knodle: Sedova, A., Stephan, A., Speranskaya, M. & Roth, B. (2021). Knodle: Modular Weakly Supervised Learning with PyTorch. Proceedings of the 6th Workshop on Representation Learning for NLP (RepL4NLP-2021), 100–111. https://doi.org/10.18653/v1/2021.repl4nlp-1.12