

# Ahmad Dawar Hakimi

ELLIS PHD STUDENT · LMU MUNICH ↔ UNIVERSITY OF COPENHAGEN

✉ ah.dawar.hakimi@gmail.com | 🏠 adhakimi.github.io | 📧 adhakimi | 🌐 adhakimi

## Education

### LMU Munich ↔ University of Copenhagen

Germany & Denmark

ELLIS PHD IN NATURAL LANGUAGE PROCESSING (ANTICIPATED GRADUATION: MARCH 2027)

April 2024 - Present

- Supervisors: Hinrich Schütze & Isabelle Augenstein
- Research Areas: Interpretability, Active Learning, Summarization, Factuality
- Current Topic: Analyzing the Learning Dynamics, Modularity, and Knowledge Representation of LLMs

### Leipzig University

Germany

M.SC. IN COMPUTER SCIENCE - FINAL GRADE: 1.2 (GPA: 4.00/4.00)

Oct 2019 - Aug 2023

- Master Thesis:** Citance-Contextualized Summarization of Scientific Papers
- Awarded **Deutschlandstipendium** - Scholarship for outstanding academic performance
- Compiled a high-quality, large-scale dataset, "Context-SciSumm" for citance-contextualized summarization
- Developed a three-step methodology to generate contextualized summaries for scientific documents using LLMs such as LLaMa, Alpaca, Vicuna, Falcon, and GPT-4, and evaluated each step through both human and automatic assessments (published at EMNLP 2023)

### Leipzig University

Germany

B.SC. IN COMPUTER SCIENCE - FINAL GRADE: 2.1 (GPA: 3.70/4.00)

Oct 2016 - Sep 2019

- Bachelor Thesis:** Using Syntactic Information for Sentence Embeddings
- Enhanced the expressiveness of sentence embeddings using syntactical information in a Weighted Bag of Words Model with chunk weighting and an Encoder-Classifer Model integrating syntactic structures for downstream NLP tasks

## Experience

### Machine Learning Engineer

Leipzig, Germany

INSTITUT FOR APPLIED INFORMATICS - RESEARCH PROJECT KIRESys

Dec 2022 - March 2023

- Developed a prototype for extracting information from PDFs of electronic parts using *Object Segmentation Models* like YOLO and R-CNNs
- Introduced a *Bert-based* keyword generation for electronic components to enhance search capabilities
- Generated comprehensive summaries for technical documents with *Transformer Models*

### Research Assistant

Leipzig, Germany

NLP DEPARTMENT AT LEIPZIG UNIVERSITY - RESEARCH PROJECT FAME

April 2019 - Dec 2022

- Leveraged syntactical information to improve the expressiveness of sentence embeddings, for sentences with similar structure
- Designed an adversarial test set for Same-Side Stance Classification with challenging linguistic cases, such as negations and opposing positions within single arguments, demonstrating a 43.2% drop in model accuracy compared to the standard set, revealing the shallow understanding of *SOTA Transformer Models* (published at EMNLP 2021)
- Introduced the novel task of Same Sentiment Classification (published at EMNLP 2021)

### Student Research Assistant

Leipzig, Germany

NLP DEPARTMENT AT LEIPZIG UNIVERSITY - RESEARCH PROJECT ILCM

Oct 2017 - March 2019

- Implemented an R6 Cooccurrence Class, a customized logging function, refined the calculation of the context volatility, optimized and extended the Text Processing Pipeline, and added a Black- and Whitelist for the tokenization step in R
- Designed test cases and technical documentation
- Organized the lab classes for the bachelor's course in "Algorithm and Data Structures 1 & 2"

## Publications

- Hakimi, A.D.**, Modarressi, A., Wicke, P. and Schütze, H., 2025. Time Course MechInterp: Analyzing the Evolution of Components and Knowledge in Large Language Models. In Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics
- Hakimi, A. D.**, Syed, S., Al-Khatib, K. & Potthast, M. (2023, November). Citance-Contextualized Summarization of Scientific Papers. In Findings of the Association for Computational Linguistics: EMNLP 2023.
- Meinecke, C., **Hakimi, A. D.**, & Jänicke, S. (2022). Explorative Visual Analysis of Rap Music. Information, 13(1), 10
- Körner, E., Wiedemann, G., **Hakimi, A. D.**, Heyer, G., & Potthast, M. (2021, November). On Classifying whether Two Texts are on the Same Side of an Argument. In Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing (pp. 10130-10138).
- Körner, E., **Hakimi, A. D.**, Heyer, G., & Potthast, M. (2021, November). Casting the Same Sentiment Classification Problem. In Findings of the Association for Computational Linguistics: EMNLP 2021 (pp. 584-590).
- Baumann, R., Wiedemann, G., Heinrich, M., **Hakimi, A. D.**, & Heyer, G. (2020). The Road Map to FAME: A Framework for Mining and Formal Evaluation of Arguments. Datenbank-Spektrum, 20(2), 107-113.

## Skills & Languages

LANGUAGES	German (Native), Dari (Native), English (Fluent), Latin (Intermediate), French (Basic), Russian (Basic)
TECH	Python, PyTorch, TensorFlow, Huggingface, Detectron2, Numpy, Pandas, R, Docker, Git
INTERESTS	NLP, LLMs, ML, Interpretability, Knowledge Representation, Summarization, Object Detection
ACTIVITIES	Running, Bodyweight Fitness, Road Cycling, Table Tennis, Basketball