Maximilian Weissenbacher

AI Engineer @PwC
Phone: +49(0)17656955718
Email: maxweissenbacher98@gmail.com
Portfolio: https://maxiweissenbacher.github.io/max-weissenbacher/
Socials: LinkedIn, GitHub, HuggingFace, Medium, Scholar, Ollama
Skills: Python, GenAI, Data Science, Machine Learning

EDUCATION

2021 - 2023 **University of Regensburg** Masters Information Science
- Grade: 1.2 (Germany), equivalent GPA 3.8
- Focus on Data Science (Top 5% of class), Natural Language Engineering with Python, Machine Learning, Deep Learning, Research

2018 - 2021 **University of Regensburg** BA Media- and Information Science
- Courses: Natural Language Processing with Python, OOP in Java, Algorithms & Data Structures, Information Retrieval (Grade: 2.3 Germany)

2016 **Gymnasium LSH-Marquartstein** A-Level (Abitur), GPA 2.7 (Germany)

WORK EXPERIENCE

2023 – Present **AI Engineer at PwC, Munich (current position)**
- Architected and implemented open-source LLM deployments (Gemma 3, Mistral Small 3.1) on secure government infrastructure, enhancing AI capabilities for multiple federal institutions.
- Spearheaded development of innovative GenAI applications for German government agencies, including the Ministry of Finance and ITZBund:
  - ***Legal impact assessment tool***: Tool to check if a ministry is affected by a law or not. Reduced analysis time by 60%.
  - ***Multimodal Confluence chatbot***: Increased employee productivity by 25% through advanced RAG techniques and agentic pipelines with tool-calling. This chatbot secured the second place in Wirtschaftswoche's magazine prestigious "Best of Consulting Awards."
  - ***Legal term modularization app***: Knowledge Graphs in combination with RAG (GraphRAG) to identify different definitions and interpretations of law terms.
- Conducted time series forecasting using XGBoost to predict passenger flow on Munich's public transport system.

2023 **Research Assistant – Machine Learning, University of Regensburg, 6 Months**
- Development of an early warning system for depression in children and adolescents using BERT (Text Classification) and GPT-3 (Data Augmentation, Few-Shot Learning), and ChatGPT (Prompt Engineering).

2022 **Data Science Intern at Amazon, London, 4 Months**
- Building multivariate time series models on AWS to forecast the incoming units per day on carrier level for the two biggest warehouses in the UK.
- Ensemble Learning with built Machine Learning (LightGBM, FbProphet)- and Deep Learning Models (Amazon DeepAR) increased WK+1 forecast accuracy by 6.6% (660bps) and Intraweek forecast accuracy by 7.73% (773bps).
- Back-Testing showed possible savings for 2021 in a seven-digit range (GBP).
- The Model got expanded to the whole UK and EU network.

2021 - 2022 **Data Analytics Working Student at SAP, Munich, 11 Months**
- Performance analytics of SAP Learning platforms (see video of myself here).

- Forecasting Data and automating processes with Python scripts.

2019 - 2020 **Oracle APEX Development Internship at BMW, Regensburg, 4 Months**
- (Re-)Engineered an Oracle APEX Application for customer orders using SQL, PL-SQL, and JavaScript. The application was exported to BMW plants in the UK.

PUBLICATIONS & PROJECTS

2024 **Publication: "MINI-KID interviews with adolescents: a corpus-based language analysis of adolescents with depressive disorders and the possibilities of continuation using Chat GPT".** *Frontiers in Psychiatry 15 (2024): 1425820. (Link). Details in Work Experience under Research Assistant.*

2024 **Publication: "Analyzing Offensive Language and Hate Speech in Political Discourse: A Case Study of German Politicians".** *Proceedings of the Fourth Workshop on Threat, Aggression & Cyberbullying@* ***LREC-COLING-2024****. 2024. (Link)*
- Developed an ensemble of transformer models achieving 0.94 F1-Macro score for hate speech detection
- Engineered data pipeline to mine and process >2.7 million tweets using Twitter API
- Created Streamlit-based annotation interface with AWS for efficient crowdsourced data labeling (Medium article).
- Conducted large-scale analysis of hate speech patterns in German political discourse

2023 **Publication: "Steps towards Addressing Text Classification in Low-Resource Languages".** *Proceedings of the 19th Conference on Natural Language Processing (****KONVENS 2023****).(Link)*
- Several approaches, like data augmentation, to increase text classification accuracy on resource-poor languages using SOTA Large Languages Models (LLM).
- Text Classification with Fine-Tuning BERT,- XLM-RoBERTa-Models

2022 **Publication: "Sentiment Analysis on Twitter for the Major German Parties during the 2021 German Federal Election".** *Proceedings of the 18th Conference on Natural Language Processing* ***(KONVENS 2022****). 2022. (Link)*
- Data Mining and Annotation of 60k texts
- Sentiment Analysis with LLMs like fine-tuning BERT
- Data Analysis with insights about the political online communication in Germany

2019 – 2023 **Key University Projects:**
- Car Price Prediction (2021): Scraped data and built ML models (Linear Regression, Random Forest) in R.
- Fake Trump Tweet Generation Chatbot (2020): Developed using Python, AWS, and ML techniques (K-Means, GloVe, PCA).
- Application Manager (2020): Built resume/cover letter generator using JavaScript and Google Firebase.
- Voluntary Study-accompanying IT education. Certified by the University of Regensburg
- Recipe Search Assistant: Information Retrieval Project with Elasticsearch, SQL and Java

SKILLS & INTERESTS

Programming Skills:	Python, R, JavaScript, SQL, Java, HTML, CSS
Technologies:	MCP, Langgraph, FastAPI, Ollama, Neo4J, Postgres, Docker, Azure, GCP, AWS, Transformers, PyTorch, Unsloth, vLLM
Languages:	German (native), English (fluent), Italian (A1), French (A1)
Interests & Hobbies:	Data Science, Machine Learning, NLP, Soccer (captain), Kitesurfing, Skiing Thank you for reading up to this point!