github.com/max-gartz
Maximilian Gartz
gartz.maximilian@gmail.com ✉

Maximilian Gartz

Software Engineer - Machine Learning, Berlin, Germany

SKILLS

- **Specializations:** Machine Learning, Software Engineering, MLOps & DevOps
- **Languages:** ***Python***, Java, Bash, SQL | German (native), English (professional proficiency)
- **Frameworks & Libraries:** FastAPI, PyTorch, TensorFlow, Huggingface Transformers, LangChain
- **Tools:** Git, Terraform, dbt, Docker, Kubernetes, Grafana, MLflow, Weights & Biases, Kubeflow, Flyte, Triton
- **Databases:** PostgreSQL, MySQL, BigQuery, Firestore, Prometheus
- **Clouds:** ***GCP***, Azure

WORK EXPERIENCE

Machine Learning Engineer @ML6, *Berlin, Germany* *06/2021 – present*

- Exploring applications of LLMs for a US Retail giant, while implementing DevOps practices across various streams of work. Deployments utilize the Azure Container Apps service.
- Leading the development of an internal standard solution for LLM and RAG projects, utilizing FastAPI and LangChain, designed to seamlessly integrate various retrievers, LLMs, and embedding models via a central configuration file. Deployments on GCP and AWS utilize various managed services, while on-prem deployments leverage Docker Compose or Kubernetes. This approach has significantly driven down development cost and enabled us to offer competitive pricing for RAG solutions.
- Took on a leading role within our internal DevOps/MLOps team, driving the formulation and implementation of standards and best practices in both software development and Machine Learning, enabling high quality project delivery.
- Served as Tech Lead on a demand forecasting project for a large retail company in Germany. Achieved a nearly 10% revenue increase by improving warehouse replenishment through the deployment of classical time series models. Lead the transition of POC stage solutions into production on GCP, leveraging Terraform, BigQuery, dbt, Cloud Run, and Vertex AI, while streamlining build and deployment processes.
- Acted as Lead Engineer for an MLOps-centric project at a Swiss insurance company. Implemented significant improvements in development workflows and established a standardized GitHub template with robust CI/CD pipelines for ML projects on GCP, adhering to stringent security protocols. Focused on end-to-end ML pipelines utilizing Vertex AI and facilitated extensive knowledge transfer to client's Data Scientists and Engineers.
- Took up an Engineer role in a POC aimed at detecting asbestos fibers in high-resolution SEM images using the Tensorflow Detection API and SAHI. Demonstrated the capability to automate fiber detection with human level accuracy, culminating in a labor reduction of up to 80%. Concurrently authored a well-received blog post, elucidating the basic approach using the MMDetection framework and SAHI for small object detection in satellite imagery.
- Repeatedly mentored new recruits during their onboarding periods and first projects

Software Engineer @InsideM2M GmbH, *Osnabrueck, Germany* *05/2018 – 12/2020*

- Developed REST APIs for client-facing applications in IoT projects utilizing Java EE and built micro-services with Quarkus. Learned about containerization with Docker and deployment on Docker Swarm managed with portainer. Ensured code quality through unit testing with JUnit.
- Conducted analysis and comparison of time series databases InfluxDB and TimescaleDB to evaluate possible performance enhancement in handling transactional data from IoT devices over basic relational database solutions. Investigated downsampling and retention policies as well as visualization with Grafana.

EDUCATION

Bachelor of Cognitive Science @University of Osnabrueck, *Osnabrueck, Germany* *10/2017 – 03/2021*

- Graduated with distinction and final grade 1.0 (German) equivalent to 4.0 GPA in USA
- Focused on Machine Learning, Bayesian statistics and Mathematics
- Bachelor's thesis: Bayesian Inference with Hamiltonian Normalizing Flows

CERTIFICATIONS

Google Cloud Certified Professional Machine Learning Engineer *since 10/2021*