Nikolai Klassen

Berlin

lorem-ipsum   mail me   www.github.com/Nikolaikl

Experience

Deutsche Rentenversicherung Bund   **Berlin, Germany**
Machine Learning Engineer / Team Lead   *Nov. 2024 – Ongoing*

- **Lead the development and deployment of a multilingual generative AI platform, covering the frontend and backend** on AWS, enabling the development of LLM workflows across multiple teams and enhancing cross-functional collaboration aimed at serving 20,000 user requests per hour.
- **Set up monitoring and LLM-engineering** platform tools to monitor and evaluate LLM-based applications for the accuracy and reliability of our service
- **Developed and deployed ETL pipelines** that transform unstructured data into structured data for optimized LLM ingestion
- **Designed and implemented scalable Retrieval-Augmented Generation (RAG) pipelines**, optimizing the processing and retrieval of insights from large-scale text datasets and significantly improving and automating data processing pipelines for unstructured data.
- **Collaborated with Architects to shape the generative AI setup**, integrating cloud technologies into both AWS and on-premise solutions, ensuring a seamless hybrid infrastructure and integration.
- **Led the usage of Requests for Comments (RFCs) and Architectural Decision Records (ADRs) for Generative AI projects**, driving the adoption of innovative practices in software development processes. Furthermore, I introduced and conducted Postmortems after outages.
- **Mentored and guided junior data scientists in both LLM and classical ML projects**, fostering their technical skills and professional growth through structured feedback and collaborative problem solving, leading to improved project results and greater confidence.

Billie.io   **Berlin, Germany**
Data Engineer / Data Scientist   *Oct. 2022 – Oct. 2024*

- **Developed and maintained backend services in Python and Kotlin** in accordance with **MLOps best practices**, achieving **99% availability** for **20,000 daily orders**.
- **Designed CI/CD pipelines** to ensure reliable testing and deployment of **ML-powered systems** utilizing **GitHub Actions**, **Kubernetes**, and **Jenkins**.
- **Collaborated with stakeholders** to gather both **technical** and **non-technical requirements** for new **ML services** and features.
- **Enhanced ML models for classification**, resulting in a **20% reduction** in defaults and fraud through improved data processing techniques.
- **Implemented real-time pipelines** to process semi-structured business events from multiple sources using **Apache Kafka** and **Apache Flink**.

Skills

Programming Languages: Python, SQL, JavaScript, Typescript, Java, Kotlin, Go, HCL
Software: Apache Kafka, Apache Flink, Python ML Stack (TensorFlow, PyTorch, Scikit-learn, XGBoost, Pandas, ...), Python Frameworks (uv, ruff, pytest, fastapi), AWS (DynamoDB, EC2, Lambda, RDS, Redshift, S3, SQS, Bedrock, IAM), , ArgoCD, Kubernetes, Snowflake, Jenkins, ELK Stack
Skills: Machine Learning, LLMOps / MLOps, Test Driven Development, DevOps, Data Engineering, Architecture Design
Languages: German (Native), Russian (Native), English (C2), Italian (B1), Spanish (B1)

Education

Eberhard Karls University of Tübingen   **Tübingen, Germany**
M.Sc. Data Science (**NC:** 1.7)   *Oct. 2019 – Oct. 2022*
University of Bayreuth   **Bayreuth, Germany**
B.A. Development Economics   *Oct. 2014 – Oct. 2018*