

Loay Nasser

AI Engineer

loay.nasserr@gmail.com | +49 15224007210 | Germany

LinkedIn: linkedin.com/in/loaynaser | GitHub: github.com/LoaiAlaa24

PROFESSIONAL SUMMARY

Experienced Software Engineer with experience in AI and Machine Learning, specializing in Python, Deep Learning, and Cloud technologies. Expertise in developing and implementing AI solutions, including LLMs and Diffusion Models. Master's degree from TUM. Strong problem-solving skills with a focus on building efficient and scalable applications.

EXPERIENCE

AI Automation Engineer at Legartis GmbH

Leipzig, Germany | December 2024 – Present

- Implemented a Multi-Modal RAG sales agent chatbot that accelerated the sales process by 70% and improved lead qualification accuracy by 95%.
- Engineered N8N workflows to automate internal processes, integrating with People Data Labs, HubSpot, Grain, and LLMs, which reduced task completion time by 99% and optimized workflow efficiency by 40%.
- Developed prototypes using ReactJS and FastAPI, achieving a 60% reduction in development time and improving API response time by 4x.

AI Engineer (Remote) at EFG Holding

Cairo, Egypt | December 2024 – Present

- Developed LangGraph RAG AI agents with FastAPI on Azure, integrated into Microsoft Power Automate to build a smart chatbot that improved EFG's internal workflow efficiency by 70% and reduced query response time by 40%.
- Engineered multi-modal RAG pipelines enabling LLMs to query across text, images, and tables for enhanced information retrieval, improving precision from 78% to 95%.
- Leveraged Langsmith (LLMOPs) to monitor, test, and debug the LLMs pipeline during development and production, reducing debugging time by 50% and achieving 85% test coverage.
- Evaluated the RAG pipeline using RAGAS metrics to ensure efficiency and accuracy, optimizing model performance and improving response relevance by 30%.
- Performed data pre-processing and transformation tasks using Pandas and SQL, ensuring data quality and feature readiness for advanced analytics, reducing data loading time by 60% and optimizing query performance by 3x.

Gen AI Researcher at Volkswagen Group

Wolfsburg, Germany | December 2023 – September 2024

- Researched and graded (1.0, highest distinction) Generative AI Diffusion Models for generating complex corner case street scenes to enhance autonomous driving algorithms, improving scene generation time by 40%.
- Implemented complex corner case street scenes in Unity game engine and transformed them into realistic scenes using generative AI, optimizing scene rendering time by 30%.
- Generated high-quality synthetic data leveraging state-of-the-art diffusion models to enhance autonomous driving testing datasets, increasing dataset size by 5x and improving model accuracy by 15%.
- Fine-tuned Stable Diffusion XL with LoRA on Audi's A2D2 Dataset, achieving superior domain adaptation for autonomous vehicle applications.

Software Engineer (AI) at Hotdesk

Dubai, United Arab Emirates | September 2022 – December 2023

- Developed Hotdesk's recommendation engine using machine learning techniques, providing tailored listing recommendations based on proximity, price, amenities, and user behavior, which improved user retention by over 20% and increased click-through rate by 15%.
- Implemented a CNN-based machine learning model with OpenCV to assess listing image quality by detecting watermarks, blur, and brightness, ensuring optimal content standards and reducing low-quality images by 40%.
- Optimized ETL pipelines on Google Cloud Platform (GCP), achieving a 65% reduction in operational costs and improving data processing speed by 50%.
- Deployed machine learning models using AWS Lambda and Flask API, improving API response time by 30% and reducing latency by 25%.
- Generated impactful visualizations with Google Data Studio to support data-driven business decisions, improving report generation time by 70% and enhancing data accessibility for stakeholders.

Software Engineer (AI) at Stadtwerke München

Munich, Germany | October 2022 – December 2023

- Designed and deployed an NLP-based email classification system, leveraging a fine-tuned German BERT model to categorize German emails into 10 distinct topics, improving organizational workflow and response efficiency by 40%.
- Developed advanced machine learning models using PyTorch, with a focus on large language models (LLMs) to address complex business challenges, accelerating model training time by 30% through optimized hyperparameter tuning.
- Implemented a robust data anonymization pipeline for PDF documents, integrating OCR and NLP entity recognition techniques using fine-tuned BERT, ensuring data privacy and reducing manual review time by 50%.

EDUCATION

Master of Science in Informatics

Technical University of Munich (TUM) | Munich, Germany | September 2024

- GPA: 1.9
- Major: Machine Learning and Analytics
- Minors: Computer Vision / Security

Bachelor of Science

German University in Cairo (GUC) | Cairo, Egypt | July 2020

- GPA: 1.8
- Computer Science and Engineering
- Thesis Grade: 1.0

SKILLS

Programming Languages: Python, JavaScript, Java

LLM Frameworks: Langgraph, LangChain, LangSmith, Pydantic

Machine Learning: CNN, NLP, Diffusion Models

Data Analysis: Pandas, NumPy, Matplotlib, Seaborn, Scipy

CERTIFICATIONS

- **Deep Learning Specialization** - Coursera Taught by: DeepLearning.AI (Dr. Andrew Ng) (2019)