Sanath Vijay Haritsa

(+49)17648994438 | sanath.vh@gmail.com
linkedin.com/in/sanath-haritsa-974686315/ | github.com/sanath95
Mannheim, Germany

Summary

MSc in Applied Data Science graduate with **3+ years of work experience** across Germany and India, specializing in generative AI. Brings a scientific approach, strong work ethic, and a passion for building reliable AI systems. Currently seeking a full-time role in data science or generative AI.

Work Experience

Schwarz IT Heilbronn, Germany
Working Student in Generative AI *01/2025–09/2025*
Designed and tested multiple PoC pipelines to explore real-world applications of large language models across automation and content-evaluation.

• **Job Reference Generation:** Built a large language model system with **few-shot examples** for compliant German job-reference creation. Implemented a second model as a judge that assigned a score (8.42/10) to ensure accuracy and policy adherence.

• **Social Media Content Automation:** Delivered a no-code **n8n** workflow integrating chat models and Google Sheets to auto-generate internal communication posts with 2-3 day development cycles, showcasing rapid prototyping.

• **Dual-Quality Detection:** Designed a semantic-similarity pipeline on **Databricks + MLflow** to flag ingredient mismatches in EU food products, achieving precision 0.82 / recall 0.80 and demonstrating multilingual generalization.

• **No/Low-Code Platform Evaluation:** Developed a **KPI-based evaluation framework** to assess tools like n8n, ServiceNow, and Google Agent Space, enabling business teams to build AI workflows without coding.

Nichesolv Private Limited Bangalore, India
Data Scientist *06/2020-09/2023*
• **Object-oriented programming (OOP)**, clean coding principles, and backend development using Java and .NET (C#)

• **Computer Vision:** Used OpenCV, YOLO, and TensorFlow to automate tennis match analysis and track player performance.

• **Time Series:** Trained ARIMA model to predict water pump cavitation in advance, enabling proactive maintenance and less downtime.

• **Explainable AI:** Benchmarked LIME, SHAP, Anchors for model interpretability, identifying SHAP as the most stable framework.

Education

SRH Hochschule Heidelberg, Germany
M.Sc. in Applied Data Science and Analytics *10/2023 - 09/2025*
Grade: 1.5

Academic Projects

• **Master's Thesis – Layout-Aware Metadata Extraction Framework** [04/2025–09/2025] Developed a metadata extraction system for scholarly PDFs comparing layout-aware GROBID against small language models(<=4B parameters), proving SLMs can outperform layout-aware parsers on variable fields (e.g., affiliations, emails) in accuracy.

• **Capstone Project (collaboration with ONTEC AG) – Making Tables Searchable for LLMs** [04/2024–09/2024] Built a table extraction pipeline that preserves cell relationships using multimodal large language model, achieving 0.90 cosine similarity and enabling multilingual table search within RAG workflows.

• **Responsible AI - Predicting Healthcare Test Results** [08/2024–09/2024] Created an explainable ML pipeline using CatBoost, SHAP, and bias analysis under EU AI Act principles, revealing critical performance–fairness trade-offs in healthcare prediction models.

• **Chatbot for Academic Support:** [07/2024–08/2024] Developed a local LLM-based chatbot using Llama 3 (Ollama) and an information retrieval (RAG) system to assist MSc students, evaluated by students on in-context and out-of-context queries, demonstrating reliable and context-aware academic assistance.

RNSIT Bangalore, India
Bachelor of Engineering - Mechanical Engineering *2013 - 2020*
Percentage: 59

Technical Skills

- **Programming Languages:** Python, R, SQL
- **Libraries:** NumPy, Pandas, Matplotlib, Seaborn, Scikit-learn, SHAP
- **Deep Learning:** TensorFlow, PyTorch, Keras
- **Generative AI:** Prompt and Context Engineering, Retrieval Augmented Generation, AI Agents
- **AI Frameworks** LangChain, LangGraph, LlamaIndex, smolagents, n8n, VertexAI
- **MLOps:** MLflow, Databricks, FastAPI, Pydantic, Docker, Git
- **Cloud Platforms:** Google Cloud Platform, Azure

Certifications

Hugging Face Agents Course 08/2025
View
• Understood the "Think → Act → Observe" loop underpinning AI agents and how large language models serve as their core reasoning engine.

• Gained proficiency in building agents that call external tools, execute multi-step workflows, and integrate with web APIs using smolagents, LangGraph, and LlamaIndex.

NVIDIA Fundamentals of Deep Learning 12/2024
View
• Understood core mechanics of training neural networks (convolutional nets, recurrent nets).

•Learned to apply data augmentation, transfer learning using pre-trained models to achieve high accuracy with limited data.

• Gained practical experience building a final project using TensorFlow/Keras on GPU-accelerated infrastructure, cementing skills for production-ready deep learning.

Languages

Kannada: Native

English: C1

German: A2

Interests

Sleight of Hand

Swimming

Table Tennis

Mannheim, Germany