



ABHISHEK SINGH KUDAVALE *AI/ML Engineer*

✉ kudavaleabhisheksingh@gmail.com

☎ +49 15758515843

📍 Wolfsburg, Germany



LinkedIn



GitHub



portfolio

PROFILE

AI/ML Engineer with 2.5+ years of experience specializing in Computer Vision and LLMs. Focus on optimizing transformer architectures, fine-tuning AI models, and prompt engineering. Skilled in implementing containerized solutions (Docker) and deploying microservices on AWS. Committed to transforming complex algorithms into reliable, high-performance software solutions.

PROFESSIONAL EXPERIENCE

AI Research Intern & Master's Thesis

05/2025 – Present | Gifhorn, Germany

IAV GmbH

- Optimizing scenario extraction through a multi-layer fusion of computer vision (CV) and vision language models (VLM), currently tracking a 23% increase in accuracy compared to internal baselines.
- Collaborating with cross-functional teams to deploy a low-latency AWS microservice connected to CARLA and dSPACE simulators for automated scenario extraction and generation.
- Developing a SaaS-based benchmarking framework to evaluate AI models across diverse public and private datasets.
- Designing integrated reasoning systems combining VLM capabilities with CV outputs through prompt optimization and temporal context to enhance semantic scene understanding.
- Implementing and optimizing benchmarked CV models on real-world off-road data, achieving significant improvements in vision tasks.
- Engineering an AI framework integrating computer vision with vision language models for automated scenario generation, delivering a comprehensive solution for off-highway setups for the IAV Mela project.

AI/ML Research Assistant

10/2023 – 11/2024 | Siegen, Germany

Universität Siegen

- Enhanced 6D pose estimation by integrating deep learning models with dynamically tuned Kalman filters, achieving 22.88% performance improvement over baseline for the TRUST-E automotive project.
- Optimized parameters and algorithm configurations to maximize tracking accuracy on real-world IMU data.
- Developed visualization tools to analyze sensor noise profiles and trajectory deviations, enabling data-driven decisions for algorithm tuning.
- Implemented robust preprocessing techniques for IMU signals that significantly reduced drift in challenging conditions.
- Designed comprehensive evaluation workflows with reproducible benchmarks to quantify improvements from noise tuning and preprocessing.

Graduate Engineering Trainee

10/2020 – 11/2021 | Hyderabad, India

Technique Design Group

- Designed 3D models and assemblies in CATIA V5, ensuring manufacturability and performance optimization.
- Analyzed fabrication metrics and defect rates to identify root causes, driving design improvements for higher yield.
- Deployed ML models (XGBoost, SVMs) on Raspberry Pi, achieving 92% accuracy for real-time 3D CAD analysis.
- Managed fabrication workflows and quality control for client projects from concept to delivery.
- Collaborated with a 5-member cross-functional team to deliver projects on time and within scope.

EDUCATION

Masters of Science – Mechatronics

10/2022 – Expected 2026 | Siegen, Germany

Universität Siegen; CGPA: 1.9 (Best: 1.0)

- Focus: ADAS, Object-oriented Programming (C++ and Python), Machine Learning, Deep Learning, Embedded Systems, Computer Vision, Control Theory, Sensorics, Electric Machine, Project Management

Bachelor of Engineering – Mechanical Engineering

08/2016 – 05/2020 | Hyderabad, India

Vasavi College of Engineering; CGPA: 8.69 (Best: 10.0)

- Focus: Programming Languages (C and C++), Database Management, Fluid Mechanics, Machine Design, Thermodynamics, FEM Simulation, CAD/CAM, Automobile Engineering, Physics, Mathematics

PROJECTS

- Advanced RAG for Research Paper Question Answering**

10/2025 – Present

Proficiencies: RAG Pipeline, LangChain, Huggingface, Prompt Optimization, BERT, Transformers

 - Architected an advanced RAG pipeline incorporating reranking, memory integration, and DSPy-based prompt optimization for research paper queries.
 - Developed custom question-answer datasets and implemented context relevance and answer correctness evaluation metrics achieving 62% precision using the Gemini-2.5 flash model.
- Fine-Tuning Large Language Model for Spam Classification**

02/2025 – 04/2025

Proficiencies: LLM Pipeline, Token Embedding, Attention Mechanism, Transformer Architecture, GPT2

 - Implemented transformer based architecture with 12 attention heads for efficient context vector generation. Optimized token embedding and attention mechanisms for classification performance.
 - Fine-tuned pretrained GPT2 model using LoRA adapters on 10,000+ spam classification examples, achieving 97.21% training accuracy and 95.67% test accuracy.
- Robustness and Fairness in Deep Learning**

05/2024 – 07/2024

Proficiencies: CNN, Python, Adversarial Training, Computer Vision, Transfer Learning, Classification

 - Performed PGD training with 3 distinct robust loss rules, achieving 84% and 76% accuracy against targeted and untargeted attacks, enhancing security for vision-based systems.
 - Developed batch sampling techniques ensuring equal class representation, improving model fairness and reducing vulnerability to adversarial attacks in production AI systems.
- Image Classification Challenge on Custom Dataset**

11/2023 – 01/2024

Proficiencies: Deep Learning, Python, Pytorch, NumPy, Pandas, Data Pipeline, Data Augmentation

 - Trained EfficientNet-b4 on custom 8-class dataset, achieving 96.20% test accuracy through hyperparameter optimization, outperforming all comparison models.
 - Implemented advanced regularization strategies and scheduling techniques applicable to production image recognition systems with limited training data.

SKILLS

AI and Machine Learning: Artificial intelligence, Machine learning, Computer vision, LLMs, RAG, ChromaDB, AI agents, LangChain, LangGraph, FastAPI, Data science, Statistical analysis, Image processing

Software Engineering: Software development, Git, Docker, AWS, Linux

Programming & Simulation Skills: C++, Python, Matlab, Simulink

ML and Data Libraries: PyTorch, Tensorflow, Keras, NumPy, Pandas, Matplotlib, OpenCV, Scikit-learn

Tools and Environments: Visual Studio Code, Jupyter Notebook, PyCharm, Linux

LANGUAGES

English: C1, **German:** A2 (currently learning B1)

ADDITIONAL INTERESTS

Travelling, Hiking, Badminton, Formula-1 enthusiast