KEDAR GAIKWAD

kedar.gaikwad@asu.edu ◇ (602) 662-6935 ◇ LinkedIn ◇ Github ◇ kedarg.tech ◇ Tempe, AZ, US

EXPERIENCE

AI Software Engineer, *Stealth Startup* — Milpitas, CA   *Feb 2025 - Present*

- Architected high-volume web scraping pipeline extracting 5,000+ jobs daily, enriching metadata using LLM's and orchestrating pipelines with Apache Airflow, deployed as cron jobs to fetch latest job data
- Deployed serverless microservices on GCP Cloud Run to host user tailored document enhancement scripts, automating code deployment using reliable CI/CD processes reducing deployment times by 70%
- Developed an adaptive mock interview simulator using GCP Vertex AI leveraging RAG framework, integrating Google TTS Neural models with SSML for natural voice interactions, providing candidates with personalized interview preparation

Research Assistant, *Exsight.ai* — Phoenix, AZ   *Oct 2023 - Present*

- Researched neuro-symbolic techniques with Object Detection models, developing Explainable AI (XAI) solutions that increased recall in geospatial images by 30%, helping secure a STTR Phase 1 Space Force contract W. P. Carey News.
- Engineered robust **stress testing framework** utilizing **Meta SAM 2** for precise segmentation, enabling targeted adversarial patch and camouflage attacks that identified and addressed key vulnerabilities in mission-critical AI systems
- Optimized **XAI models** for **Nvidia Jetson** edge deployment enabling accurate real-time inference capabilities through **quantization, pruning, and layer fusion**

AI Researcher, *Raga AI* — Fremont, CA   *Jun 2024 - Aug 2024*

- Contributed to RagaAI Catalyst an observability platform that aids RAG system debugging by providing granular insights through timeline and execution graphs, facilitating to identify and resolve performance bottlenecks
- Collaborated on LLM evaluation framework to rigorously assess model quality across 50+ evaluation metrics and guardrails, establishing crucial quality benchmarks for production RAG applications and improving LLMs response accuracy
- Led end-to-end testing of RAG and Agentic workflows across multiple LLM providers (OpenAI, Vertex AI, Ollama, Groq), developing optimal prompting strategies and establishing performance benchmarks for production deployments

Senior Data Scientist *Raga AI* — Bangalore, IN   *Jan 2022 - Aug 2023*

- Led implementation of custom **autoencoder** network for drift tracking and outlier detection in **ADAS**, achieving 95% test accuracy which was installed on Ambarella CV22 **edge-AI** platform and featured at **2023 CES** in Las Vegas
- Collaborated cross-functionally to create the RagaAI platform for computer vision **drift detection** using CNNs and **anomaly detection**, directly contributing to secure $4.7 million in seed funding
- Transformed client's image processing workflow by implementing **CLIP-based semantic search**, reducing image tagging processing time by 70% while maintaining high accuracy and relevance
- Developed LLM testing platform with metrics (context precision, etc.) and implemented RAG model (VectorDB, GPT-4 API) for Raga's LLM framework

Deep Learning Engineer, *Uncanny Vision* — Bangalore, IN   *Jun 2019 - Jul 2021*

- Developed and enhanced edge-AI simulation environment by optimizing system-level and module-testing processes, resulting in a tenfold increase in test output speed and improved memory efficiency using multi-threading and .
- Adapted and integrated major neural network architectures (ResNet, YOLO, SR-GAN, VGG, Mask-RCNN, MobileNet, SqueezeNet) from frameworks like TensorFlow and PyTorch into a custom edge-AI platform, significantly boosting compatibility, reducing deployment complexity, and enhancing overall device performance.
- Authored training manuals, upskilled annotation team—improving image/video annotation speed by 20%; engineered a face mask recognition model (EfficientNet) with 97% accuracy tailored for low-resolution COVID scenarios, and developed boom barrier monitoring model achieving 95% accuracy for automated parking checkout.

EDUCATION

MS, Robotics Autonomous Systems (Artificial Intelligence) | Arizona State University   Aug 2023 - May 2025
Courses: Frontiers in GenAI, Operational Deep Learning, ML accelerator design

BE, Computer Engineering | University of Mumbai   Aug 2015 - May 2019
Courses: Machine Learning, Artificial Intelligence

SKILLS

Programming	Python, C/C++, Matlab, SQL, Flask, REST APIs, Git, Docker, Linux, GitHub Actions, Jira
ML	PyTorch, TensorFlow, ONNX, TensorRT, Scikit-Learn, Ollama, Langchain, OpenAI API, Diffusers
Data Science	Pandas, Numpy, Matplotlib, Seaborn, Tensorboard, SciPy, OpenCV, MLOps, GCP, AWS
Computer Vision	UNet, EfficientNet, ResNet, YOLO, VAE, GAN, ViT, ImageJ, Ambarella CV22, Nvidia Jetson