

Kedar Deepak Gaikwad

✉ kedar.gaikwad@asu.edu 📞 +16026626J35 in LinkedIn 🐙 Github

Current role	AI/ML Software Engineering Intern at Stealth Startup and Research Assistant at exsight.ai, ASU
Technologies	GCP, Python, Airflow, Azure, ChatGPT, AWS, Docker, PyTorch, Scikit-Learn

Work experience

AI/ML Software Engineering Intern, Stealth Startup 🏢 Feb 2 '25 - Present (1m)

GCP Python Airflow Azure ChatGPT

- As the first AI/ML developer, led the creation of core application functionalities.
- Developed a robust web scraping pipeline for data extraction, utilizing LLM APIs for metadata enhancement and automating database uploads. Deployed and orchestrated the scraper using Apache Airflow for continuous monitoring and scheduling.
- Designed and implemented AI-driven document enhancement endpoints, deploying them via GCP Cloud Run. Integrated with GitHub for streamlined CI/CD, ensuring rapid and reliable deployments.
- Created an audio podcast generator leveraging Gemini 1.5 Pro within a Retrieval-Augmented Generation framework. Integrated Google Text-to-Speech Neural and Studio voice models, employing Speech Synthesis Markup Language (SSML) for natural and expressive voice output.

Research Assistant, exsight.ai, ASU 🏢 Oct 2 '23 - Present (1y 4m)

AWS Docker Python PyTorch

- Achieved 8 % reduction in deep learning model size through pruning and quantization, ensuring performance while improving data processing efficiency, leading to a 3 % decrease in deployment costs
- Integrated neuro-symbolic methods with Explainable AI (XAI) to enhance neural network performance, resulting in a 3 % increase in object detection recall for geospatial imaging, securing an Air Force/Space Force STTR Phase I contract
- Created stress-testing framework for object detectors with targeted adversarial patch attacks, camouflage, and contrast modifications to simulate real-life scenarios and assess model robustness

AI Research Intern, RagaAI 🏢 Jun 2 '24 - Aug 2 '24 (2m)

Python PyTorch ChatGPT

- Developed RagaAI Catalyst, an observability tool for trace recording in RAG applications with one-click deployable solution for LLM models
- Collaborated on Raga LLM Hub development, implementing metrics to evaluate and establish guardrails for LLMs and RAG applications, resulting in a robust open-source framework with 100+ tests
- Benchmarked and optimized custom RAG pipelines for improved response quality and reduced token costs across Llama, Gemma, and Mistral models

Senior Data Scientist, RagaAI

- Collaborated on developing RagaAI Platform for computer vision drift detection through CNNs and anomaly detection, contributing to securing \$4.7 million in seed funding.
 - Conducted research on Out-of-Distribution (OOD) detection and AI stress testing across various sectors, facilitating company growth to engage with 8 organizations.
 - Designed and implemented an API pipeline with a dashboard for interactive visualization and clustering of DNN embeddings using t-SNE, UMAP, and PCA, enabling real-time analysis of high-dimensional data.
 - Implemented MMD and Kolmogorov-Smirnov tests for drift detection in image datasets, enhancing model stability and reducing undetected data drift.
 - Utilized AE, VAE, and VAE-GMM algorithms to improve anomaly detection accuracy by 4 % in identifying outliers within high-dimensional datasets.
-
- Adapted seven major neural network architectures from eight DL frameworks to a custom framework, reducing model sizes by 5 % and increasing operational speed for FPGA devices
 - Conducted continuous research in computer vision and provided consultation on advancing the custom DL framework
 - Developed and deployed an annotation tool that automated processes, increasing throughput by 4 % and enhancing team proficiency
 - Engineered a face mask recognition model using the EfficientNet series tailored for lower resolution images, achieving an exceptional 97% accuracy rate during COVID challenges
 - Trained boom barrier monitoring model with 95% accuracy to ensure smooth operations in automated parking checkout

Snippets

Application Enhancement Tool @ [github.com](#)

Designed an AI-powered tool to refine resumes and cover letters by analyzing job descriptions and extracting

key requirements. [AssignAI](#) [Clapnet](#) [Coopwe](#) [Ceyce](#) [Ase](#) [Bos](#) job applications [Donot](#) %x