Adithya S Kolavi

GitHub | LinkedIn | adithyask.com | Blog

Location: Bengaluru
Email: adithyaskolavi@gmail.com

TECHNICAL SKILLS

- **AI|ML|DL:** PyTorch, Transformers, PEFT, Bitsandbytes, Ray
- **Python Libraries:** FastAPI, Flask, OpenCV, BeautifulSoup, Selenium, Pandas, Poetry, Langchain, LlamaIndex
- **Web Frameworks:** React.js, Next.js, Express, Node.js, Vue.js, Bootstrap, Tailwind
- **Cloud:** Azure, Azure Machine Learning, AWS, AWS SageMaker, Docker, Kubernetes(AKS)
- **Big Data:** Databricks, Azure Data Factory, Apache Spark, Hadoop, Kafka
- **Databases:** MongoDB, PostgreSQL, Firebase, Redis, MySQL, Supabase, Pinecone, FAISS, Qdrant, ChromaDb, LanceDB
- **Languages:** HTML, CSS, JavaScript, TypeScript, Python, C/C++, SQL

EXPERIENCE

Microsoft Research — **Bengaluru**
Research Fellow — August 2025 to present

- Leading research on **real-time LLM applications**, developing **low-latency pipelines** optimized for streaming inputs, and advancing **long-context memory retention** with compression techniques to support live conversational scenarios.
- Designing and deploying **end-to-end LLM integrations** within the Microsoft 365 ecosystem including **Teams** and **Stream** to enable intelligent meeting summarization, context-aware recommendations, and retrieval-augmented interactions.

Apple — **Bengaluru**
ML Intern — January 2025 to present

- Contributing to the **Unified Intelligence Team**, specializing in **leveraging large language models (LLMs) for advanced knowledge disambiguation** and **constructing high-fidelity knowledge graphs**.
- Working on **information extraction** and **retrieval at scale** to enhance data integration and intelligence across Apple's ecosystem.
- Researching advanced model distillation techniques using self-supervised learning and reinforcement learning (RL)

CognitiveLab — **Bengaluru**
Founder , AI Researcher — January 2023 to present

- Leading an open-source AI research lab that received the **LLaMA Impact Grant** a six-figure, non-dilutive award by Meta for our **Nayana** initiative, building a multimodal vision-language model covering 22 languages and driving breakthroughs in OCR and cross-lingual reasoning Specializing in **synthetic data generation pipelines** for large-scale fine-tuning; created one of the **largest multilingual, multimodal document datasets** using synthetic data **2TB+**.
- Pioneered **India's first Kannada-English bilingual LLM — Ambari-7B**, achieving GPT-3-like performance for cross-lingual tasks and an 85% improvement in Kannada tokenization efficiency; also developed **Indic Eval / Leaderboard**, the first open-source benchmark suite for Indic LLMs with language-specific evaluation metrics.
- Developed open-source projects with ~10,000 GitHub stars, actively used by thousands of developers worldwide.

TurboML — **Bengaluru**
AI Developer — June 2024 to November 2024

- Architected and implemented a comprehensive **LLMOps observability platform**, enabling real-time monitoring and evaluation of large language models. Developed custom instrumentation for tracking inference latency, token usage, and model performance.
- Built an advanced **distributed tracing system** for AI workflows, reducing debugging time by 60% and enabling granular analysis of model behavior across multiple deployment environments.
- Pioneered automated evaluation frameworks integrating multiple metrics (ROUGE, BLEU, RAG Metrics), enabling continuous assessment of model updates. Reduced manual evaluation effort by 75% while maintaining robust quality standards.

RESEARCH

Indic Eval/Leaderboard (Open Source) — **spaCy ,NLTK ,Transformers, SkyPilot, Azure**

- Developed a first-of-its-kind evaluation framework for Indic Large Language Models, addressing the current English-centric bias in performance measurement across tasks like machine translation and cross-lingual question answering.
- The framework, compatible with various datasets, integrates TGI to speed up evaluation by 45%, offering metrics like accuracy, F1 score, perplexity, ROUGE, and BLEU. It also supports multi-cloud environments for easy execution.

Vision-Augmented Retrieval and Generation (VARAG) — **LLaVa, Visual RAG, LLama-index, Qdrant**

- An innovative system integrating textual and visual information, enhancing conventional Retrieval-Augmented Generation (RAG) byt 35% and improving contextual precision by 60%.
- Implemented comprehensive pipeline enabling VARAG to seamlessly handle text-heavy documents such as PDFs, textbooks, and research papers, leveraging Vision-Language models for efficient processing. Increasing RAG performance on documents by 35%

VLM for Unified visual understanding of Documents (ViViD) — **PyTorch, PEFT, Distributed Training, HPC**

- A unified Vision Language Model (VLM) that consolidates multiple document analysis tasks (layout detection, OCR, table extraction, math expression recognition, image captioning) into a single model(700 M parameter) .
- Improved document-to-markdown conversion accuracy, with automatic handling of complex elements like tables and equations.
- Fine-tuned VLM using novel multi-task strategies (full fine-tuning, LoRA), improving performance across diverse document structures.Reduced infrastructure costs by 20% through efficient memory utilization and multi-task capabilities.

PUBLICATIONS

- **A Multi-Agent Approach for Iterative Refinement in Visual Content Generation** *AAAI 2025 | Advancing LLM-Based Multi-Agent Collaboration.* ***Accepted***
- **CAPTAIN: Continuous Automated Planning Through Autonomous Internet Navigation** *AAAI 2025 | Large Language Models for Planning (LM4Plan).* ***Accepted***
- **Nayana OCR: A Scalable Framework for Document OCR in Low-Resource Languages** *NAACL 2025 | Language Models for Underserved Communities.* ***Accepted***
- **ViViD - Vision Language model for Unified Visual Understanding of Documents** *CVPR 2025 | Emergent Visual Abilities and Limits of Foundation Models (EVAL-FoMo 2025)* ***Accepted***
- **Nayana – A Unified Foundation Model for Multilingual, Multimodal, and Multitask Intelligence** *LlamaCon 2025 | Winner of 2024 LLaMA Impact Grant from Meta* ***Accepted***
- **Nayana: A Foundation for Document-Centric Vision-Language Models via Multi-Task, Multimodal, and Multilingual Data Synthesis** *ICCV 2025 |Workshop on Computer Vision for Developing Countries (CV4DC)| Proceedings Track* ***Accepted***

PROJECTS

OmniParse(6K stars on Github) **Python, FastAPI, Transformers, Redis Queue, Containerization, K8s**

- Open Source Data Ingestion and Processing platform, achieving 85% accuracy in structured data extraction from diverse data source like Documents , Media , Web for GenAI applications
- Integrated cutting-edge models for OCR, captioning, and transcription, improving data quality for AI applications by 90%
- Optimized for T4 GPUs deployment, reducing infrastructure costs by 40% compared to cloud-based alternatives.

GitVizz(Open Source) **Python, TypeScript, Langraph**

- an open-source tool for developers to intuitively explore codebases through **interactive dependency graphs** visualizing files, modules, and commit relationships in a clear, navigable structure.
- Designed a powerful **graph-centric UI** enabling fast search, filtering, and navigation making it effortless to understand complex repositories at a glance.

Cognitune - End-to-End LLMops Platform **Python, FastAPI, Huggingface, Transformers, LLMs,Containerization**

- All-in-one platform for LLMops, featuring distributed dataprocessing, multi-GPU fine-tuning, dynamic evaluation, and one-click high-throughput API deployment for **enterprises**. Achieved a **60% reduction in deployment time** for production-ready LLMs.
- Enables seamless swapping between fine-tuned LLMs and implements Retrieval Augmented Generation to reduce hallucinations.
- Cloud-agnostic design for versatile deployment across different cloud environments.

EXTRA CURRICULAR/AWARDS

Google Developers Student Club Lead **University Lead**

- Led the Google Developers Student Club, fostering a collaborative environment for technology enthusiasts at PES University.
- Organized events during Hacktoberfest, including hands-on Git and GitHub workshops, to promote open source contributions.
- Facilitated industry collaborations, guest lectures, and networking opportunities for club members.

Samarpana, Shunya **Technical Head**

- Directed technical teams for Samarpana (a fundraising marathon for families of martyrs) and Shunya (Math club) events.
- Led website development projects to enhance registration awareness and streamline participant engagement.

Hackathons **Participant**

- Won National Level Hackathons in Generative-AI - GenAI-Rush and Kodikon3
- Participated in more than **25 Hackathons** in the span of 2 years (adithyask.com) with a 90% finalist selection rate.

EDUCATION

- BTech, Computer Science, PES University
- 11th/12th, Physics | Chemistry | Mathematics | Electronics, VVS Sadar Patel
- Schooling, CBSE Syllabus, KLE Society's School