ARAV ADIKESH RAMAKRISHNAN

+1 (413) 210-6592 | aravadikesh@gmail.com | linkedin.com/in/aravadikesh | github.com/aravadikesh | aravadikesh.vercel.app/

EDUCATION

University of Massachusetts - Amherst **Sep 2024 - May 2026**
Master's, Computer Science *GPA: 4*

- Coursework : Machine Learning, Modern Neural Networks, Reinforcement Learning, Security & Privacy in GenAI, Advanced Natural Language Processing, Algorithms for Data Science

University of Massachusetts - Amherst **Aug 2021 - May 2024**
Bachelor's, Computer Science *GPA: 3.93*

- Coursework : Data Structures & Algorithms, Software Engineering (Typescript), Algorithms, Game Programming (C#), Operating Systems (C/C++), Database Management (SQL), Intelligent Visual Computing, Search Engines

SKILLS

- **Languages:** Python, TypeScript, JavaScript, Java, C#, C/C++, SQL
- **Frameworks & Tools :** Pytorch, NumPy, Pandas, Express.js, Flask, Git, AWS, Kubernetes, MySQL, Postgres, OpenCV, Unity, Apache Spark, Hadoop, Ollama, Unsloth

PROFESSIONAL EXPERIENCE

UMass BioNLP Lab **Amherst, MA, USA**
AI Translation Engineer *Jan 2025 - Present*

- Contributed to a Multi-Agent System Framework for translating clinical notes across multiple languages, enhancing global healthcare accessibility.
- Established comprehensive dataset processing pipelines and conducted rigorous evaluations of base models and fine-tuned variants including Llama 3.1B, Phi4 14B, and GPT4o-mini via Ollama and Unsloth.
- Implemented custom evaluation frameworks by modifying MedPrompt, APO, and MAPS methodologies. Achieved superior translation quality with the GPT4o-mini MedPrompt approach, demonstrating a BLEU score of 51.7 and ROUGE-2-F score of 0.573, outperforming all baseline models across multiple evaluation metrics.

Prime Focus Technologies **Los Angeles, CA, USA**
AI/ML Intern *May 2024 - Sep 2024*

- Led development of a production-ready conversational GenAI-driven image generation system with enterprise integration
- Engineered distributed architecture with JavaScript frontend, SpringBoot microservices on Kubernetes, and Flask/Python APIs on AWS Lambda, integrated with PostgreSQL/Redis achieving 100ms response times and 99.9% uptime
- Reduced image generation time by 40% and infrastructure costs by 35% through optimized processing, successfully demonstrating the solution at International Broadcasting Convention (IBC)
- Designed and deployed a scalable Retrieval-Augmented Generation (RAG) chatbot leveraging LangChain, FAISS vector store, and fine-tuned HuggingFace models (RoBERTa, DeBERTa) to handle 1K+ daily customer queries
- Achieved 92% customer satisfaction across 5 media clients, reducing annual support costs by $20K and improving ticket resolution times by 44%.

UMass Rescue Lab **Amherst, MA, USA**
Independent Researcher *Aug 2024 - Dec 2024*

- Architected a production-grade deepfake detection system using PyTorch with parallel CNN ensembles (EfficientNet-B7 and XceptionNet), BlazeFace/MTCNN for face detection, and custom data augmentation techniques, achieving 96% accuracy on DeepFake TIMIT (10% higher than baseline) and processing 100+ videos/hour
- Engineered a containerized MLOps pipeline with Flask RESTful APIs, RetinaFace for precise facial landmarks, and dlib for face alignment, reducing model deployment time by 65% and enabling automated batch processing of 1000+ videos/day for the RescueBox platform

UMass Amherst **Amherst, MA, USA**
Undergraduate Course Assistant (UCA) *Sep 2023 - May 2024*

- Instructed 100+ students in JavaScript, OpenGL, and foundational computer graphics concepts, through engaging lectures and hands-on projects.
- Conducted 5+ weekly office hours, assisting 50+ students with clarifying doubts and troubleshooting code, while providing supplementary resources to enhance learning outcomes.
- Graded 200+ assignments with precision, offering constructive feedback to support student growth and ensure academic excellence.

PROJECTS

YOLO Knowledge Distillation *Nov 2024 - Dec 2024*
ML Researcher

- Collaborated in a team of two to develop an optimized knowledge distillation framework for YOLOv8 using PyTorch, achieving 37.61% accuracy improvement on Oxford Pets while compressing model to 2.7M parameters
- Reduced model FLOPs by 95% (from 99.7B to 4.3B) and increased inference speed by 2.4x, enabling efficient deployment across CIFAR-10, Tiny-ImageNet, and Oxford Pets datasets.

UMass Outing Club Gear Locker Project *Sep 2023 - Dec 2023*
Lead Backend Developer

- Led the development of a scalable REST API using Express.js/TypeScript with Firebase Real-time Database, handling 100+ daily transactions and reducing data fetch latency by 60%
- Managed a cross-functional team of 3 developers, implementing Agile methodologies and CI/CD pipeline with GitHub Actions, resulting in 30% faster feature delivery and zero critical production bugs

Run! *Oct 2023 - Dec 2023*
Environment and Gameplay Engineer

- Engineered procedural terrain generation using Unity's Compute Shaders and custom C# algorithms (Perlin noise, Voronoi diagrams)
- Implemented performant AI pathfinding using multithreaded A* algorithm with Unity's NavMesh system, handling 100+ simultaneous AI agents while maintaining stable framerate