Baaz Jhaj

baazjhaj@gmail.com | linkedin.com/in/baaz-jhaj | + 1 (503) 847-4450 | https://github.com/bjhaj

Arizona State University - Ira A. Fulton Schools of Engineering | Barrett the Honors College **Tempe, AZ**
Bachelor of Science, Computer Science **2021 - 2025**
GPA: 3.75

PROFESSIONAL EXPERIENCE

DARPA TIAMAT - Transfer from Imprecise and Abstract Models to Autonomous Technologies **Tempe, AZ**
Research Assistant **August 2024 - Present**

- Actively contributing to TANR2L: Transfer of Autonomy using Neuro-symbolic Representation-based Reinforcement Learning project.
- Developing neurosymbolic representation-based reinforcement learning for autonomous robot navigation.
- Contributed to creating scene graphs and integrating large language models for decision-making in simulated and real-world environments, focusing on sim-to-real transfer.

REACH **Phoenix, AZ**
Software Engineering Intern **June 2024 - August 2024**

- Developed a tool to analyze historical seasonal data, helping trucking service centers predict high-demand equipment needs based on factors such as location, date, and weather.
- Built an OCR tool capable of reading handwritten documents from service centers, automating transcription tasks and reducing human effort and time spent on data entry.

PROJECTS

SmokeNet (Pytorch, LoRa, Edge AI, Hardware Acceleration, Model Optmization)

- Developed a wildfire detection system optimized for remote, no-service areas, achieving over 6% higher accuracy than a recent benchmark study (*Forests 2023*) despite using a smaller, resource-efficient model.
- Implemented knowledge distillation and quantization-aware training pipeline to deploy a high-performance model on low-power camera boxes.
- Designed a LoRa-based communication architecture enabling multiple camera nodes to relay fire alerts without traditional network coverage.
- Built a modular library to streamline future edge-device model deployment.

Translatica (React, Vercel, NextJS, SupaBase, AWS Lambda, Python, Pytorch, Git)

- Placed 2nd in the AI in Education Hackathon hosted by ScaleAI, receiving $5,000 in awards.
- Developed a modular speech-to-speech translation (S2ST) pipeline that preserves speaker identity, tone, and emotional prosody across 20+ languages.
- Evaluated trade-offs between direct and modular S2ST approaches, achieving a balance of accuracy, voice fidelity, and flexibility for future real-time deployment
- Integrated Whisper for robust ASR, GPT-based models for context-aware translation, and TTS for expressive voice cloning; optimized pipeline latency and prosody retention through modular component tuning.
- Designed a microservice architecture with Supabase (database, auth, storage) and AWS Lambda for scalable batch video translation; deployed frontend on Vercel with a React/Vite UI.

SelfAlign (LoRA/QLoRA, RLHF, Synthetic SFT)

- Building a fully customizable alignment pipeline for large language models, giving users control over model persona and ideological alignment.
- Implementing synthetic supervised fine-tuning (SFT) and self-RLHF pipelines, including dataset cleaning (WizardLM, Alpaca), synthetic data generation, and DPO/PPO-based preference optimization.
- Designed to produce lightweight, persona-controlled adapters for user-defined instruction-following models, with a planned visual interface for tracking “alignment drift” as models are trained.

PUBLICATIONS

“On the Impact of Pre-Training Datasets for Matching Dendritic Identifiers Using Residual Nets” **June 2024**
Geometric Media Lab - Arizona State University Tempe, AZ

- Applied machine learning to naturally random structures to create a highly accurate, secure identifier.
- Accepted for presentation at AI-SIPM 2024, with professors Pavan Turaga and Micheal Kozicki.

AWARDS

AI in Education Hackathon – ASU ScaleU | 2nd Place ($5,000) **October 2023**
Designed and built the backend of Translatica, a tool to translate lectures while maintaining context, tone, and voice.

FURI Scholar **January 2022**
Awarded $1500 in funding to research dendrite usage for item authentication in the supply chain with Dr. Michael Kozicki and Dr. Pavan Turaga

SKILLS

Programming Languages: Python (advanced ML and systems), C++, Java, Assembly

Machine Learning: PyTorch, TensorFlow, Scikit-Learn, Keras, Hugging Face Transformers; expertise in evaluating and benchmarking models, constructing datasets, and exploratory data analysis; Knowledge Distillation (KD), Quantization-Aware Training (QAT), LoRA/QLoRA fine-tuning, RLHF (DPO/PPO), Synthetic SFT pipelines; speech and audio generation and analysis (ASR, diarization, voice cloning, prosody-preserving TTS); convolutional and transformer-based architectures for vision and speech

Systems & Deployment: Microservice & cloud-native architectures (Supabase, AWS Lambda, Vercel), Edge AI deployment on Raspberry Pi & Jetson, LoRa-based multi-node communication systems, scalable video/audio batch processing pipelines

Research & Communication: Skilled at translating academic research into practical systems, strong academic reading and technical writing, experience presenting and publishing research

Soft Skills: Critical thinking, Leadership, Creative Problem-Solving, Time Management