Aktan Azat

aazat@ucdavis.edu | 310-405-1476 | Davis, CA
Seeking Summer 2026 Internship

Education

University of California, Davis Davis, CA
B.S. in Computer Science & Engineering; GPA: 3.93 *Graduating June 2026*

- **Honors & Activities:** Tau Beta Pi Engineering Honor Society, IEEE
- **Coursework:** Algorithm Design & Analysis, Operating Systems, Probability & Statistics, Computer Architecture, Software Engineering

Technical Skills

Languages: Python, Rust, Go, JavaScript/TypeScript, SQL, C/C++, LaTeX
ML/NLP: PyTorch, scikit-learn, Hugging Face Transformers, spaCy, PubMedBERT, LoRA/PEFT, Librosa
Infrastructure: AWS (ECS/Fargate), Azure ML/AKS, Docker, FastAPI, Celery, Redis, PostgreSQL, Palantir Foundry
Frontend & Tools: React, Next.js, Git, Prometheus/Grafana, OpenTelemetry, Pinecone, SMPL-X

Research Experience

Undergraduate Researcher, Complex Care Laboratory (UC Davis Health) May 2025 – Present
Advisor: Dr. Clodomir Santana *Sacramento, CA*

- Fine-tuned **spaCy v3** NER model with custom tok2vec embeddings on 4,200 clinical notes, achieving 91% F1 (vs 76% baseline); deployed pipeline processing 850+ notes/week, flagging 127 previously undetected high-risk conditions.
- Integrated **caseOLAP** framework with UMLS taxonomy via scispaCy for polarity scoring; identified 6 salient cardiovascular risk factors correlating with readmission ($p < 0.01$).
- Built **PubMedBERT** sentence classifier (focal loss, data augmentation) to distinguish true protein-disease associations from incidental mentions in HFpEF literature; achieved 69.3% accuracy / 0.656 macro F1 on held-out set (n=176), filtering 500 false-positive proteins (14%) from CaseOLAP rankings across 25,000 PubMed abstracts.

Undergraduate Researcher, Ben-Shalom Lab (UC Davis Health) Sep 2025 – Present
Advisor: Mandar Patil *Sacramento, CA*

- Built automated spike-sorting and signal analysis pipeline for 120+ neurons from high-density MEA recordings; statistical models explain 18–32% of neural threshold variance from spatial features, validated against immunostaining (89% concordance).
- Fit multi-compartment biophysical models predicting 15–40% threshold shifts from axon initial segment reorganization, consistent with observed plasticity after 72h protocols.
- Derived artifact amplitude model $A = K(V\tau)^{\beta}e^{-d/\lambda}$ and fit via log-linear least-squares on 300k+ channel-pulse observations (R^2 up to 0.51); built signal processing pipeline parsing Maxwell .raw.h5 recordings with 300–4500 Hz bandpass and automated artifact prediction.

Undergraduate Researcher, Motion Lab Sep 2025 – Present
Advisor: Dr. Michael Neff *Davis, CA*

- Built end-to-end pipeline testing spatial consistency of co-speech gesture under coreference across the BEAT2 motion capture corpus; extracted 2,400+ coreference-gesture pairs from **spaCy + neuralcoref** entity chains filtered by prosodic prominence.
- Implemented stroke detection from **SMPL-X** forward kinematics (elbow angle, wrist speed, forearm extension); phase segmentation achieves 83% agreement with manual annotations. Bounding-box and L1 divergence analysis shows coreferent gesture pairs have 28–41% higher spatial overlap than non-coreferent baselines.
- Built **Matplotlib**-based labeler GUI with synced audio/3D skeleton animation and visualization suite (hand-density heatmaps, 6-DOF animations, periphery plots) for qualitative validation and parameter tuning.

Undergraduate Researcher, Z Lab Jan 2026 – Present

Advisor: Dr. Jie Zheng *Davis, CA*

- Building multimodal analysis platform for objective autism assessment, replacing subjective ADOS scoring with data-driven behavioral quantification from speech, prosody (**Librosa**), and semantic features (**HuggingFace Transformers**).
- Training regression/classification models (**scikit-learn**, **PyTorch**) to predict severity scores from extracted behavioral features; integrated optional gaze fixation metrics for multimodal signal fusion.
- Building clinician-facing **React/FastAPI** application for session review, data visualization, and model inference with strict PII safeguards for protected health information.

Industry Experience

Member of Technical Staff July 2025 – Present

Ovavision – fertility health *New York City, NY (Remote)*

- Developing **PyTorch** temporal models for reproductive cycle forecasting using irregular time-series data; validation on clinical cohort shows 10–25% reduction in false-positive alerts vs rule-based system, improving predictive reliability ahead of beta launch.
- Architected async model serving with **FastAPI**, Celery+Redis, deployed on **AWS ECS/Fargate**; p95 inference latency at 210ms (down from 850ms) with Prometheus/Grafana monitoring and OpenTelemetry tracing.
- Leading UI/UX redesign of consumer-facing **React/Next.js** application; designing updated information architecture, component system, and interaction flows in collaboration with clinical advisors through iterative UAT cycles.

Data Engineering Intern June 2025 – September 2025

Rigetti Computing *Fremont, CA*

- Built anomaly detection pipeline validating 60M+ daily sensor records; automated QA replaced manual spot-checks covering <5% of data, catching 47 critical anomalies before fabrication impact.
- Designed and hand-soldered passive LC low-pass filters (4th-order Butterworth, 80 MHz cutoff) for cryogenic signal lines; bench-tested with VNA showing 38 dB stopband rejection at 250 MHz, reducing thermal noise floor by 6 dB across 12 qubit readout channels.
- Aggregated fab sensor, cryostat telemetry, and process control data into **Palantir Foundry** pipelines; built interactive dashboards tracking 15+ KPIs (chamber pressure drift, etch uniformity, deposition rates) used by 3 engineering teams for daily production decisions.
- Optimized time-series database with partitioning and composite indexes; 65% faster queries (2.1s→0.74s) enabled real-time monitoring dashboard for fab operations.

Software Engineer Intern January 2025 – May 2025

European Bank for Reconstruction and Development (EBRD) *London, United Kingdom*

- Fine-tuned LLMs with **LoRA/PEFT (Hugging Face)** on proprietary bank documents; achieved 18–35% accuracy gains on domain QA benchmarks (baseline GPT-4: 64%, fine-tuned: 82–86%) with content-policy filtering for sensitive data.
- Built NLP pipeline (**spaCy + transformers**) for topic and sentiment classification on 30k–70k articles/day; 86–93% precision on risk labels, flagging 340+ high-risk mentions missed by manual review.
- Deployed RAG system using **Azure OpenAI (GPT-4)** + Cognitive Search; 83% usefulness rating vs 61% for keyword search, cutting analyst research time by 2.3 hours/week.
- Implemented MLOps on **Azure ML** with gated evaluation and blue-green inference on AKS; automated model validation cut deployment cycle from 2–3 days to 3–6 hours.

Data Science Intern June 2024 – August 2024

Quantum Brains *London, United Kingdom (Remote)*

- Built RAG system with hybrid retrieval (**BM25 + sentence-transformers**, Pinecone vector DB, cross-encoder re-ranking); reduced query response time from 4.2min to 18sec across 800 daily interactions.
- Designed evaluation framework with few-shot prompts and SME reviews; guardrails for prompt-injection detection and PII redaction maintained hallucination rate below 7% on 2,400 production queries.
- Deployed inference service with **FastAPI** + Redis caching; monitoring pipeline enabled rapid rollbacks when retrieval quality degraded after corpus updates.