Aktan Azat

aktanazat@pm.me | 310-405-1476 | Davis, CA
linkedin.com/in/aktanazat | github.com/aktanazat | aktanazat.github.io

Education

University of California, Davis — Davis, CA
B.S. in Computer Science & Engineering; GPA: 3.9/4.0; Tau Beta Pi; IEEE — *Expected June 2026*

De Anza College — Cupertino, CA
Transferred to UC Davis — *Sep 2022 – June 2024*

Foothill College — Los Altos Hills, CA
Transferred to UC Davis — *June 2023 – June 2024*

Research Experience

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

- Fine-tuned **spaCy v3** NER with custom tok2vec embeddings on 4,200 clinical notes, improving F1 from 76% to 91%; deployed the pipeline to process 850+ notes/week and surface 127 previously unflagged high-risk conditions.
- Integrated **caseOLAP** with UMLS taxonomy via scispaCy for polarity scoring, identifying 6 salient cardiovascular risk factors associated with readmission ($p < 0.01$).
- Trained a **PubMedBERT** sentence classifier (focal loss, data augmentation) to separate true protein-disease associations from incidental mentions in HFpEF literature, achieving 0.656 macro F1 on a held-out set (n=176) and removing 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*

- Developed an 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 72-hour protocols.
- Derived an artifact amplitude model $A = K(V\tau)^{\beta}e^{-d/\lambda}$ and fit it via log-linear least-squares on 300k+ channel-pulse observations (R^2 up to 0.51); built a signal processing pipeline for Maxwell `.raw.h5` recordings with 300–4500 Hz bandpass filtering and automated artifact prediction.

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

- Designed an end-to-end pipeline to test spatial consistency of co-speech gesture under coreference across the BEAT2 motion capture corpus, extracting 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, and bounding-box plus L1 divergence analysis shows 28–41% higher spatial overlap for coreferent gesture pairs than for non-coreferent baselines.
- Created a **Matplotlib**-based labeler GUI with synced audio and 3D skeleton playback, plus a visualization suite including hand-density heatmaps, 6-DOF animations, and periphery plots for qualitative validation and parameter tuning.

Undergraduate Researcher, Z Lab — Jan 2026 – Present
Advisor: Dr. Jie Zheng — *Davis, CA*

- Developing a multimodal analysis platform for objective autism assessment, replacing subjective ADOS scoring with data-driven behavioral quantification from speech, prosody (**Librosa**), and semantic features (**Hugging Face Transformers**).
- Trained regression and classification models (**scikit-learn**, **PyTorch**) achieving 86% accuracy in ASD severity prediction from extracted behavioral features, with optional gaze-fixation metrics for multimodal fusion.

- Building a 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 on irregular time-series data; validation on a clinical cohort shows a 10–25% reduction in false-positive alerts versus a rule-based baseline.
- Architected asynchronous model serving with **FastAPI**, Celery, and Redis on **AWS ECS/Fargate**; reduced p95 inference latency from 850 ms to 210 ms with Prometheus/Grafana monitoring and OpenTelemetry tracing.
- Leading the UI/UX redesign of a consumer-facing **React/Next.js** application, defining updated information architecture, component systems, and interaction flows in collaboration with clinical advisors.

Data Engineering Intern June 2025 – September 2025
Rigetti Computing *Fremont, CA*

- Developed an anomaly detection pipeline validating 60M+ daily sensor records; automated QA replaced manual spot-checks covering under 5% of data and caught 47 critical anomalies before fabrication impact.
- Designed and hand-soldered passive LC low-pass filters (4th-order Butterworth, 80 MHz cutoff) for cryogenic control/baseband signal lines; VNA testing showed 38 dB attenuation at 250 MHz, contributing to a 6 dB reduction in measured downstream noise across 12 channels.
- Aggregated fab sensor, cryostat telemetry, and process-control data into **Palantir Foundry** pipelines; built dashboards tracking 15+ KPIs used by 3 engineering teams for daily production decisions.
- Optimized a time-series database with partitioning and composite indexes, improving query latency by 65% (2.1 s $\rightarrow$ 0.74 s) and enabling real-time monitoring 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 banking documents, improving domain QA accuracy by 18–35% over baseline GPT-4 while applying content-policy filtering for sensitive data.
- Designed an NLP pipeline (**spaCy + transformers**) for topic and sentiment classification over 30k–70k articles per day, reaching 86–93% precision on risk labels and surfacing 340+ high-risk mentions missed by manual review.
- Deployed a RAG system using **Azure OpenAI (GPT-4)** and Cognitive Search; internal evaluations showed 83% usefulness versus 61% for keyword search, reducing analyst research time by 2.3 hours per week.
- Implemented MLOps on **Azure ML** with gated evaluation and blue-green inference on AKS, cutting model deployment cycles from 2–3 days to 3–6 hours.

Data Science Intern June 2024 – August 2024
Quantum Brains *London, United Kingdom (Remote)*

- Implemented a RAG system with hybrid retrieval (**BM25 + sentence-transformers**), Pinecone vector search, and cross-encoder re-ranking, reducing response time from 4.2 minutes to 18 seconds across 800 daily interactions.
- Designed an evaluation framework with few-shot prompts and SME review; prompt-injection guardrails and PII redaction kept hallucination rates below 7% on 2,400 production queries.
- Deployed an inference service with **FastAPI** and Redis caching; monitoring enabled rapid rollback when retrieval quality degraded after corpus updates.