

Haoran "Tommy" Fang

832-331-3168 | tommyfang2004@gmail.com | LinkedIn | GitHub | tfangu.com

EDUCATION

The University of Texas at Austin

B.S. Statistics and Data Science | Minors: Computer Science, Business | GPA: 3.93

December 2026

Austin, TX

EXPERIENCE

Revantage (Blackstone)

AI Engineering Intern

June 2026 - August 2026

Dallas, TX

- Built and handed off a Python/FastAPI/React document-intelligence system for deployment in a secure non-production Azure environment, converting 400-800-page loan agreements into 22 source-linked fields in 1-2 minutes and cutting review time from 30-60 minutes to under 5.
- Implemented retrieval, reranking, source attribution, answer-key validation, and reusable JSON correction rules with FAISS and PostgreSQL, achieving 90%+ agreement; executives estimated 300+ hours of potential time savings.
- Built a React chatbot over a proprietary insurance API so engineering and transactions teams could query operational data in natural language, then completed a non-production handoff.

Glynac AI (Startup)

Initial Member of Technical Staff

August 2025 - January 2026

Remote

- Built a Flask REST inference service for 20+ investment intents and 2,000+ labeled examples, serving a 12-layer TensorFlow Conv1D classifier with 89% held-out accuracy.
- Implemented LangChain/Qdrant semantic retrieval over 100+ investment documents, returning source-grounded evidence through vector embeddings and cosine-similarity ranking.

Springer Capital

Data Science Intern

January 2025 - August 2025

Remote

- Engineered a Next.js/TypeScript workforce-intelligence application with Prisma and Azure PostgreSQL to centralize workforce analysis and surface sentiment-derived employee-risk signals.

JPMorgan Chase

Software Engineer Intern

June 2022 - August 2022

Houston, TX

- Built a Java/React banking dashboard backed by Cassandra, implementing editable account workflows and automated interest calculations for internal users.

AI EVALUATION

Mercor

AI Model Evaluator

September 2026 - Present

Remote

- Evaluate and refine model responses used for personality training, applying structured rubrics across tone, behavioral consistency, instruction adherence, naturalness, and conversational quality under NDA.

Handshake AI

LLM Evaluator

October 2025 - Present

Remote

- Designed adversarial source-extraction and response-generation tasks to elicit failures in advanced GPT/ChatGPT models, evaluating 100+ outputs against 10-criterion rubrics and documenting reasoning, grounding, and generation failures under NDA.

PROJECTS

Listing Finder

Cloudflare Workers, Hono, OpenAPI

- Co-developing a deployed real-estate search and alerting platform with public read APIs, API-key-protected writes, 15-minute scheduled evaluations, 72-hour freshness expiration, and auditable new-listing, price-change, and status-change rules.

Embedding Inference Service

Python, FastAPI, PyTorch

- Built and benchmarked a self-hosted batch embedding API with CPU/GPU execution, model caching, health checks, telemetry, and load tests, documenting cost-latency tradeoffs against hosted inference.

ECG Neural Network Classifier

Python, PyTorch, 1D CNN, Data Pipelines

- Built a reproducible pipeline over 100K+ ECG recordings to predict moderate-or-greater structural heart disease from waveform tensors, achieving 0.821 test AUROC, 0.791 AUPRC, and 0.018 expected calibration error.

SKILLS

Languages: Python, TypeScript/JavaScript, SQL, Java, C++

AI/ML: RAG, LLM and multimodal evaluation, agentic systems, PyTorch, TensorFlow, embeddings, reranking, FAISS, Qdrant

Engineering: FastAPI, Flask, Pydantic, REST APIs, React, Next.js, PostgreSQL, Azure, Docker, Terraform, Git, CI/CD