

DongYeop Lee

dongyeop0810@gmail.com | (619) 870-6702 | [LinkedIn](#) | [Github](#) | [Portfolio](#)

EDUCATION

University of Southern California, B.S. in Physics/Computer Science, GPA: 4.0/4.0 December 2028

- **Honors & Awards:** Dornsife Dean's List, Phi Kappa Phi Recognition Award, USACO Platinum Division
- **Member:** ShiftSC, ScopeSC, HackSC, Avenues Consulting Club

PROFESSIONAL EXPERIENCE

Hemut (YC P25)

Fresno, CA

AI Engineer

April 2026 – Present

- Extended an MCP copilot with per-aspect semantic retrieval and approval-gated multi-tool execution, and built a 333-question regression bank across 20 routing and context categories.
- Reduced required LLM generations by 50% for Voice AI Agents on eligible status-lookup turns through deterministic tool execution, and bypassed a confirmed-transfer planning stage with a recorded 310ms median time-to-first-speech.
- Reduced duplicate load-detail narration from 83% to 0% in replays by strengthening tool-result guards, and hardened a 3-round negotiation state machine with load-scoped agreement locks.

Sense AI (a16z)

Los Angeles, CA

Founding Software Engineer

September 2026 – Present

- Built and secured a 15-tool MCP server for a 25-team user-testing platform for coding agents, backed by 400+ tests: SHA-256-hashed bearer tokens, account scope taken only from the token, and foreign test IDs rejected.
- Shipped a real-time AI voice interviewer (streaming STT, low-latency LLM inference, prompt caching, 3-provider TTS failover) that placed 291 calls and ran 129 interviews (16+ hours) for 21 teams for a Hackathon during SF Tech Week.

KrootMe

Los Angeles, CA

Founding Software Engineer

September 2026 – Present

- Hardened PostgreSQL authorization with function permissions and booking-state guards, reducing successful unauthorized operations from 7/10 to 0/10 in synthetic regression tests while preserving legitimate workflows.
- Built a deterministic job-enrichment pipeline syncing Greenhouse, Lever, and Ashby hourly for 1,000+ active users, extracting compensation and sponsorship with source evidence, ambiguous-match rejection, and failure-aware retries.

IDX Exchange

Los Angeles, CA

Natural Language Processing Engineer

February 2026 – May 2026

- Implemented hybrid property search using MiniLM embeddings, FAISS cosine-similarity retrieval, and BM25 keyword ranking, combining normalized scores to return ranked listings with structured property signals.
- Raised entity-extraction micro-F1 from 78.2% to 88.1% and reduced false-positive price extractions 75% on a fixed synthetic benchmark through spelled-number handling and contextual filtering.

PROJECTS

[Mend](#) | *Swift, Supabase, PostGIS, PyTorch, Core ML*

Finalist, F26 Origin Hackathon

- Trained an on-device damage classifier (EfficientNet-B0, Core ML, 8 MB), chosen over ConvNeXt-tiny at 0.949 vs 0.954 mAP for 5x fewer parameters; caught train/test leakage across drone video frames and re-split by frame block.
- Built the server verdict on a vision LLM with schema-enforced JSON (type, severity 1-5, confidence), fencing reporter notes as untrusted input and routing results under 0.6 confidence to human review.

[EcoPrompt: AI Energy Efficiency](#) | *Next.js, React, TypeScript, AWS Bedrock, DynamoDB, Tailwind CSS*

- Developed EcoPrompt, a Next.js and TypeScript AI chat app on AWS Bedrock (Titan embeddings) and DynamoDB, featuring semantic-deduplication caching, prompt-complexity-based model right-sizing, and a real-time energy dashboard.
- Cut per-query LLM energy consumption by up to 98.6% on cache hits and 90% on right-sized queries, quantified via EPA eGRID factors.

TECHNICAL SKILLS

- **Languages:** Python, Java, C++, TypeScript, JavaScript, SQL, Bash, HTML, CSS
- **AI & ML:** PyTorch, Anthropic API, MCP, RAG, FAISS, Embeddings, LLM Evals, Structured Outputs, Voice AI
- **Backend & APIs:** FastAPI, LiveKit, Twilio, REST APIs, Node.js/Express, PostgreSQL, Supabase, Redis, DynamoDB (NoSQL), SQLAlchemy, Pydantic, pandas
- **Cloud & DevOps:** AWS (S3, ECS/Fargate, ECR, RDS, DynamoDB, Secrets Manager, Bedrock), AWS CDK, Docker, Kubernetes, Linux, Terraform, GitHub Actions, CI/CD, Git, Claude Code, Codex, GitHub Copilot
- **Frontend & Testing:** React, Next.js, Tailwind CSS, Zustand, Zod, pytest, Jest, Vitest