

NANTHA KUMAR

AI / Machine Learning Engineer

Irving, TX (Open to Relocation) | •Phone: 817-819-7132 | • Email: reachnanduu24@gmail.com
• LinkedIn: [linkedin.com/in/nanthakumarashokanand](https://www.linkedin.com/in/nanthakumarashokanand) | • Github: github.com/Nanduu24

SUMMARY

AI and Machine Learning Engineer with a master’s in computer science from UT Arlington who ships LLM systems into production. Built and maintains a live agentic AI platform on LangGraph with pgvector memory, voice and multi-provider model routing, plus a document question-answering agent that cites its sources and abstains without evidence. Published ML researcher (IEEE I2CT, 2024).

SKILLS & INTERESTS

Languages: Python, TypeScript, JavaScript, SQL, Java, C++, C | **Frameworks:** Next.js, React, FastAPI, Node.js, Express, REST, GraphQL
AI and LLM Engineering: LangGraph, LangChain, Retrieval-Augmented Generation (hybrid and vector), Prompt Engineering, Function Calling, Structured Outputs, Multi-Provider LLM Routing, Grounding and Hallucination Guardrails, Speech-to-Text and Text-to-Speech, LangSmith, Anthropic Claude, OpenAI, Google Gemini, Groq, Llama 3.3
Machine Learning: scikit-learn, XGBoost, Random Forest, TensorFlow, Keras, PyTorch, CNN, LSTM, Optuna, SHAP, Cross-Validation, Hyperparameter Tuning, Model Interpretability, NumPy, pandas
Data and Cloud: PostgreSQL, pgvector (HNSW), Supabase, Vector Embeddings, Docker, Vercel, Hugging Face Spaces, Git, GitHub Actions, Stripe, Clerk
Testing and AI Tooling: Jest, Pytest, Selenium, Claude Code, Cursor, GitHub Copilot
Interests: Agentic AI systems, retrieval-augmented generation, LLM evaluation and grounding, applied ML research

EXPERIENCE

AI / Machine Learning Research Assistant Aug 2025 - Present
iResearchE3 Lab, University of Texas at Arlington | Arlington, TX

- Built the lab's end-to-end machine learning pipeline in Python (scikit-learn, XGBoost, TensorFlow, Keras) to predict rock permeability from NMR T2 measurements across 720 carbonate core samples from two Kansas wells, handling targets spanning six orders of magnitude with log-space modelling and CDF-based feature encoding.
- Tuned XGBoost, Random Forest and deep neural network regressors with nested Optuna search and 5-fold cross-validation, reaching a test R-squared near 0.60 against the lab's SDR physics baseline, then ran 50-iteration resampling across seven training sizes (120 to 720 samples) to map accuracy scales with data volume.
- Used SHAP and Spearman rank correlation to show 4 of 11 features drive roughly 95% of output; co-authoring the manuscript.

PROJECTS

Mindful AI | Agentic AI Platform with Voice and Memory | Next.js, FastAPI, LangGraph, pgvector, Claude, Gemini, Stripe

- Shipped a six-stage LangGraph agent to production with provider-agnostic routing across Claude, Gemini 2.0 Flash and Llama 3.3 70B, switchable through one environment variable, holding cost near \$0.0002 per session.
- Built RAG memory on Postgres with pgvector and HNSW indexing (768-dimension embeddings, custom cosine-similarity SQL) returning results in under 200 ms and added browser push-to-talk voice via Gemini speech-to-text and Eleven Labs.
- Deployed across Vercel, Hugging Face Spaces and Supabase with LangSmith tracing, Clerk auth and Stripe billing.

FleetIQ | Grounded Document Q&A Agent, Buildathon Dallas 2026 (32-hour build) | LangGraph, FastAPI, Hybrid RAG

- Built a hybrid retrieval agent over 1,014 fleet documents spanning 12 real carrier form types that answer with citations and stays silent when the evidence is missing.
- Generated the corpus from 12 templates with realistic entity variation to create known ground truth for measuring retrieval quality; delivered backend, dispatch interface and pitch inside 32 hours.

Job Agent | Autonomous Job Application Pipeline | Python, LangGraph, Greenhouse / Lever / Ashby / Workable APIs

- Built a pipeline that ingests postings from four job boards, scores them with a heuristic filter then an LLM fit review, tailors resume through a LangGraph workflow and render PDFs; ran end to end against live postings.
- Added a fact-verification node that checks every generated claim against a structured profile and blocks invented skills, titles or dates before the document is rendered.

AI Career Job Advisor | LLM-Powered Career SaaS, 3-person Agile team | OpenAI, FastAPI, Next.js, PostgreSQL

- Owned resume-to-job-description matching, ATS scoring and cover-letter generation with schema-validated outputs; wrote 25+ backend tests and contributed to a Selenium end-to-end suite.

EDUCATION

Master of Science in Computer Science, AI/ML Specialization Aug 2024 - May 2026
University of Texas at Arlington | Arlington, TX | GPA: 3.60 / 4.00

Bachelor of Technology in Information Technology Aug 2020 - May 2024
Rajalakshmi Engineering College | Chennai, India | CGPA: 8.28 / 10.00 (First Class)

PUBLICATION

Sign Language Caption Generation Using LSTM, IEEE International Conference for Convergence in Technology (I2CT), 2024. Real-time sign-language captioning using MediaPipe Holistic keypoints and LSTM action-detection model. ieeexplore.ieee.org/document/10543890