

SUJENDRA JAYANT GHARAT

Boston, MA | (857) 930-1933 | gharat.su@northeastern.edu | linkedin.com/in/sujendra-gharat/ | github.com/suju297

TECHNICAL SKILLS

Languages: Python, JavaScript, TypeScript, Bash

Backend & Frameworks: FastAPI, Flask, Node.js, gRPC, asyncio

ML & AI: PyTorch, TensorFlow, scikit-learn, Transformers (Hugging Face), LangChain, Autogen

GenAI & LLMs: Prompt Engineering, RAG Pipelines, Agentic Workflows, Pinecone, OpenAI SDK (GPT-4o, GPT-5), Llama 3

Cloud & DevOps: AWS, Docker, Kubernetes, Git, GitHub Actions

Databases: MySQL, PostgreSQL, SQLite, MongoDB, Redis

EXPERIENCE

Software Engineer Intern (Co-op)	Jan 2025 – Aug 2025
Xellar Biosystems	Boston, MA
• Architected and optimized a FastAPI microservice with Redis -backed caching to convert and persist 16-bit images on first load, enabling a median response time of < 0.4 s for repeat requests	
• Automated high-throughput S3 ingestion using boto3 with concurrent upload orchestration, retry logic, and conflict resolution policy (overwrite, rename, skip) for terabyte-scale datasets	
• Standardized raw microscopy data into a canonical S3 structure and auto-generated a per-file manifest used by the FastAPI service for Redis lookups and reliable retrieval	
Graduate Research Assistant – AI-CARING	Feb 2024 – Aug 2025
Northeastern University – Khouri College of Computer Sciences	Boston, MA
• Deployed a multi-stage pipeline that extracts reminder intent from caregiver conversations and compiles it into AST-validated Python functions	
• Designed a feasibility checker using the OpenAI o4 mini reasoning model that validates reminders against available sensors and a home layout graph before generation, preventing non-feasible reminders	
• Orchestrated a low-latency AWS IoT Core to gRPC middleware streaming data from over 120 sensors (about 3,000 events per second) and achieved a 50% reduction in end-to-end latency	
Senior Software Engineer	Feb 2022 – Aug 2023
Capgemini	Mumbai, India
• Engineered a Python/Flask inference gateway for multimodal AI workloads, integrating model endpoints and optimizing request handling for 40% lower response latency	
• Scaled inference workloads on Kubernetes with autoscaling and monitoring to ensure 99.9% uptime	
Software Engineer	Aug 2018 – Feb 2022
LTIMindtree	Mumbai, India
• Integrated 30+ REST/SOA APIs in Node.js with async patterns and MongoDB caching , improving performance by 15% and cutting API calls by 30%	
• Implemented graceful shutdowns and monitoring to reduce recovery time by 50% and maintain 99.9% uptime	

ACADEMIC PROJECTS

Time-Series Forecasting with Market Indicators (TSLA)	Sep 2024 – Dec 2024
• Built a macro-finance dataset and developed a time-series ML pipeline using OLS, ElasticNet, Random Forest, and XGBoost, achieving R2 = 0.88 and 50% lower RMSE vs baseline models	
• Applied PCA and K-Means/GMM clustering with a backtesting system to evaluate trading signals (945 trades, 57% win rate) and validate model stability across regimes	
Quant Stack Exchange Chat Assistant	Sep 2024 – Oct 2024
• Developed FinRobot, a generative AI chat assistant for quantitative finance using AutoGen with RAG , delivering context-aware answers with approximately 85% response accuracy	
• Engineered a scalable knowledge layer with a daily Python scraping pipeline that ingests Quant Stack Exchange into Neo4j (knowledge graph) and Pinecone (vector index) to enable fast, entity-aware retrieval	

SUJENDRA JAYANT GHARAT

Boston, MA | (857) 930-1933 | gharat.su@northeastern.edu | linkedin.com/in/sujendra-gharat_ | github.com/suju297

Cloud Native Web App

Jan 2024 – Apr

2024

- Provisioned and deployed a cloud-native web app using pre-configured machine images with **Packer** and **Terraform**, achieving a **75% faster** setup and integrating a serverless user verification system with **Cloud Functions** and **Cloud SQL**

EDUCATION

Northeastern University , Boston, MA	Expected	Dec
2025		
Master of Science in Information Systems	GPA :	
3.72/4.0		May
University of Mumbai , Mumbai, India		
2018		
Bachelor of Engineering, Electronics Engineering		