

# ADHISH THITE

[jobs.adhish@gmail.com](mailto:jobs.adhish@gmail.com) | Pune, India | [linkedin.com/in/adhish-thite/](https://www.linkedin.com/in/adhish-thite/) | <https://adhishthite.com/>

Generative & Agentic AI | Multi-Agents Systems | Cloud-Native AI Infrastructure | Applied AI Engineering | Systems Architecture

Lead Generative AI Engineer with over ten years of experience designing intelligent systems that combine applied machine learning, modern cloud architecture, and thoughtful user experience. Skilled at turning complex ideas into scalable, reliable products that improve efficiency, security, and decision-making. Known for a mix of technical depth, product sense, and leadership that helps teams move fast, align on goals, and deliver meaningful impact.

## Skills:

Core AI & ML	LLMs, RAG Systems, Agentic AI Frameworks, LLM Fine-Tuning, Prompt Engineering, Context Engineering, MCP
System Architecture & DevOps	Microservices, API Design, Distributed Caching, Async Processing, CI/CD (GitHub Actions, Jenkins), Monitoring (Elastic, DataDog)
MLOps & Infra	MLFlow, AWS SageMaker, Vertex AI, Azure, Docker, Kubernetes, Cloudflare Workers, Durable Objects, R2, KV, Vercel Stack
Programming & Data	Python, JavaScript/TypeScript (Next.js, React, Node.js), Rust, Go, SQL & NoSQL DBs (PostgreSQL, MongoDB, Redis, Elasticsearch, Convex)
Innovation Areas	Agent Swarms, Multi-Component AI Pipelines, Synthetic Data Generation, Edge AI.

## PROFESSIONAL EXPERIENCE

Lead Engineer (AI / GenAI) | Elastic Co. (San Francisco, CA) [Remote from Pune, India] 2/24 – Present

- “ElasticGPT” Deployment: Architected multi-agent system combining RAG, tool-use, and custom LLM fine-tuning for internal knowledge retrieval, automation, and decision support. Deployed across the enterprise; handles 50K+ queries/month with <1s TTFT latency. [1][2][3]
- Bespoke AI Deployment: Designed and implemented fine-tuned LLMs & RAG pipelines, deploying custom AI solutions to multiple business units with secure data interaction protocols, leading to a 60% improvement in data security and operational workflows.
- Business Process Optimization: Developed GenAI-driven internal tools that streamlined business processes, resulting in a 25% reduction in sales cycle time and a 40% increase in customer support efficiency. [1]

Senior ML Engineer | Synapse (San Francisco, CA) [Remote from Pune, India since 2023] 12/19 – 2/24

- Expedited KYC Process: Developed an AI-driven ID verification system, reducing onboarding time from hours to seconds and streamlining the authentication of users' government-issued documents.
- Innovative System Design: Created a “Video Auth” solution using speech-to-text and facial recognition, validating users in under 20 seconds by comparing videos with their respective government-issued IDs.
- Advanced Automation: Designed an NLP and ML-based automation system to optimize due diligence procedures, assessing **platform eligibility in just 15 seconds** through web scraping, sentiment analysis, and industry categorization.

- **Enhanced Security:** Developed user-friendly products like Remote Deposit Capture, real-time SSN Verification, and Sanctions Monitoring, **increasing compliance effectiveness by 2x** by mitigating risks of fake accounts and fraudulent transactions.
- **System Scalability:** Rearchitected systems to seamlessly manage over a **million requests per day** by moving from **on-premises to AWS**, implementing efficient caching and async processes, developing robust APIs, and optimizing database implementation.
- **GenAI Applications:** Developed modules on GPT-3.5 and GPT-4 LLMs, achieving a **40% improvement in customer support** and **minimizing dataset labeling** duration from **1 day to 5 minutes** for every 1000 images.

**Machine Learning Engineer | Gemini Data Inc.** (San Francisco, CA)

4/19 – 12/19

- **Pattern Extraction:** Enhanced log analytics performance by **80%** by implementing a sophisticated **Python** and **Apache Calcite** pipeline to **generate regular expressions** (regex) from unstructured log data to parse & store real-time streams automatically.

**Machine Learning Engineer (Intern) | Welch Labs** (Charlotte, NC)

10/18 – 4/19

- **Modular Algorithm Design:** Partnered with **Microsoft & Mariner** to identify defective windshields with **96%** accuracy by implementing an ensemble of Convolutional SVMs and CNNs using **OpenCV** and **Keras** in Python.
- **Increased Savings:** Projected to save **USD 4 million per year**. Optimized the model to have an **8s inference time on edge** after deploying as an **Azure Machine Learning Web Service**.

**Machine Learning Engineer (Intern) | Zuora, Inc.** (San Francisco, CA)

6/18 – 8/18

- **Business Process Improvement:** Reduced Zuora's live support agent involvement by **75%** by building an NLP Topic Modelling pipeline in **Java** and **Python** to correlate customer support tickets with internal knowledge base content.
- **Driving Customer Engagement:** Accelerated ticket response time by **90%** by leveraging Deep Learning algorithms to automate access validation to Salesforce.com orgs in Zendesk.

**Application Development Analyst | Accenture** (Pune, India)

4/15 – 7/17

- **Application Development & Maintenance:** Led the end-to-end delivery of an e-commerce platform for Splunk. Acted as the Lead Salesforce.com Developer and Technical Team Lead for Salesforce's **first-ever implementation of a cloud-on-cloud model**.
- **System Overhaul:** Facilitated complete overhaul of a Purchase Order flow by developing key delivery components. Augmented Sales Reps' efficiency by **50%** by deploying advanced automation processes via Salesforce.com customization & configuration.

---

## EDUCATION

**Master of Science (Computer Science) | University of North Carolina at Charlotte, USA**

May 2019

**Bachelor of Engineering (Computer Engineering) | University of Pune, India**

May 2014