Sanghita Chakraborty

(929)-855-6830 | sanghita.chakraborty@nyu.edu | LinkedIn | GitHub | Address - San Jose, California

PROFESSIONAL SUMMARY

Software Engineer with 3+ years of experience building scalable backend systems and production-grade ML-powered services. Skilled at owning problems from concept to production — driving design decisions, defining service boundaries, and building reliable, high-performance distributed systems.

TECHNICAL SKILLS

- **Language Skills:** Python, C#, C++, SQL
- **Software Engineering Skills:** Kubernetes, Docker, PostgreSQL, CI-CD Pipelines, Kafka, Terraform, Prometheus, AWS, Azure Services, Google Cloud Services, Redis, Rest API, Object Oriented Programming, Database Systems, Operating System, DevOps, Distributed System, Git, Linux
- **AI System Skills:** Retrieval-Augmented Generation (RAG), LLM Integration, LangGraph

EXPERIENCE

Intuit   June 2025 — present
Software Engineering(AI/ML) Intern   *Mountain View, CA*
TurboTax Business Knowledge Retrieval System
- Developed an AI-powered Slackbot for 1000+ employees in Intuit's TurboTax Organization, for querying business KPI information like product performance insights, user churn rates, and SKU conversion rates.
- Enabled real-time concurrent user query processing across 10+ Slack channels with Slack's Asynchronous SocketMode SDK.
- Reduced downtime of the query processing service by 95% with Kubernetes health checks and auto-restart mechanism.
- Engineered an OpenSearch RAG-based library to decouple knowledge retrieval from Slack query handling, with response latency <2s.
- Minimized the occurrence of stale data in the results by 90% through a production pipeline that preprocesses and ingests multimodal unstructured data present in 30GB+ presentation files with graceful error handling.

Automated ML Insights Service
- Built an agentic backend API integrated with the TurboTax Data Analytics dashboard to automatically summarize ML interpretability plots, reducing manual error; This was adopted by 8 Data Science teams across TurboTax.
- Designed RAG and MCP(Model Context Protocol) retrieval pipelines to generate summaries from ML experiment metadata, tax season trends and TurboTax metrics (example SKU conversion rates, upsell trends).
- Reduced end-to-end request latency from 30s to 2s by caching business context using Redis.

Microsoft   June 2022 — June 2024
Software Engineer   *India*
- Enhanced accessibility of Microsoft Visio Application for visually impaired customers (8% of the userbase) with speech navigation, customized contrast, and smooth zoomed navigation using React.
- Improved scalability and decreased overall system downtime by 85% for Microsoft Employee Performance Platform by refactoring the MVC-based backend to a containerized Microservice architecture.
- Engineered a data quality monitoring system to validate employee payroll data migration from SAP-based to internal system, reducing data discrepancy tickets from employees by 28%, implemented using Spark and Scala.

Google   May 2020 — July 2020
Software Engineering Intern   *India*
- Developed a functional testing algorithm for memory read/write transactions between Google Memory Controller and RAM in devices like Pixel and Chromebook with 78% coverage.

PROJECTS

Adaptive Flatmate Recommendation Platform
Associated with NYU Courant
- Built and deployed a two-sided flatmate recommendation system on Google Cloud, scaling from 40 real users to 5K+ synthetic profiles to tackle cold-start and data sparsity in early matching.
- Implemented a production ML pipeline with Vertex AI and MLflow, running 20+ experiments across 4 model classes, improving precision@10 by 18% and reducing recommendation drift impact by 25% via automated retraining.

Healthcare Insurance Decisioning Platform
Associated with NYU Courant
- Built a backend platform that converts unstructured clinical notes into structured, queryable health features, enabling automated, explainable insurance risk assessment and underwriting workflows.
- Designed a workflow-driven OLTP/ODS system with cloud storage, scheduled batch processing, and ORM-backed relational persistence, ensuring reproducibility, auditability, and efficient cohort and risk queries.

EDUCATION

NYU Courant   Sep 2024 - March 2026
Master of Science - Computer Science
BITS Pilani   Aug 2018 - July 2022
Bachelors of Engineering - Electronics and Instrumentation