

# YASH JAIN

Brooklyn, NY, 11209 | +1 646 246 5651 | [yashjain778@gmail.com](mailto:yashjain778@gmail.com) | [linkedin.com/in/yashjain0112](https://www.linkedin.com/in/yashjain0112) | [github.com/yj3335](https://github.com/yj3335)

## EDUCATION

### New York University

Sep 2025 – Present

Master of Science, Computer Science

- **GPA:** 3.9/4 | **Achievements:** Merit Based Scholarship | **Coursework:** Operating Systems, Algorithms, Machine Learning, Software Engineering, Cloud Computing, Big Data

### Manipal University Jaipur

Jul 2019 – Jul 2023

Bachelor of Technology, Computer Science and Engineering, Minor Specialization - Cloud Computing

- **GPA:** 9.07/10 | **Achievements:** Student Excellence Award, 3rd Best Minor Project (Research), 1st Position SudoHacks | **Coursework:** DBMS, Object Oriented Programming, Computer Networks, Data Science

## WORK EXPERIENCE

### NYU Department of Computer Science and Engineering

Sep 2025 – Present

Graduate Course Assistant

Brooklyn

- Conducted weekly recitations for 80+ students in CS-UY 3224 Intro to Operating Systems, covering process scheduling, memory management, and synchronization; contributed to a 15% increase in average assignment scores.
- Graded assignments and exams achieving 99% accuracy with actionable feedback; mentored students through office hours on labs and coursework.

### Zscaler

Aug 2023 – Jun 2025

Software Development Engineer

Bangalore

- Developed the ZIA Nanolog cloud-native microservice using Go, Kubernetes, and REST APIs to enable scalable API-driven query execution, supporting 300+ parallel queries and reducing average query latency by 20%.
- Engineered the Zscaler Workplace Insights data backend using Hadoop, Spark, Airflow, Python, and SQL, generating \$4M ARR and reducing workflow runtime by 75%; secured early customer conversions via Google Sheets POC.
- Constructed a Zagent metrics dashboard monitoring 14K+ nodes across resilient distributed systems using Python, Elasticsearch, Fluentd, and Kibana, cutting debugging time by 30% and enhancing system-wide observability.

### Zscaler

Feb 2023 – Aug 2023

Software Development Engineer Intern

Bangalore

- Optimized large-scale data pipelines processing 1B+ daily transactions using Hadoop, Python, and Shell scripting, accelerating data availability by 40% for Zscaler's revenue-generating Workplace Insights BI product.
- Restructured database schemas across distributed systems to accommodate evolving analytics requirements, reducing query execution time by 25% for the BI team.
- Architected a Kubernetes-based EFK monitoring pipeline, reducing error rates by 18% across all applications and enabling early incident detection for a 50+ member engineering team.

### Programming Pathshala

Mar 2022 – Jun 2022

Software Development Engineer Intern

Remote

- Shipped React Native and Node.js features for 5,000+ students, increasing engagement by 30%; automated a GCP storage pipeline eliminating 10+ hours of weekly manual work.
- Partnered with cross-functional teams to troubleshoot and resolve production issues, reducing page load times by 35% and improving platform reliability.

## SKILLS

**Programming Languages:** Python, Go, C++, C, JavaScript, TypeScript, SQL

**Web & Mobile:** ReactJS, Django, NodeJS, ExpressJS, React-Native, Redux Toolkit, MongoDB, PostgreSQL, Redis

**Cloud & Data:** AWS (ECS, Elastic Beanstalk, CDK), GCP (Cloud Run, Vertex AI, BigQuery), Kubernetes, Docker, Kafka, PySpark, Hadoop, Airflow, Cassandra, DuckDB, Elasticsearch, Fluentd, Kibana

**AI/ML & DevOps:** RAG Pipelines, LLM Agents (Ollama, Gemini, Google ADK), Prompt Engineering, AI Coding (Claude Code, Copilot, Cursor), CI/CD (Travis CI), Git, Bash

## PROJECTS

**CodeCollab – Real-Time Collaborative Code Editor** | TypeScript, Yjs, Redis, AWS ECS, AWS CDK

- Engineered the real-time collaboration backend using Yjs CRDTs and Redis pub/sub for conflict-free concurrent editing across users; coordinated shared API contracts across a 5-service monorepo (frontend, collab, execution, translation, infra).
- Integrated with an ECS-based sandboxed code execution service and a Gemini Lambda translation layer within AWS CDK-provisioned infrastructure, supporting multi-language Python and Node.js Docker runners.

**HeatDebt – NYC Urban Heat Inequality Auditor** | Python, DuckDB, Flask, MapLibre GL, Ollama

- Fused 6 public datasets (Street Tree Census, PLUTO, ACS income, Landsat 8 satellite temps, census tracts) into a DuckDB spatial database across 2,327 NYC tracts, engineering a custom borough-to-FIPS join across incompatible geographic identifiers.
- Computed a composite Heat Debt Score via percentile-ranked ML components and shipped a Flask + MapLibre GL 3D choropleth; integrated a local LLM agent (Ollama + Mistral 7B) for natural language to SQL queries. **Code4City Hackathon 2026.**

**Subway Dash – Real-Time MTA Congestion Pipeline** | Kafka, PySpark, Cassandra, MongoDB, FastAPI

- Designed a Lambda architecture storage layer: MongoDB TTL-indexed speed layer for live congestion scores and Cassandra batch layer for historical aggregates, ingesting MTA GTFS-RT feeds via Kafka and PySpark.
- Developed a FastAPI serving layer and Streamlit + PyDeck live dashboard rendering per-station subway congestion across NYC on an interactive map.