SIDDHARTH JAIN

Seeking ML Research Infrastructure Engineer Role

Email: *sj99@illinois.edu* | **Phone:** +1 (602) 596-7491

LinkedIn: /in/sjain166 | GitHub: /sjain166

EDUCATION

University of Illinois Urbana-Champaign (UIUC)

Master of Computer Science

• Courses Taken: Deep Learning For Computer Vision, Advanced Distributed Computing, Computer Security, Cloud Networking

Arizona State University (ASU)

August 2020 - May 2024

Expected: December 2025

Bachelor of Science in Computer Science

4.00 GPA

Courses Taken: Data Structures & Algorithms, Distributed Systems, Operating Systems, Software Design Principles, Computer Networks

WORK EXPERIENCE

Software Engineer Intern

May 2025 – Current *Scottsdale, AZ*

PayPal Inc.

Nexus

• Engineering Spring Batch pipeline to automate reconciliation of over 6.5M+ failed financial events, reducing manual SQL fixes.

- Implementing event requeue logic based on lifecycle state transitions and timestamps to ensure correctness and idempotency at scale.
- Troubleshooting **QA** failures using distributed log analysis and event state inspection; validating batch behavior across systems.
- Writing unit and integrated tests using JUnit, Mockito, and WireMock for end-to-end validation; building observability dashboards.

Cloud/DevOps Engineer

January 2025 - April 2025

Remote, USA

- Defined AWS architecture for multi-user coding platform, enabling CI/CD, ECS scaling, and region-based production readiness.
- $\bullet \ \ \text{Reduced deployment time by 90\% with 3 GitHub} \rightarrow \textbf{CodeBuild CI/CD pipelines} \ \text{for frontend, backend, and infrastructure}.$
- Refactored EC2-based stack to serverless AWS (ECS, EFS, S3, Route53, DynamoDB) with CDK, reducing infra overhead by 30%.
- Delivered production-ready deployment in 6 weeks by aligning infra roadmap with dev workflows and cross-team priorities.

Software Development Engineer Intern

May 2023 - August 2023

San Diego, CA

Amazon.com

- Built automated orchestration with **Lambda**, **SQS**, **Java & TypeScript** to back-fill a million records, saving developer time by 99%.
- Architected 1-click reusable cloud stack using AWS:CDK to perform batching for downstream services with rate-limited TPS.
- Constructed robust error handling mechanism with **automated retries** for transient errors reducing failure, ensuring 100% data integrity.
- Tested system end-to-end using JUnit5 achieving 100% code coverage. Deployed real-time monitoring dashboard using CloudWatch.

PROJECTS

PitchPal | ML Engineer | PyTorch, Torchvision, OpenCV, CNN, MediaPipe

- Designed and implemented an AI-powered multimodal analysis system that evaluates elevator pitches using speech and facial cues.
- Built a nervousness classification model using EfficientNet-B0 trained on custom-labeled facial expressions, achieving 80% accuracy.
- Integrated eye contact detection using a **pre-trained CNN-based** facial gaze tracker.
- Developed a fully automated local inference pipeline for video analysis with PyTorch, OpenCV, dlib, and MediaPipe.

360Torrent | Distributed Systems Engineer | Python (asyncio), Ansible, tmux, P2P Systems, BitTorrent

- Developed a **Python-based P2P CDN** for 360° video, reducing download latency by 42% via geo-aware chunk prioritization.
- Designed churn-resilient downloader using asyncio, adaptive health checks, and rarest-first strategy; improved reliability by 35%.
- Deployed system on 20 VMs using **tmux**, **tc**, and **Ansible**, simulating 100+ clients with region-aware network conditions.
- Integrated popularity matrix at tracker; pre-seeded hot chunks across peers, cutting startup latency by 28%.

Banking Easy | Network Engineer | Java, Network Programming, Sockets, TCP/IP

GitHub

- Implemented P2P banking application in **Java** (Sockets) leveraging network programming to perform parallel processing transactions.
- Created a shared customer cohort with a bank to perform 100+ transactions and update the system state in real-time.
- Reduced error rates by 30% by deploying advanced checkpoint and rollback algorithms, managing transaction conflicts, data corruption, and network failures.

SKILLS

Languages: Python, Java, C/C++, TypeScript, JavaScript

ML & Research: PyTorch, Torchvision, OpenCV, CNN, MediaPipe, CoreML Infrastructure & Tools: Git, JUnit, Ansible, tmux, Linux, XML, JSON, MySQL, JDBC

Cloud & Systems: AWS {CDK, Lambda, SQS, S3, RDS, EC2, ECS, EFS, API-Gateway, CloudWatch}