

Reet Nandy

Portfolio: reetnandy.com **Phone:** [+1 5189306116](tel:+15189306116) **LinkedIn:** in/reetnandy **Mail:** reet.nandy@nyu.edu **Github:** [techpertz](https://github.com/techpertz)

Software + AI Engineer with **3 years** of internship experience across 6 organizations in **distributed systems, cloud infrastructure and Machine Learning**. Expertise in system optimization, fault-tolerant microservices, and DevOps practices. Skilled in **developing and deploying AI models**, including large language models (**LLMs**), with proficiency in fine-tuning, embeddings, and quantization for **building multi-agent systems**.

Education

New York University | *Master of Science in Computer Science*

Expected May 2025

- *Relevant Coursework:* Data Structures and Algorithms, Cloud Computing, Machine Learning, Big Data Analytics, Database Systems, Software Engineering, Object-oriented Design, Probability and Statistics
- *Graduate Teaching Assistant* for CS GY 6233 - Operating System, CSCI UA 0310 - Algorithms

Manipal University Jaipur | *Bachelor of Technology in Computer Science*

Technical Skills

Languages: Python, C/C++, Java, HTML/CSS, Javascript, Typescript, Flutter.

Frontend / Backend: React.js, Next.js, Tailwind CSS, Django, Flask, FastAPI, Node.js, Express.js, Spring

AI / ML: Scikit-learn, TensorFlow, PyTorch, Hugging Face, LLM, RAG, Model Deployment, AI Workflow Automation

Cloud / DevOps: AWS (EC2, S3, Lambda, API Gateway, SQS, SageMaker), GCP, Azure, Docker, Kubernetes, Jenkins, CI/CD Pipelines

Databases & APIs: PostgreSQL, MongoDB, DynamoDB, Cassandra, Redis, Elasticsearch, Kafka, REST/GraphQL APIs

Technical Experience

Mobility Intelligence

June 2024 - December 2024

Backend Developer Intern (AI) - React.js, Django, PostgreSQL, Apache Airflow, Regression, TimeSeriesSplit

New York City, NY

- Developed a Django-ReactJS app to enable real-time ML model integration, delivering predictions with **sub-1 second** latency
- Engineered ETL pipelines processing **15M+ records** daily, automating ingestion, transformation, and validation, **saving 3 man hours** per dataset
- Developed and deployed ML models for refurbished phone price prediction, **reducing error rates to +/- 5%** over a 90 days horizon

Defence Research & Development Organisation (Govt. of India)

January 2023 - June 2023

Software Engineering Intern (R&D) - Python, Shell, PyQt, Matplotlib, SerialPy, SQL, Docker

India

- Built a Python desktop app under 30MB to decode **1.5M+ bytes/sec** from RS232 ports, enabling real-time LiDAR and GPS data analysis
- Engineered **25+ algorithms** for rut measurement analysis using LiDAR, enhancing sensor data interpretation and visualization accuracy

Solar Industries India Ltd

April 2022 - December 2022

Software Engineering Intern - Django, FastAPI, SQL, AWS, Docker, Jenkins, Kafka, Redis, Elasticsearch

India

- Led 5 interns to automate workflows for 25+ plants, deploying 5 Python projects with **75K+ lines of code**
- Engineered scalable systems to process **100K+ requests/day** and execute operation on **2.5M rows/sec** with Cassandra and Kafka
- Partnered with 5 cross-functional teams to redesign backend workflows, reducing latency for **high-traffic APIs by 20%**

Projects

[**AI/Fullstack**] AI Automation to streamline tasks from multiple communication channels [[GitHub](#)] **Stack: Python, Flask, LLMs, Postgres, ReactJs**

- Developed an open-source app that reads communication from gmail, slack, trello and categorizes as per importance.
- Achieved capabilities to store conversation-specific and global context to update existing tasks based on new information.

[**AWS/Infra**] Cloud App with Alerting and Monitoring [[GitHub](#)]

Stack: Flask, Docker, Kubernetes (Minikube, AWS EKS), Prometheus

- Containerized a Flask-MongoDB app with Docker and deployed on Kubernetes with replication, rolling update and health probes
- Integrated Prometheus and Slack for real-time alerting, ensuring scalable and monitored deployments on Minikube and AWS EKS

[**AI/AWS**] AI-Fitness Analytics Dashboard [[GitHub](#)]

Stack: ReactJs, Django, PostgreSQL, DynamoDB, EC2, Sagemaker, SNS, SQS

- Architected a hybrid AWS infrastructure to host a Django application on Beanstalk with Google Fit API for health metrics
- Designed serverless data pipelines with AWS Lambda and deployed a KNN-based exercise recommendation model on SageMaker

[**Core**] Unix Command Execution Framework [[GitHub](#)]

Stack: C++, Shell Scripting, File Handling, Multi-threading, Error Handling

- Developed a C++ framework to interpret and execute command sequences with piping, redirections, concatenations, and error handling.
- Implemented dynamic file parsing for complex command execution workflows, enabling shell operations and structured output handling

[**Core**] Performance Optimization and Analysis of Disk I/O [[GitHub](#)]

Stack: C++, File I/O Optimization, Cache Management, Benchmarking

- Optimized disk I/O performance using multi-threading, block size tuning (65536 bytes), and system call enhancements
- Achieved 15,003 MiB/s cached reads and 3000 MiB/s under cache-cleared conditions, improving throughput by 5x using Google Benchmark

[**Fullstack**] Real-Time Collaborative Whiteboard [[GitHub](#)]

Stack: Java, Spring Boot, Java Swing, WebSockets, JSON, PostgreSQL

- Designed a collaborative whiteboard with real-time updates using Swing and WebSockets for multi-user synchronization
- Enhanced performance by integrating multi-threading, batch processing, and asynchronous communication for efficiency