Shubham Rastogi

New York City, NY | shu.rastogi337@gmail.com | sr7421@nyu.edu
LinkedIn:srastogi97 | Github: SentientRamen12 | Phone: +1 (917) 818-7191

Experience

Wadhwani AI | Building AI Solutions with Social Impact, backed by BMGF, USAID, Google, Meta — New Delhi, India
Software Development Engineer 3 — May 2023 – Aug 2024

- Spearheaded the development of an innovative **AI-based oral reading fluency (ORF) solution**, revolutionizing Indic language assessment in Indian public schools and significantly impacting education quality.
- Architected and implemented a scalable Kubernetes infrastructure with **ASR ML pipelines**, enabling the processing of **3 million assessments** across Gujarat and multiple states, while reducing processing latency by 80%. Used **Kubeflow, Kubernetes, FastAPI, RabbitMQ**.
- Led a cross-functional team of education vertical in designing and developing modular, extensible services that facilitated seamless cross-platform integrations and optimized performance for low-bandwidth environments, enhancing accessibility for underserved schools.
- Orchestrated a complex migration of education infrastructure from **AWS to GCP** with minimal downtime, resulting in a 40% reduction in infrastructure costs and improved system efficiency.
- Contributed to the remediation component of the solution, leveraging GenAI to create personalized learning experiences for students. This innovative approach won the prestigious **Llama Grant for Education,** recognizing its potential to transform learning outcomes.

Ula | Sequoia funded start-up revolutionizing marketplace and B2B E-commerce in Indonesia — Bengaluru, India
Founding Engineer — Sep 2020 - May 2023

- Spearheaded the development of critical **social commerce** features, increasing monthly users onboarded by around 20%.
- Architected and implemented a robust data pipeline using **Golang, Kafka, BigQuery, and GraphQL**, aggregating live data from multiple microservices. This solution enabled custom data insights, dramatically reduced report generation latency, and enhanced overall data accessibility.
- Engineered an innovative inventory prediction system leveraging **Amazon Forecast**, which evaluated multiple factors to estimate required goods for upcoming purchase orders (POs). This significantly streamlined the PO evaluation and prediction process.
- Led a team of three engineers in extending the platform to online retailers, enabling them to manage and fulfill end-customer orders efficiently, driving a 62% increase in digital platform usage for order placement, management, and digital payments.
- Augmented existing order, catalog, inventory, and logistics services while introducing new product listing, discovery, and promotional capabilities. These enhancements increased app usage from 10% to 80% and contributed 25% to the monthly Gross Merchandise Value (GMV).

Undostres | Building D2C Fintech services for Mexico — Gurgaon, India
Software Engineer — Feb 2019 - Aug 2020

- Architected and implemented the core database and **distributed logging layers (AWS Cloudwatch)** for a cutting-edge Wallet as a Service (WaaS) product, enabling seamless multi-tenant D2C and B2B capabilities.
- Engineered robust payment processing systems, incorporating advanced features such as chargeback handling and subscription management for **digital consumer payments**, integrating with major platforms including PayPal and various credit card merchants.
- Collaborated cross-functionally to ensure the WaaS product met stringent security and compliance standards essential in the fintech industry.
- Optimized system performance and reliability, resulting in improved transaction processing times and enhanced user experience for both D2C and B2B clients.

Pingal Technologies | Building B2B NLP-based data analytics tools. — Mumbai, India
Software Engineering Intern — Jun 2018 - Aug 2018

- Engineered an automated real-time customer tracker to measure and evaluate product interest levels in convenience stores by developing a hybrid ML implementation using **Yolov3** for multiple human detection and **Deep-sort + KLT** for multiple object tracking.

Education

New York University, Tandon School of Engineering — New York, NY
M.S. in Computer Science ***(Expected Graduation May 2026)*** — Sep 2024

- Courses: Machine Learning, Applied Cryptography, Software Engineering

Delhi Technological University — New Delhi, India
B.Tech in Computer Engineering — Aug 2015 – May 2019

- Courses: DSA, Big Data, Machine Learning, Computer Vision, Computer Graphics,Distributed Systems, Artificial Intelligence

Skills

Languages: C++, Python, Golang, Typescript (Nodejs), Java
Database: MySQL, Postgres, BigQuery, MongoDB, ElasticSearch, Redis, VectorDB (Pinecone)
Frameworks: FastAPI, Django, React, Springboot, GraphQL
Cloud and Devops: Queues (SQS, RabbitMQ, Kafka), Docker, Kubernetes, Sockets, Kubeflow, Torchx, AWS, GCP

Research

A Study on Neural Networks Approach to Time Series Analysis | IEEE ICISC Paper Submission

- R. Katarya and S. Rastogi, "A study on neural networks approach to time-series analysis," 2018 2nd International Conference on Inventive Systems and Control (ICISC), Coimbatore, India, 2018, pp. 116-119, doi: 10.1109/ICISC.2018.8399024.

LANS Lab Delhi Technological University | Student Researcher (Machine Learning) — Sep 2018 – May 2019

- The research group focuses on affective computing, medical aid devices, and machine learning applications in epilepsy detection.
- Implemented a Tensorflow online prediction model with signal processing that analyzed live EEG data to predict seizures in epileptic patients. Achieved an accuracy of 81% (B.Tech Thesis).