John Doe

City, ST | 123-456-7891 | johndoe@gmail.com | LinkedIn | GitHub | Medium

Education

Redacted University, City, ST – MS in Computer Science   May 2025
Relevant Coursework: Distributed Systems, Design Patterns, Algorithms for Scalable Computing, Large Language Models
Redacted University, City, ST – BS in Computer Science   May 2021

Experience

Software Engineer, Company 1 – Cambridge, MA   May 2024 – August 2024
- Implemented **Python** tool to automate **ETL workflows** using **Airflow** by automating data ingestion and transformation processes of over **150 clinical datasets**, resulting in a **100% faster access** to insights to researchers, and data scientists.
- Developed a robust storage schema within the **Databricks Lakehouse** architecture enabling efficient cross-trial inferences and enhancing data accessibility for researchers.

Senior Software Engineer (Backend), Company 2 – Bangalore, KA   July 2021 – July 2023
- Scaled **Product Analytics SaaS platform** from ground up to over **100 customers**, developing no-code event capture system, query infrastructure over Clickhouse, supporting robust data insights over any Web/Mobile application.
- Co-Designed scalable 5+ backend services using **JAVA** (**Spring Boot**) deployed in a **multi-cloud setup** over AWS and Azure with automated **CI/CD pipeline** using **Jenkins** and **Ansible**, achieving independent **deployments every 2 weeks.**
- Optimized **database schemas** and indexing strategies in **ClickHouse**, resulting in a 60% improvement in query performance and supporting real-time analytics capabilities.
- Implemented a **Superset** cluster on ClickHouse for internal data quality and performance metrics, providing internal insights for over **10 engineering teams**. Awarded a **Spot Award** for enhancing engineering team productivity.
- Enhanced dashboard performance by consolidating three SQL queries into a single, efficient query using **window functions** for total and partitioned counts, reducing server load by nearly 50% and accelerating data retrieval times.
- Achieved **90% test coverage** by developing **integration test** cases on the Playwright framework for the newly developed features – Form Field Capture, End User SSO Authentication ID Capture, Google Analytics Event Forwarding.
- Designed and Implemented **export APIs** for User Journey analytics feature, integrating **gRPC layer** within Analytics microservice for communication with **Scheduler services** to trigger export jobs according to scheduling rules.
- Awarded as **Success Engineer of the Quarter** in 2022 for resolving customer issues with utmost priority and resolution, helping bring down churn rate. Resolved 5+ P0 issues relating within deadline with support from multiple stakeholders.
- **Mentored 5+ junior engineers** and interns providing guidance on design, code architecture, best practices and code reviews.

Software Engineer Intern, Company 2 – Bangalore, KA   January 2021 – June 2021
- Modernized and created reusable charts (Funnel, Trend, Bar) and table components in **React** (**TypeScript**) using FusionCharts and AG Grid, while gradually deprecating legacy GWT code.
- Created detailed **API documentation** using **Swagger**, improving onboarding and reducing support **queries by 30%.**

Research Assistant, Redacted University – City, ST   January 2020 – May 2020
- Conducted research on collaborative **distributed computing** across IoT swarms, crafting a **Python**-based benchmarking framework to evaluate performance and implementing simulation controls for key **network parameters**.
- **Co-authored two peer-reviewed papers** and presented findings at the Proceedings of International Conference on Paradigms of Communication, Computing and Data Sciences.

Skills & Certifications

Programming Languages: JAVA, Go, C++, Python, HTML/CSS, Javascript/Typescript
Technologies: Docker, Kubernetes, Ansible, Jenkins, Kafka, Spark, Spring Boot, Django, REST API, Superset, GWT, React, Fusioncharts, Git, Nginx, Azure/AWS, Postman
Databases: Clickhouse (OLAP), Cassandra, PostgreSQL, Redis

Projects

Distributed Message Queue System
- Built my own mini "**Kafka**" from scratch using **Raft consensus** in **Java**, implementing key distributed systems concepts such as leader election, log compaction, and fault tolerance, handling of over **10K+ messages per second** with an 8-core machine.
- Utilized SofaJ-Raft library for **log replication**, consensus among distributed services, automated multi-service deployment using **Docker Compose**, and implementing 200+ JUnit test cases covering system logic, with **JMeter** scripts for performance testing.

News Application
- Developed an **iOS news app** using **Swift** with a TikTok-style interface, fetching Google News RSS feeds and leveraging **GPT (LLM)** to generate concise summaries for enhanced user engagement.
- Engineered a **Docker** based **Flask** backend on **AWS EC2** with data persistence in **PostgreSQL**, on a secured **VPC** network.
- Trained a **LSTM (Neural Network)** classifier to classify news articles into 5 categories using **PyTorch**.