NIKHIL DINESH KAMATH

+44 7435 912307 • nikhildineshkamath@gmail.com
linkedin.com/ndkamath • github.com/ndk03

EDUCATION

University at Buffalo – State University of New York, USA
Master of Science, Computer Science and Engineering
Courses: Analysis of Algorithms, Deep Learning, Pattern Recognition, Computer Vision, Reinforcement Learning
SJB Institute of Technology, India
Bachelor of Engineering, Computer Science and Engineering
Courses: Data Structures and Algorithms, Object-Oriented Programming and Technology, Database Management Systems

TECHNICAL SKILLS

Languages: Java, Python, SQL, Groovy
Libraries and Frameworks: NumPy, Matplotlib, OpenCV, TensorFlow, Pandas, Keras, PyTorch, SpringBoot, Spark, Kafka, Hadoop, Presto, REST, Jersey
Database Technologies: AWS Singlestore, SybaseIQ, Trino, Elasticsearch, MongoDB
Build and CICD Technologies: Gradle, GitLab, Jenkins, Maven, Docker
Tools: Windows, Linux/Unix, Git, VS Code, IntelliJ, Terraform, AI Agents

EXPERIENCE

Engineering Associate, Goldman Sachs **01/01/2023 – Present**

- Designed and implemented a **new low-latency data ingestion and delivery platform** processing **1.5M+ rows of structured data**, integrating **three external vendor feeds** for family office reporting. Led **data modelling**, **integration design documentation**, **core development**, and **end-to-end testing**.
- Designed and delivered a **high-performance, low-latency REST API** to model **complex account hierarchy data**, consistently achieving **sub-200 ms response times** with **high availability and failure resilience**. Enabled **automated document generation** for clients' **alternative investment products** by accurately identifying in-scope parties and accounts.
- Built and optimized a **Spark-based data transformation pipeline** using **Trino**, integrating data from multiple heterogeneous sources and reducing **data delivery time per consumer from ~60 minutes to under 10 minutes**.
- Developed backend services for a **Data Quality Platform**, a **data aggregation and validation solution** focused on **data integrity and reporting consistency**, reducing impacted accounts to **<5%** and achieving an **85% year-over-year reduction in data quality issues**.
- Reduced **false positives in data quality monitoring** by implementing **rule-based controls using SparkSQL and Groovy**, aligned with Operations requirements. Decreased **data breaks from ~100,000 to 500 within 4 months**, significantly improving operational efficiency and issue resolution.

Engineering Analyst, Goldman Sachs **07/26/2021 – 01/01/2023**

- Implemented and validated **vendor-built data quality controls in production**, ensuring timely **detection of data issues** and **stale forward data** based on **pre-defined business and technical criteria**.
- Designed and implemented **low-latency, near-real-time data acquisition and processing pipelines** using **Kafka and Apache Spark**, **modelling external vendor data** to align with internal data schemas and applying **business logic transformations** to deliver **client-ready datasets to the reporting portal within 15 minutes**.
- Reduced **data latency and architectural redundancy** by replacing a **dual-persist architecture** with a **Kafka-based messaging queue**, enabling the **decommissioning of a redundant MemSQL (SingleStore) database** and simplifying the overall data platform.
- **Mentored and guided interns** to onboard multiple datasets onto the Data Quality platform, identify data issues, and **design and implement automated data quality rules** to detect anomalies and inconsistencies, strengthening data governance, compliance, and reporting reliability.