

ROHIT KAMINENI

Hartford, CT | kaminenirohit1@gmail.com | +1 4847441785 | [LinkedIn](#)

Results-driven Data Engineer with 4+ years of experience delivering scalable data and ML solutions in enterprise cloud environments. Proven expertise in developing real-time data pipelines, implementing MLOps architectures, and optimizing cloud infrastructure. Recognized for transforming business processes through automation, data-driven decision-making, and cross-functional collaboration. Strong foundation in AWS, GCP, Kubernetes, and big data tools, with a track record of reducing operational overhead and improving forecasting accuracy. Adept in Python, SQL, Spark, Kafka, and MLOps using Vertex AI and Airflow.

WORK EXPERIENCE

Kroger Technology & Digital (Consulting through ZopSmart)

Bangalore

Data Engineer, Kroger Technology

Aug. 2019 – June 2023

- Transitioned seamlessly into the Data Engineering team after leading the migration of the Products BFF microservice from PCF to on-prem Kubernetes, embedding monitoring hooks and custom observability features that enhanced SLA tracking by 40%.
- Designed and deployed scalable ETL pipelines on Databricks and Airflow using PySpark, processing over 2TB/day of product interaction logs and driving a 23% increase in personalization accuracy across Kroger's digital storefront.
- Engineered a scalable data ingestion framework on GCP, integrating Kafka, REST APIs, and GCS to support real-time analytics; enabled marketing and ops teams to run advanced ML workloads in BigQuery with sub-minute latency.
- Architected a cost-optimized data lake using GCS, Dataflow, and Glue Catalog, reducing data pipeline failures by 30% and slashing report generation times from hours to minutes for 10+ BI stakeholders.
- Delivered interactive dashboards using Tableau and BigQuery, enabling 50+ category managers to monitor customer behavior and SKU trends across 100+ stores—increasing engagement with data-driven decisions by 35%.
- Led MLOps enablement for sales and loyalty models on Vertex AI, implementing CI/CD and model retraining pipelines; cut model drift by 18% and accelerated deployment cycles from weeks to days.
- Developed reusable SQL and dbt models for handling SCD Type-2 transformations, aggregation logic, and anomaly detection, automating weekly reporting workflows and saving ~60 engineer hours/month.

Software Development Engineer, ProdSquad (CX Web Products And Search), Kroger Technology, CX vertical

Aug. 2019 – Dec. 2019

- **Products BFF:** Supported 15 million daily requests on the Kroger website by building and scaling a SpringBoot microservice (Products BFF), migrating APIs to a 3-tier architecture, and setting up CI/CD workflows on On-Prem Kubernetes.
- **Radon Framework:** Enabled multiple services in the Java vertical by developing the Radon framework, incorporating company-specific features like annotation processing and Kafka support, and maintaining comprehensive documentation.

Awards, Hiring, Mentorship, and Training

- **Pivot Polaris - ZopSmart:** Recognized by clients for delivering high-impact engineering solutions that drove measurable business outcomes.
- **Talent Acquisition:** Spearheaded 300+ technical interviews across SDE and data engineering roles, helping scale talent across multiple Kroger tech divisions.
- **Internship Mentorship:** Led onboarding and project mentoring for two cohorts (10+ interns each), with 80% receiving full-time offers or extensions after training in enterprise-level software practices.

Tcube

Hyderabad, India

Data Engineer

Jan. 2019 – June 2019

- Built batch data pipelines using Python and SQL to aggregate and standardize logistics data from multiple carriers, enabling daily freight performance reports and improving data availability for downstream analytics teams.
- Collaborated with ERP consultants to automate data ingestion and cleansing for last-mile logistics operations; enabled real-time KPI dashboards that accelerated decision-making for transportation and warehousing clients.

EDUCATION

UNIVERSITY OF CONNECTICUT

Hartford, CT

Masters in Business Analytics and Project Management

Aug. 2023 – Dec 2024

- **GPA:** 3.65 / 4.0
- **Key Coursework:** Statistics in BA, Data Management & Business Process Modeling (SQL); Predictive Modeling (SEMMA, Machine Learning Models); Data Science using Python (EDA, Pandas, Numpy, Scikit-learn); Big Data Analytics using Cloud (Hadoop, Cloud Computing); Data Mining & Business Intelligence (Time-Series Forecasting and Text Mining); Visual Analytics (Tableau, Power BI)
- **Capstone Project:** *AI-Powered Chatbot for Cryptocurrency Platform (with Ananda, mentored by Prof. Sudip Bhattacharjee)*
Developed an LLM-integrated chatbot that analyzed dynamic trading patterns and user behavior to surface timely, personalized insights; contributed to architecture design, model selection, and deployment strategies to enhance user engagement and align outcomes with business goals.
- **Research Project:** *End-to-End Analytics of E-Commerce Platforms*
Enabled data-driven decision-making by building models for price and sentiment analysis, leveraging Hadoop and MLOps pipelines to process and update datasets, resulting in a 30% improvement in predictive accuracy.
Enhanced e-commerce insights through Tableau visualizations, improving stakeholder engagement by 40%, after scraping and analyzing data from multiple websites.

- **GPA:** 3.0 / 4.0
- **Key Coursework:** Machine Learning, Artificial Intelligence, Object Oriented Programming, Linear Algebra, Operating Systems, Data Mining.
- **Leadership:** *Coordinator, Student Aid Fund (SAF):* Led a student committee that evaluated and helped 120 students needing financial aid to pursue their academics.
Captain, College Snooker Team: Led the snooker team for a year during which time we participated in, and won events at 3 National level sports festivals.

PROJECTS

Databricks Retail Insights – Customer Segmentation at Scale

Developed a scalable ETL pipeline on **Databricks** using **PySpark** and **Delta Lake** to process **millions of daily transactions** across multiple retail regions; applied clustering techniques to segment customers, driving **20% improvement in targeted marketing ROI**.

Uber Data Analytics – GCP-Hosted End-to-End Pipeline

Orchestrated ingestion and transformation of ride-level Uber-like datasets using **Mage**, **BigQuery**, and **Looker** on **GCP**; created a modular pipeline to extract insights on supply-demand mismatches, peak pricing zones, and operational bottlenecks across city clusters.

StackOverflow Developer Trends – BigQuery at Scale

Analyzed 2.8M+ open-source GitHub projects using **Google BigQuery** to identify developer trends like weekend commit patterns, language shifts, and tab vs. space preferences; engineered SQL-based data marts to deliver curated views and dashboards for codebase analysis.

Scraping Inflation Indicators from CommonCrawl

Designed a Spark-based ETL workflow to extract product pricing data from **CommonCrawl** archives, using **Amazon Athena** and **EMR** to transform WARC files into structured **Parquet**, and applied **Pandas** to compute inflation indices across categories and time horizons.

Cryptocurrency Price Prediction

Built a predictive ML model to forecast crypto price movements over a two-week horizon; achieved an F1 score of **76.5%** by applying advanced **feature engineering**, model tuning, and evaluation strategies; improved accuracy to **74.5%** after iterative enhancements using ensemble techniques.

McDigest – McDonald's Reviews Analysis

Uncovered **20% improvement opportunities** in customer satisfaction by extracting actionable insights from **33,000+ reviews** using **SAS Enterprise Miner**; led sentiment analysis and unsupervised clustering to inform regional operational strategy across U.S. outlets.

Sales Forecasting for a Retail Superstore

Conducted time-series analysis on historical sales data to optimize inventory planning; experimented with **ARIMA** and **Exponential Smoothing**, selecting the most accurate model and enabling data-driven marketing campaigns by uncovering **seasonal and regional demand trends**.

Travelogy – SQL-Based Travel Booking Engine

Improved booking efficiency and personalization for an online travel agency by designing a normalized **SQL database** and ERD; integrated service and feedback data to enable dynamic customer insights and reduce data retrieval times by **30%**.

SKILLS

Languages & Programming: Python, SQL, Java, R, Go, C++, HTML, CSS, JavaScript

Big Data & Data Engineering: PySpark, Apache Airflow, Databricks, Kafka, Hadoop, Hive, dbt, SparkSQL, Dataflow, GCS, Glue Catalog, BigQuery

Cloud & MLOps: GCP (BigQuery, Vertex AI, Dataflow, GCS), AWS (Lambda, S3, RDS, Sagemaker), Azure, Docker, Kubernetes, CI/CD, On-Prem Kubernetes, Databricks Workflows

Databases & Storage: PostgreSQL, Oracle, MySQL, MariaDB, MongoDB, Snowflake

Machine Learning & Analytics: Scikit-learn, TensorFlow, PyTorch, Keras, Pandas, NumPy, Matplotlib, Seaborn, SAS, JMP

Visualization & BI Tools: Tableau, Power BI, MS Excel

Frameworks & APIs: SpringBoot, REST APIs, Swagger, Postman, NodeJS, Microservices

Tools & Platforms: Git, GitHub, JIRA, IntelliJ, VS Code, Jupyter, Colab, PCF

Other: Agile Methodologies, Project Management, MLOps, ETL/ELT Pipelines, Data Modeling, Observability & Monitoring

CERTIFICATIONS

- **Udemy:** Data Science using Python; GCP - Google Cloud Professional Data Engineer Certification; MLOps with AWS
- **Coursera:** Deep Learning Specialization