

ML Engineer/Data Scientist with 6+ years developing evaluation infrastructure and production ML systems.

Technical Skills

Languages (Experienced)

Python, Scala, Java, C/C++, SQL

ML

PyTorch, cudNN, Spark

Infrastructure

Kubernetes,
OpenShift Container Platform

Databases & Data

Db2, PostgreSQL, Elasticsearch,
Lance DB

Development

MCP, REST, gRPC

Education

Dual Masters

Applied Statistics, UT Austin
2019

BA, Classical Languages

Hunter College, CUNY
2015

Languages

- English (Native)
- Arabic (Professional)
- Japanese (N2)

Professional Experience

ML Engineer & Data Scientist

IBM Expert Labs | 2021 – Present

Develop AI solutions and tooling and support production deployment systems for enterprise clients.

AI Agent Evaluation System

- Develop Agentic procurement platform integrating Dun Bradstreet data with ERP systems (Ariba, Workday). Implemented quality assurance tracking tool to track execution accuracy, response grounding, and task completion rates using Langfuse and custom tooling.
- Built observability layer using Prometheus/Db2 for tracking agent behavior and performance in production. Real time monitoring of Precision/Recall for NAICS/UNSPC semantic search.
- Created automated testing pipelines for input validation and edge case handling. Integrated evaluation metrics into CI/CD workflow
- Evaluated performance of RAG

LLM Serving Infrastructure on OpenShift AI

- Managed MAAS (Model as a Service) platform deployed to Openshift. Used Nvidia GPU Operator to split 8 H200 GPUs and automate assignment of models to MIG instances.
- Built scheduling logic for optimal GPU allocation across concurrent inference workloads.
- Deployed and managed various Llama models as well as GPU accelerated TTS, reranker and embedding models to serve upstream APIs
- Built Grafana dashboards and PromQL queries for tracking token and GPU consumption, accuracy across various use cases

Model Evaluation & Testing Framework

- Built evaluation platform in Scala/C++ for standardized model testing across ML projects. Implemented in-db SIMD accelerated metrics for accuracy, bias detection, and performance consistency
- Implemented continuous evaluation system detecting model drift and data quality issues. Built automated testing for new model versions before production deployment
- Implemented in-database analytics functions in C++ for Db2 MPP cluster processing multi-terabyte datasets. Optimized query performance for parallel data processing workloads

ML Engineer

IBM Cloud Pak Acceleration Team | 2019 – 2020

Developed containerized ML applications and deployment pipelines on OpenShift.

- Built MLOps pipeline using Kubeflow on OpenShift for model training and deployment. Implemented A/B testing framework for model evaluation in production
- Developed Scala microservices for ML model serving with authentication and rate limiting. Deployed containerized applications on Kubernetes handling production inference workloads
- Created data integration pipelines connecting federated Db2 warehouse with external data sources (MySQL, Hive) for feature engineering

Data Specialist

Princeton University | 2018 – 2019

Built data quality and statistical analysis pipelines for research projects.

- Developed automated data validation system in Python and R for survey data quality assessment.
- Built ETL pipelines for data cleaning and standardization. Created reproducible analysis workflows supporting research publications
- Implemented statistical methods for anomaly detection and consistency checking. Conducted regression and clustering analysis on survey datasets