BHARATHRAAJ NAGARAJAN

United States (Open to Relocate) | +1 716-228-7052 | bharathraajnagarajan@gmail.com | LinkedIn | GitHub | Portfolio

SUMMARY

Machine Learning Engineer with 3.5 years of experience building and deploying production ML systems across distributed, cloud-native environments. Specialized in model serving, inference optimization, and end-to-end ML pipelines, with hands-on experience in PyTorch, TensorFlow, Spark, and Kubernetes. Proven track record of operationalizing ML solutions with high reliability, low latency, and scalable data infrastructure across AWS, GCP, and Azure.

EDUCATION

University at Buffalo, The State University of New York **Aug 2024 - Dec 2025**
Masters in Data sciences and Applications (GPA: 3.9/4.0)

- **Coursework:** Database Management Structure, Data Intensive Computing, Machine Learning, Data Modeling, Cybersecurity

Anna University - Sri Sivasubramaniya Nadar College of Engineering **Aug 2017 - May 2021**
Bachelor of Engineering, Minor in Computer Science

WORK EXPERIENCE

Tata Consultancy Services | *SYSTEMS ENGINEER (Machine Learning)* **Oct 2021 - Apr 2024**

- Designed, built, and deployed **production-grade ML systems** supporting enterprise-scale analytics and decision workflows, owning the full lifecycle from data ingestion and feature engineering to model serving and monitoring.
- Optimized large-scale **data preparation and ETL workflows** for ML training and inference, improving processing efficiency by ~47% through execution tuning and pipeline redesign.
- Deployed **real-time and batch inference services** on Kubernetes, optimizing for **low-latency and high-throughput**; implementing containerized model serving with health checks, logging, and automated restarts to maintain **99.9% production uptime**.
- Built and optimized **data and feature pipelines** using Spark (PySpark, Spark SQL), Kafka, HDFS, HBase, and Kudu to support 100M+ record datasets, reducing query latency by ~40% for real-time ML workloads.
- Collaborated with cross-functional engineering and product teams to translate business requirements into **scalable ML system designs**, participating in design reviews, debugging production issues, and improving system reliability.
- Supported **model evaluation, debugging, and iteration** in production environments by instrumenting metrics, validating data quality, and diagnosing inference and pipeline failures across distributed systems.

U.S. INTERNSHIP EXPERIENCE

Media Sales Plus Inc. | *Data Engineering Intern - Analytics & ML Systems | Buffalo, NY* **Aug 2025 - Dec 2025**

- Built **data ingestion and transformation pipelines** in Python to ingest, validate, aggregate, and backfill multi-client datasets with heterogeneous schemas, enabling downstream analytics and ML-driven insights.
- Designed and implemented a **cloud-deployed analytics and forecasting platform** (Streamlit-based) to support executive reporting, anomaly detection, and time-series forecasting.
- Integrated **CI/CD pipelines** using GitHub Actions and Docker to support repeatable deployments and automated execution of data and ML workflows.
- Collaborated with stakeholders to standardize data models and pipeline logic, improving data consistency and reliability for analytics and future ML use cases.
- Designed an LLM-ready, retrieval-augmented interface for metric querying, enabling natural-language interaction with analytical systems and laying the foundation for agentic ML workflows.

SELECTED ENGINEERING PROJECT

Production ML Pipeline: End-to-End Inference & Monitoring | Link

- Designed and built a production-grade ML pipeline covering data ingestion, model training, versioned deployment, and monitoring, with automated workflows and rollback-aware deployment strategies.

TECHNICAL SKILLS

- **Programming & Systems:** Python, SQL, R, Bash, Scala, Data Structures & Algorithms, Distributed Systems, Performance Debugging, API Design
- **Machine Learning & Inference:** PyTorch, TensorFlow, Scikit-learn, Transformers, NLP, CNN, OpenCV, Reinforcement Learning (exposure), Prompt Engineering, Model Serving, Batch & Real-Time Prediction Systems, Inference Optimization, Model Evaluation, Feature Engineering
- **Data & Streaming Systems:** Hadoop, Apache Spark (PySpark, Spark SQL), Kafka, HDFS, Hive, HBase, Kudu, Batch & Streaming Pipelines
- **Cloud Platforms:** AWS, Azure, Google Cloud Platform (BigQuery, Compute), Microsoft Fabric, Databricks, Snowflake
- **ML Infrastructure & MLOps:** Kubernetes, Docker, CI/CD (GitHub Actions), MLflow, Model Deployment, Versioning, Rollback, Monitoring, Drift Detection (data & prediction)
- **Certifications:** AWS Certified Solutions Architect - Associate (In Progress)