

MICHAEL GURULE

Applied Scientist and ML Architect

Denver, CO

michaelgurule1164@gmail.com

linkedin.com/in/michael-gurule-denver

github.com/michael-gurule

michaelgurule.com

SUMMARY

Principal Applied Scientist and ML Architect with nearly 10 years of experience designing statistical models and the distributed cloud systems that run them at scale. Specializes in high-throughput, sparsely labeled, and non-stationary operational data. Currently owns the end-to-end modeling, AWS architecture, and technical roadmap for an enterprise anomaly detection, root-cause correlation, and forecasting platform across 32M+ broadband endpoints. Combines deep mathematical foundations (MS in Applied Mathematics) with a track record of leading greenfield ML systems across telecommunications, financial services, and industrial commodities.

EXPERIENCE

Principal Engineer II, Telemetry and AI Operational Intelligence

Oct 2025 – Present

Spectrum

Lead Applied scientist and ML architecture for an anomaly detection, root-cause correlation, and forecasting platform across 32M+ broadband endpoints.

- Architected the end-to-end AWS platform spanning streaming ingestion, real-time inference, automated retraining, and drift monitoring, leading the migration of legacy on-premises telemetry and ML workflows to a hybrid cloud architecture.
- Co-authored the multi-phase platform roadmap with the Director of Operational Intelligence, scaling from initial POC to cross-unit forecasting, and currently defines execution scope and safety guardrails for autonomous remediation agents.
- Eliminated signal truncation caused by legacy rule-aggregated inputs by re-architecting the ingestion pipeline to consume raw, full-resolution telemetry directly, decoupling deterministic threshold alerting from statistical detection.
- Engineered a custom ensemble model for high-frequency network telemetry, synthesizing mathematical methodologies from industry time-series literature into a proprietary framework tailored to non-stationary fleet data and benchmarked against a deterministic rules baseline.
- Designed a peer-group scoring framework that evaluates heterogeneous endpoints and hubs against cohorts in equivalent operating states, running Random Cut Forest and seasonal residual analysis on a five-minute cadence to eliminate false positives from diurnal load swings without the overhead of fitting individual per-device models.
- Compressed feature dimensionality to bound compute overhead and satisfy strict five-minute fleet-wide inference SLAs, reducing memory and latency bottlenecks while preserving the high-leverage fault signatures that separate true hardware and network anomalies from noise.
- Developed a cross-topology and hardware correlation engine with confidence weights calibrated from operator feedback, consolidating isolated network alerts into ranked root-cause candidates.
- Govern production model releases and evaluation strategy, establishing ground-truth labeling frameworks, inference SLAs, and retraining cadences aligned with underlying statistical assumptions while using production diagnostics to enforce upstream data validation gates.

Senior Data Scientist and Machine Learning Engineer

Mar 2020 – Aug 2025

Sedgwick

- Architected a serverless AWS Lambda inference pipeline that automated risk scoring across 3,000+ municipal jurisdictions, implementing strict version control across training datasets and Bayesian model priors.
- Quantified incremental marketing lift and portfolio profit drivers using Bayesian structural time series (BSTS) and multivariate regression, directly contributing to a six-figure ROI improvement.
- Developed probabilistic models for attribution and client segmentation. Implemented Bayesian neural networks, Latent Dirichlet Allocation (LDA) topic models, and Gaussian mixture models in PyMC and TensorFlow.
- Automated covariate and concept drift monitoring using Kolmogorov-Smirnov (KS) tests and Population Stability Index (PSI), triggering automated retraining pipelines to preserve model calibration across shifting regulatory regimes.
- Unified fragmented ERP systems and public-sector APIs into a centralized AWS data lake, reducing end-to-end data latency for portfolio risk monitoring by 20%.

JS Held

- Delivered custom time-series and econometric regression models across 13+ advisory engagements, forecasting multimillion-dollar financial and operational trajectories for enterprise industrial and financial clients.
- Led cloud migration scope assessments and schema architecture during a major M&A integration, mapping legacy data systems to AWS and relational SQL infrastructure.

SELECTED PROJECTS

Open-source implementations and reference architectures available on GitHub.

- LOCANT: Multi-Sensor & State Estimation.** Fuses asynchronous RF positioning (TDOA/FDOA), thermal imaging, and telemetry streams via an Extended Kalman Filter (EKF) for real-time target tracking.
- CONSTELLATION: LEO Satellite Fleet Diagnostics.** Detects telemetry anomalies and predicts subsystem time-to-failure across low-Earth-orbit constellations using an Isolation Forest and LSTM autoencoder ensemble paired with Temporal Fusion Transformers and Cox proportional hazards models. Engineered specifically to maintain calibration under sparse failure labels and cyclical orbital load.
- MERIDIAN: Regime-Switching Portfolio Optimization.** Integrates hidden regime detection, dynamic covariance estimation, and GARCH(1,1) return forecasting into a convex portfolio optimizer accounting for realistic transaction costs across 25+ assets and 10 years of market data.
- HYPERION: Multi-Agent Reinforcement Learning (MARL).** Coordinates autonomous UAV swarms in Ray RLLib and PyTorch, coupling physics-informed neural networks (PINNs) for trajectory dynamics with graph neural networks (GNNs) for decentralized agent communication. Achieves sub-100ms inference latency and an 85% interception rate across 20+ concurrent agents.

TECHNICAL SKILLS

- Statistical Modeling:** Regularized regression, GAMs and splines, local regression, Bayesian and hierarchical models, Bayesian structural time series (BSTS), survival analysis, change-point detection, Kalman filtering and state estimation, hypothesis testing
- Machine Learning:** Gradient boosting, random forests, Isolation Forest, Random Cut Forest, LSTM autoencoders, Temporal Fusion Transformers, Gaussian mixture models, reinforcement learning, graph neural networks (GNNs), physics-informed neural networks (PINNs)
- Evaluation and Model Ops:** Feature engineering, regularization, model selection, cross-validation for time-dependent data, probability calibration, drift detection (KS, PSI), hyperparameter tuning, experiment tracking, model versioning
- Mathematics:** Convex and non-convex optimization, information theory, estimation bounds, Fisher information, numerical linear algebra
- Cloud and Systems:** AWS (SageMaker, Bedrock, MSK, Managed Apache Flink, Lambda, S3, Glue, Athena, Neptune, EventBridge), Kafka, Flink, Druid, Spark, Docker, Kubernetes, CI/CD, FastAPI
- Languages:** Python (NumPy, SciPy, pandas, scikit-learn, PyTorch, TensorFlow, PyMC, CVXPY), SQL (PostgreSQL, MySQL), JavaScript

EDUCATION AND CREDENTIALS

- | | |
|---|------|
| MS in Applied Mathematics, University of Colorado Boulder | 2024 |
| BS in Finance and Risk, University of Colorado Denver | 2017 |
| AWS Certified Machine Learning Engineer, Associate | |