

Nishanth Jayram

📍 Berkeley, CA ✉ nishanthjayram@gmail.com 📞 (408) 763-6586 🌐 newtrino.ink 📄 nishanthjayram
🔗 nishanthjayram

Summary

Data and Machine Learning Infrastructure Engineer specializing in large-scale ingestion systems, production ML pipelines, and validation frameworks across healthcare and consumer engagement data.

Experience

Protege, Data Engineer - San Francisco, CA / Remote Jan 2026 – present

- Standardized healthcare data ingestion across provider integrations by designing production pipelines for 500+ TB EHR and imaging datasets, improving reliability and maintainability at scale.
- Developed reusable raw-to-conformed ingestion and normalization frameworks to onboard heterogeneous provider data through a shared processing path, reducing one-off implementation work across sources.
- Strengthened data quality and validation through schema checks, normalization safeguards, and drift monitoring, reducing recurring downstream pipeline failures by 75%.
- Improved production operations across 5+ pipelines with Dagster-based orchestration, observability, alerting, runbooks, and identity-mapping migrations that strengthened reliability and made downstream joins more deterministic.

Mastercard, Machine Learning Engineer - San Francisco, CA Aug 2022 – Jan 2026

- Architected a consumer engagement forecasting ecosystem that improved targeting accuracy by 50% using hierarchical transaction and redemption propensity models.
- Engineered feature pipelines from six sources, including transaction history, offer attributes, redemption patterns, and merchant/category signals, improving merchant-level model performance by 15%.
- Designed and deployed end-to-end training and inference pipelines across 20+ platforms, supporting production ML releases for hundreds of millions of users.
- Migrated core ML data workflows to Delta Lake, improving data versioning, reliability, and deployment speed across multi-platform releases.
- Built React and TypeScript dashboards for model performance, feature importance, and production insights while mentoring junior engineers and supervising intern-led analytics platform development.

Education

BS **University of California, Santa Cruz**, Computer Science & Mathematics Santa Cruz, CA
Sept 2019 – June 2022

Projects

Karektar

Built a full-stack bitmap font editor with canvas-based rendering, interactive editing workflows, and OpenType export using React, TypeScript, and opentype.js.

Automated Retraining of Cyber-physical Systems (Senior Thesis)

Senior thesis investigating counterexample-guided training for CPS robustness using Scenic and probabilistic sampling.

Skills

Languages: Python, SQL, TypeScript, JavaScript, Go

Data & ML: PySpark, Pandas, Scikit-learn, PyTorch, TensorFlow, Delta Lake

Infrastructure: AWS S3, Glue, Batch, Athena, Dagster, GCP

Databases & Web: PostgreSQL, Snowflake, React, GraphQL, Vite