

ALANKRIT SRIVASTAVA

469-604-8655 | alankrit04jan@gmail.com | <https://www.linkedin.com/in/alankrit-srivastava-91407b116/> | <https://github.com/alankrit04>

Machine Learning Engineer specializing in time-series forecasting, financial anomaly detection, and explainable ML systems. Experienced in deploying drift-aware ML models, continuous retraining pipelines, and SHAP-based governance frameworks across multi-million record transaction datasets. Strong background in Python-based production ML, statistical modeling, and financial data pipelines. MS in Advanced Data Analytics (GPA 4.0).

EDUCATION

University of North Texas, Denton, TX

Master of Science, Advanced Data Analytics (GPA: 4.0/4.0)

Aug 2023 – May 2025

SKILLS

Languages: Python, SQL, Java

ML & AI: PyTorch, XGBoost, LSTM, Time-Series Forecasting, Anomaly Detection, Regime Modeling

LLM & Retrieval: RAG, FAISS, ChromaDB, Prompt Engineering, HuggingFace

Model Validation & Risk: Walk-Forward Validation, OOD Detection, Drift Monitoring, SHAP, Failure Analysis

MLOps & Cloud: AWS, MLflow, ONNX Runtime, Docker, CI/CD

Data Systems: PostgreSQL, Snowflake, Streaming Pipelines

PROFESSIONAL EXPERIENCE

Data Engineer Intern

Verity Advisor LLC, Denton, TX

Jan 2025 – May 2025

- Led the design and delivery of real-time anomaly detection and regime classification systems across 1M+ daily financial transactions, enabling early-warning risk signals used by multiple business stakeholders.
- Owned end-to-end development of drift-aware time-series models with out-of-distribution validation, improving early risk detection precision by 18% under non-stationary data conditions.
- Built automated evaluation, retraining, and performance regression pipelines, enabling daily production refresh with fully reproducible training and monitoring.
- Optimized inference pipelines for real-time deployment, ensuring scalable and stable predictions under continuously changing data distributions.
- Guided non-technical stakeholders through model interpretation and anomaly diagnostics, translating ML outputs into actionable client-ready decisions and reducing manual investigation time by 12+ hours per week.
- Delivered production ML systems used by external business users for real-time operational decision-making and risk monitoring.

Graduate Teaching Assistant

University of North Texas, Denton, Texas

Aug 2024 – May 2025

- Built Python-based validation, anomaly detection, and clustering templates used across 300+ student ML submissions, reducing evaluation workload by 40%.
- Designed reproducible ML workflow scaffolds for dataset validation, mislabeled sample detection, and model diagnostics.
- Mentored 30+ advanced students on error analysis, feature leakage, and model interpretability, reinforcing production-grade ML practices.

DevOps Engineer

Pitney Bowes, New Delhi, India

Dec 2020 – Aug 2023

- Engineered distributed Java/Python platforms processing 10M+ operational events/day, supporting real-time ML inference availability.
- Built streaming ingestion and anomaly monitoring pipelines that reduced data-quality failures by 33% across logistics workflows.
- Implemented microservices and optimized feature pipelines, improving ML inference feature availability by 30%.
- Automated CI/CD pipelines (Docker, Jenkins), reducing batch processing cycles by 25% and operational failures by 20%.

PROJECTS

Explainable Financial AI Analyst

Jan 2025 – Present

- Built an explainable demand forecasting system using XGBoost + SHAP + FAISS-backed RAG over 2.5M+ time-series records.
- Achieved 95%+ feature attribution stability, improving forecast traceability by 42%.
- Developed LLM-based explanation and anomaly diagnostics agents, reducing manual interpretation by 60%.
- Integrated 6+ data sources with automated rolling feature generation and nightly retraining, enabling sub-5-minute inference.

Finance Sentiment & Forecasting System

Aug 2023 – Dec 2024

- Built scalable NLP ingestion and preprocessing pipelines for 100K+ financial news articles, aligning sentiment signals with ticker-level market movements.
- Trained and optimized LSTM and XGBoost models combining sentiment embeddings with technical indicators, improving directional accuracy by 20% over price-only baselines.
- Reduced total training and evaluation runtime by 29% through vectorized rolling windows, feature pruning, and strict walk-forward validation to eliminate leakage.

Formula One Performance Analytics Dashboard

Jan 2024 – May 2024

- Aggregated, cleaned, and harmonized **18 years (2003–2021) of Formula One driver and constructor performance data** into a **regulation-aware analytical dataset**.
- Built an **era-normalized driver efficiency index** correcting **aerodynamic and engine-era distortions** across regulation cycles.
- Designed and deployed an **interactive Tableau dashboard** enabling **multi-season performance decomposition across drivers, teams, and eras**.
- Delivered a **data-validated performance comparison framework** separating **driver skill from car-era dominance** across the full 18-year window.

CERTIFICATIONS

DataBricks: Generative AI Fundamentals