DHRUV PATEL

Atlanta, Georgia, United States · darup67@gmail.com
linkedin.com/in/darup89 · github.com/darup67 · darup67.github.io/dhruvpatel-site

AI/ML & DATA ENGINEERING · FORWARD DEPLOYED ENGINEERING

PROFILE

Hands-on AI/ML engineer who writes and ships LLM-agent and machine learning systems in Python, with 11+ years building production data and ML platforms in Fortune 100 environments, including Azure Machine Learning, Microsoft Foundry and Palantir Foundry. Most recently a Sr. Forward Deployed Engineer at Sherwin-Williams, where I built the Freight Rate, Batch Size Rate and demand-forecasting models behind pricing. I manage AI/ML delivery across the full software development lifecycle (scoping and data design, model development and evaluation, CI/CD and MLOps, release, and production monitoring), including leading teams of 25+ engineers and data scientists at T-Mobile. On the agentic AI side, outside work I design and operate a 15+ agent autonomous paper-trading and research system (no broker path) with Claude Code and MCP integrations, with scoped tool permissions, human-approval gates and kill switches, where every agent output is graded against statistical baselines before it is trusted: evaluation harnesses, guardrails and monitoring, written and debugged by hand.

Earlier work spans Generative AI and LLM solutions (RAG and vector-database patterns) and developer-productivity tools, MLOps and model monitoring, explainability and responsible AI (Global Payments), TensorFlow and spaCy NLP models (Compliance.AI), and event-driven data platforms (Kafka/Confluent, Spark, Databricks, Snowflake, ClickHouse) on AWS, Azure and GCP. Comfortable as the engineer embedded with business and external stakeholders, from first prototype through production.

CORE COMPETENCIES

- Large Language Models & Generative AI
- RAG & Vector Databases
- LLM Agents & Tool Use
- MCP Servers & Integrations
- LLM Evaluation, Guardrails & Permissions
- AI Coding Tools (Claude Code, GitHub Copilot, Cursor, OpenAI Codex)
- Machine Learning & NLP (spaCy, TensorFlow, AutoGluon)
- MLOps, CI/CD & Model Monitoring
- Explainability & Responsible AI
- Demand Forecasting
- Pricing & Freight Rate Modeling
- Forward Deployed Engineering
- Python
- SQL
- Scala / Spark
- Node.js
- Apache Airflow
- Apache Kafka / Confluent
- Databricks / Delta Lake
- Snowflake / ClickHouse
- Azure Machine Learning
- Microsoft Foundry (Azure AI)
- Palantir Foundry
- Palantir Data Integration
- Palantir Ontology & Semantic Layers
- Vertex AI / BigQuery ML
- AWS / Azure / GCP
- Terraform / Docker / Kubernetes
- Data Governance & Security (SOX, GDPR, Collibra, Purview, Azure RBAC)
- Informatica MDM / Reltio

HANDS-ON LLM & AGENT ENGINEERING

- **Autonomous trading & research ecosystem (15+ agents, running 24/7, paper-only):** Built and operate in Python and Node.js: a shared SQLite 15m bar store (Alpaca IEX + Yahoo fallback); signal agents (SuperTrend flip watcher on 75 symbols, 11-sector IV spread scanner, 102-coin crypto monitor, Kalshi BTC caller); research tools (paper-lab for BTC/ETH and 4 CME micros with no broker path, sector-breadth study, real-estate digest); all on launchd with a watchdog that repairs missed runs against a NYSE calendar. Source: github.com/darup67 (trade-core, market-lab, kalshi-btc-agent).
- **Statistical evidence gating:** Every signal kind (equity flip, IV spread, crypto pick, coin launch) is graded net of costs with Wilson lower bounds (≥30 signals, ≥10 days); only PROVEN kinds fire live alerts, the rest are logged and withheld.
- **Tool use over MCP:** Wired agents to MCP servers for brokerage (Robinhood Agentic MCP: read-only scopes, confirm-before-order rule), charting (TradingView 84-tool MCP), email and job-search data. Enforced permission boundaries in code: credentials kept in OS keychain, never in prompts or logs.
- **LLM output in shadow mode:** Integrated a hosted LLM headline-judging service to label topic, direction and materiality, storing each judgment beside the outcome so it is graded before it can influence any decision.
- **Models with guardrails:** Trained AutoGluon classifiers with held-out testing, minimum-AUC gates and kill switches that turn a model off when live results decay; diagnosed and fixed failures such as leaky cross-validation folds and stale features.
- **Text signals:** Built a sentiment pipeline over StockTwits tags, news, Reddit, CoinGecko and Telegram with a crypto lexicon and negation handling, logged each reading with the price, and tested it against what followed instead of assuming it predicts.

PROFESSIONAL EXPERIENCE

Sr. Forward Deployed Engineer | SHERWIN-WILLIAMS **Jun 2026 — Oct 2026**

SHORT-TERM CONTRACT

Forward Deployed Engineering

- **Forward Deployed Delivery:** Work directly with internal business stakeholders and external partners to turn business problems into deployed tools, technology and AI/ML-driven platforms that the business consumes directly and that support revenue generation.
- **Stakeholder Deliverables:** Gather requirements, define deliverables and drive them to production with business users, engineering and external stakeholders, and support adoption after release.

Pricing Structure & Demand Forecasting ML Models

- **Freight Rate Model (FRM):** Build and operate the Freight Rate Model, a machine learning model that estimates freight rates and supplies the freight component of the pricing structure, covering feature engineering, training, validation, deployment and ongoing performance monitoring.
- **Batch Size Rate (BSR) Model:** Build and operate the Batch Size Rate model, a machine learning model that sets rates by batch size and feeds the pricing structure, owning the full model lifecycle from data preparation through production release and monitoring.
- **Demand Forecasting:** Design, train and productionize machine learning models that forecast demand from historical data, supporting planning and inventory decisions and giving pricing a forward view of volume.
- **Pricing Structure:** Combine freight-rate, batch-size-rate and demand-forecast outputs so the pricing structure is built from data-driven, repeatable models instead of manual rate setting.

Ontology, Semantic & Data Layers; Azure ML & Microsoft Foundry

- **Palantir Ontology & Semantic Layers:** Model the pricing, freight and demand domain as ontology objects and links (products, plants, lanes, customers) over raw, curated and semantic data layers with lineage and access controls, so the models and the business share governed definitions.
- **Azure ML & Microsoft Foundry:** Train, register, deploy and monitor the pricing and forecasting models on Azure Machine Learning, and build and evaluate LLM and agent solutions in Microsoft Foundry alongside them.

Principal AI/ML Engineer | T-MOBILE **Jun 2022 — Dec 2025**

AI/ML Platform Engineering & Delivery

- **ML Delivery:** Led 25+ engineers and data scientists (budget over $15M) to architect and deploy machine learning solutions for telecommunications operations; AI/ML adoption across 15+ business units saved $2.3M a year through end-to-end automation, with models built, deployed and monitored on Azure Machine Learning and Microsoft Foundry.
- **Forward Deployed Delivery:** Worked directly with internal business units and external vendors and partners to scope, build and deliver tools, technology and AI/ML-driven platforms for business consumption and revenue generation, translating stakeholder needs into production deliverables.
- **CI/CD & MLOps:** Optimized deployment pipelines and integrated automated testing, achieving a 40% reduction in model deployment cycles while sustaining 99.8% system uptime.
- **Cloud Infrastructure:** Delivered trusted analytics via semantic data products and event-driven architectures across AWS and Microsoft Azure, leveraging GitOps and Infrastructure-as-Code (IaC) for scalable environment management.
- **High-Scale Microservices:** Managed the end-to-end lifecycle for MDM/RDM platforms through 12 major feature releases, supporting a microservices architecture that handled 10 million daily transactions for 50,000+ internal users.
- **Golden Record & Matching:** Built Golden Record solutions with fuzzy logic and machine learning matching algorithms consolidating 30+ sources: data quality up 35%, data reconciliation time down 60%, manual corrections down 45%.
- **Real-time Data Fabric:** Built integration patterns with Apache Kafka and REST APIs to process 500GB+ of daily data, improving data discovery 50% through automated lineage tracking.
- **Security & Migration:** Designed security architecture for the migration of 200+ applications and databases to Azure (assets over $100M); automated data classification and Privileged Identity Management cut data exposure incidents 80% and unauthorized access 95%. Put SOX and GDPR controls around enterprise ML and data platforms.

Data Integration (Palantir Foundry)

- **Palantir Foundry Integration:** Connected and interlaced 100+ disparate source systems and databases in Palantir Foundry, joining data across mismatched schemas, formats and system owners into one governed, queryable foundation.
- **Pipelines & Data Modeling:** Built and maintained the Foundry pipelines, transforms and shared data models that reconciled records across those sources.
- **Palantir Ontology & Semantic Layers:** Modeled the unified Foundry datasets as an ontology of business objects, properties, links and actions, with layered (raw, curated, semantic) data products, governed definitions, lineage and permissions, so analytics, ML and applications consume consistent semantic data.

Senior ML Data Engineer | GLOBAL PAYMENTS **Dec 2019 — May 2022**

Generative AI, MLOps & Responsible AI

- **Forward Deployed Delivery:** Partnered with internal teams and external stakeholders to define, build and deliver tools, technology and AI/ML-driven platforms for business consumption and revenue generation, from requirements through production adoption.
- **Generative AI / LLM Solutions:** Developed Generative AI and LLM-based solutions, including internal developer-productivity tools that streamline engineering workflows.
- **MLOps & Model Monitoring:** Implemented MLOps practices and model monitoring for production ML models, tracking model health and performance after release.
- **Explainability & Responsible AI:** Applied model explainability and responsible-AI practices so automated decisions could be reviewed, audited and trusted.

Data Platform, Cloud & Machine Learning

- **High-Scale Pipeline Design:** Architected ETL pipelines processing $50B+ annual transaction volumes using Python, Scala, and Apache Spark, migrating core workflows to GCP to improve system resiliency.

- **GCP Infrastructure Management:** Spearheaded the transition from AWS (S3, EMR) to GCP-native services, utilizing BigQuery and Cloud Storage to achieve a 45% reduction in query response times.
- **AI/ML Development (GCP):** Engineered end-to-end machine learning lifecycles within Vertex AI, deploying models for predictive analytics and automated decision-making.
- **Advanced Fraud Detection:** Enhanced real-time fraud detection systems to 98.5% accuracy and a 25% reduction in false positives by implementing BigQuery ML and streaming data transformations.
- **Cost Optimization & Migration:** Led a strategic migration initiative from legacy on-prem/AWS systems to GCP, realizing a 30% reduction in annual infrastructure overhead through resource rightsizing and BigQuery slot optimization.
- **LookerBI & Business Intelligence:** Orchestrated the integration between back-end infrastructure and BI systems, leveraging LookerBI to build automated reporting layers and real-time executive dashboards.
- **Data Mappings & Preprocessing:** Developed complex data mappings and schema definitions to align disparate financial datasets and optimized high-velocity data streams, preparing massive datasets for downstream AI/ML consumption while ensuring integrity across the enterprise data lake.
- **Delivery & Reliability:** Directed a cross-functional engineering team using Agile/Scrum, accelerating delivery 20% through sprint planning and CI/CD automation; scaled cloud-native back-end systems for high availability in mission-critical financial processing.

Master ML Data Engineer (MDM/RDM) | NOVELIS
Mar 2019 — Sep 2019

- **MDM/RDM Hub:** Implemented Informatica MDM hub for 15 global manufacturing sites, improving data quality from 78% to 96%.
- **Automation & Governance:** Developed Python/SQL workflows processing 2M+ records monthly; established governance processes and trained over 25 business users.

Data Engineer / ML Engineer | COMPLIANCE.AI
Sep 2017 — Feb 2018

- **NLP ML Models:** Developed TensorFlow and spaCy models for regulatory compliance, reaching 92% accuracy on document classification.
- **System Performance:** Built Airflow/Docker pipelines and REST APIs with 99.9% uptime and <100ms response times.

EDUCATION

Georgia Institute of Technology — BS in Computer/Information Technology Administration & Management — May 2019

University of Georgia — BBA in Actuarial Science, Risk Management & Insurance — May 2014

CERTIFICATIONS

- MIT Professional Certificate in ML & AI
- IBM AI Engineering Professional
- Google TensorFlow Developer
- DeepLearning.AI ML & Deep Learning Specializations
- AWS ML Specialty
- Azure AI Engineer Associate (AI-102)
- Google Cloud Professional Data Engineer