

Principal Machine Learning Engineer

Purpose of the role

This is the most senior engineering role in the group and the technical owner of the agent platform itself. Where the pairing Principal Data Scientist role asks “is the model doing the right thing,” this role asks “can we build, ship, run and trust this at scale, safely, and again next week.”

The successful candidate will set the technical direction for how agents and conventional ML models are built, deployed, observed and operated across the estate, and will be the escalation point when something in production needs deep infrastructure expertise rather than model expertise.

Key responsibilities

- Own the end-to-end architecture of the agent estate/foundry: build, deployment, monitoring, and reliability of agentic and conventional ML systems in production.
- Set standards and reusable patterns for CI/CD, environment management, and release processes for both traditional ML models and agentic services.
- Design and evolve the platform's use of MCP servers and A2A-based agent-to-agent communication so agents can be composed, discovered and orchestrated reliably.
- Establish production observability and evaluation practices for LLM-based and conventional ML systems (MLOps/LLMOps), including monitoring, alerting, drift detection and rollback strategies.
- Lead integration of new, greenfield agentic components with legacy services and existing data sources, balancing modern methods against operational constraints.
- Act as the senior technical authority on Azure (preferred) or AWS AI infrastructure decisions, including evaluation of Azure AI Foundry where relevant.
- Mentor the Senior Machine Learning Engineer and set technical direction across the wider engineering function, while pairing closely with the data science leadership on shared priorities.
- Represent engineering feasibility, cost and risk in architectural decisions taken jointly with the Principal Data Scientist / AI Engineer.

Essential experience

- 8+ years in machine learning or platform engineering with direct, hands-on ownership of production systems (not purely advisory or research-only experience).
- Proven experience building AI infrastructure on Azure (strongly preferred) or AWS at meaningful scale.
- Experience deploying and running conventional ML models alongside agentic/LLM-based systems side by side in the same estate.
- Track record of greenfield delivery using current methods, integrated with legacy services and data sources — not just building on a clean slate.
- Direct, practical experience with the realities of CI/CD in an ML/AI context (not just familiarity with the concept).

Essential skills

- Python at a senior/architect level.
- Docker and/or Podman for containerisation.

- Terraform and CI/CD pipeline design and ownership.
- SQL and database management.
- MCP server development and deployment.
- A2A (Agent2Agent protocol) — hard requirement, not negotiable.
- MLOps and LLMOps, including production observability and evaluation of ML/LLM systems.
- Experience with at least one agent framework or SDK (Claude Agent SDK experience is a bonus, but any comparable framework is acceptable).

Desirable / nice-to-have

- Azure AI Foundry experience.
- RAG systems and vector stores (pgvector, Pinecone).
- ETL and DAG orchestration tooling.
- Serverless architectures.

Senior Machine Learning Engineer

Purpose of the role

This role has the same remit as the Principal Machine Learning Engineer — the engineering backbone of the agent estate — but at a contributor rather than owner level. The successful candidate will implement and help evolve the architecture set by the principal role, with a growth path toward taking on more ownership over time.

Key responsibilities

- Build and maintain components of the agent estate/foundry: deployment pipelines, monitoring, and reliability tooling for agentic and conventional ML systems.
- Implement CI/CD pipelines and infrastructure-as-code for ML and agentic services under architectural direction from the Principal MLE.
- Develop and deploy MCP servers and integrate A2A-based agent communication into existing services.
- Deploy conventional ML models to production and support their ongoing operation alongside newer agentic components.
- Contribute to greenfield build-outs that integrate with legacy systems and data sources, flagging integration risk early.
- Support production observability and evaluation practices (MLOps/LLMOps) — monitoring, logging, alerting — for models and agents in the estate.
- Work day-to-day alongside the data science group (Senior Data Scientist / Senior AI Engineer) to turn model and agent designs into deployed, reliable services.

Essential experience

- 4+ years in machine learning or platform engineering.
- Experience building AI infrastructure on Azure (preferred) or AWS.
- Direct experience deploying ML models to production (not just training/experimentation).
- Some exposure to greenfield development integrated with legacy systems.
- Familiarity with the practical challenges of CI/CD in an ML context.

Essential skills

- Python.
- Docker and/or Podman.
- Terraform and CI/CD.
- SQL.
- MCP server deployment.
- A2A (Agent2Agent protocol) — hard requirement.
- Working knowledge of MLOps and LLMOps.

Desirable / nice-to-have

- Azure AI Foundry exposure.
- RAG systems and vector stores (pgvector, Pinecone).

- ETL and DAG orchestration.
- Serverless architectures.
- Claude Agent SDK or another agent framework/SDK.