

ML Platform and MLOps

Tracking, versioning, registries, canary rollouts, and drift monitoring. The skill is noticing a model has quietly gone wrong in production before a user has to tell you.

Intermediate

5 phases

26 sessions

39 contact hours

Syllabus

5 phases, 26 sessions. Each phase ends in something you have built.

1

Reproducible or It Did Not Happen

Make any training run re-creatable by someone else, months later, without you in the room.

5 sessions

\$72 alone

1.1 What MLOps Actually Owns

The boundary: the engineering around a model, not the model, and who gets paged when it rots.

[FREE PREVIEW](#)

1.2 Tracking Runs in MLflow

Runs, params, metrics, artifacts, and the LoggedModel entity that ties code to a deployed thing.

1.3 Versioning Data, Not Copying It

DVC pointers, lakeFS branches, or an Iceberg snapshot ID: pick one, record it, stop duplicating.

1.4 The Reproducibility Tuple

Code SHA, data snapshot, config, environment digest, artifact digest, written to one manifest.

1.5 Environments That Rebuild

Lockfiles, container digests, and the CUDA mismatch that makes a rerun quietly different.

By the end: **A run manifest and one command that rebuilds the same model from the same data.**

2

From Notebook to Pipeline

Turn a training script into scheduled tasks that retrain without introducing skew.

5 sessions

\$72 alone

2.1 Choosing an Orchestrator

Airflow 3.3 assets against Dagster and Prefect, and what the July 2026 merger changes for you.

2.2 The Training Pipeline

Ingest, features, train, evaluate, register: five retrievable tasks instead of one long notebook.

2.3 Training and Serving Skew

One transform, two call sites, and the future timestamp that inflated your offline score.

2.4 Point-in-Time Joins

As-of joins in SQL over lakehouse tables, proving no row saw a value it could not have had.

2.5 Do You Need Feast?

Feast 0.64 when you need millisecond online lookups, and the batch cases that never justified it.

By the end: **A scheduled retraining pipeline with point-in-time correct features and no leakage.**

3

The Registry and the Gate

Stop a bad model from reaching users, and roll one back without a retraining run.

5 sessions

\$72 alone

3.1 Aliases, Not Stages

Immutable versions, champion and challenger aliases, and why MLflow stages died in 2.9.

3.2 The Promotion Gate

Absolute floors plus non-regression against the incumbent, on a frozen holdout and a live slice.

3.3 CI for Models

GitHub Actions with OIDC, actions pinned by SHA, and a protected environment that needs a human.

3.4 Packaging for Serving

KServe, BentoML, or Ray Serve, and deploying a container digest rather than a branch name.

3.5 Shadow, Canary, Blue-Green

Mirror traffic first, ramp one to five percent, and roll back by moving an alias, not retraining.

By the end: **A promotion pipeline that registers, gates, canaries, and reverts in one step.**

4

Operating a Model

Notice a model degrading before the business does, then say what it cost.

5 sessions

\$72 alone

4.1 Log Every Prediction

One prediction ID joining input, model version, cohort, output, and the label that lands later.

4.2 Performance First, Drift Second

The alert hierarchy, and why a KS test crossing 0.05 is a dashboard entry, not a page.

4.3 When Labels Arrive Late

NannyML CBPE and DLE estimating accuracy with no targets yet, then reconciling when labels land.

4.4 The Degradation Drill

A model rots quietly, you get paged, and you separate data, feature, infra, and logic causes.

4.5 Cost Per Prediction

Cost per thousand predictions and per successful outcome, GPU idle time, and a budget alert.

By the end: **Prediction logs, an alert hierarchy, a runbook, and a postmortem from a live drill.**

5

Someone Else's Model

Run a feature on a vendor LLM API where you own everything except the weights.

6 sessions

\$72 alone

5.1 Pin the Whole Release

Model snapshot, prompt version, index, tools, and decoding parameters in one release manifest.

5.2 Traces as the Unit

OpenTelemetry GenAI spans for tokens, cost, and latency, with attributes still marked Development.

5.3 Prompt Registry and Rollback

Versioned prompts with aliases in MLflow, and diffing two versions over the same replayed traces.

5.4 Evals That Gate Merges

Promote real production failures into a versioned suite that blocks the merge when it regresses.

5.5 The Forced Migration

Sixty days of notice, incumbent against successor on your suite, canary, then move the pin.

5.6 Token Cost, Reconciled

A price table with effective dates, cost per successful task, and telemetry matched to the invoice.

By the end: **A pinned LLM release with a trace pipeline, a regression suite, and a migration plan.**

Tools you will use

MLflow 3.15

Apache Airflow 3.3

DVC 3.67

Apache Iceberg 1.11

Feast 0.64

Evidently 0.7

NannyML 0.13

KServe 0.18

GitHub Actions

OpenTelemetry GenAI

What you will build

1 Reproduce It in Six Months

One manifest, one command, the same model back

2 The Gate That Says No

A promotion run blocked by a cohort regression

3 The Quiet Degradation

Paged, triaged, rolled back, and written up

Registration

Phase 1: Reproducible or It Did Not Happen	\$72
--	------

Phase 2: From Notebook to Pipeline	\$72
------------------------------------	------

Phase 3: The Registry and the Gate	\$72
------------------------------------	------

Phase 4: Operating a Model	\$72
----------------------------	------

Phase 5: Someone Else's Model	\$72
-------------------------------	------

Whole track, all 5 phases	\$300
----------------------------------	--------------

Buying every phase separately costs \$360, so the whole track saves \$60.

No payment is taken online. Registering creates an order and payment instructions are sent with it. Access opens once payment is confirmed.